Thèse n° 9807

Computational Models of Mutual Understanding for Human-Robot Collaborative Learning

Présentée le 16 mai 2023

Faculté informatique et communications
Laboratoire d'ergonomie éducative
Programme doctoral en informatique et communications

pour l'obtention du grade de Docteur ès Sciences

par

Utku NORMAN

Acceptée sur proposition du jury

Prof. R. West, président du jury
Prof. P. Dillenbourg, directeur de thèse
Prof. O. Çeliktutan, rapporteuse
Dr S. Lemaignan, rapporteur
Prof. F. Kaplan, rapporteur

École
polytechnique
fédérale
de Lausanne

2023

For everything in the world,
for civilization, for life, for success,
the truest guide is science;
seeking a guide other than science is
heedlessness, ignorance, and fallacious.
— Mustafa Kemal Atatürk

To my one and only Gülseren, and beloved Mom and Dad

Acknowledgements

First and foremost, I am deeply grateful to my thesis supervisor, **Pierre Dillenbourg**, for mentoring me, and providing me with the freedom to explore and pursue different directions while keeping a focus, that made the road to this thesis a precious and delightful experience.

I would like to express my gratitude to the members of my thesis jury, **Robert West**, **Frédéric Kaplan**, **Oya Çeliktutan**, and **Séverin Lemaignan**, for taking the time to evaluate this work, and giving their insightful comments that helped improve the final manuscript of this thesis.

I am indebted to **Barbara Bruno** for her continuous support and rigorous constructive feedback, without whom this thesis would not be completed within a reasonable time frame. It has been an absolute pleasure working with her; I had the opportunity to learn so much from her.

This research was possible thanks to the European Union's Horizon 2020 ANIMATAS Project[I], in which I had the pleasure of creating wonderful memories with **Rebecca Stower**, **Sooraj Krishna**, **Natalia Calvo**, **Karen Tatarian**, **Sahba Zojaji**, **Sebastian Wallkötter**, **Manuel Bied**, **Ramona Merhej**, **Silvia Tulli**, and **Maha El Garf**. Special thanks to **Sera Büyükgöz** that could brighten any day, and **Tanvi Dinkar** with whom I was lucky to work with and enjoy endless discussions: they made me and Gülseren feel at home in Paris. I thank **Chloé Clavel** for supporting our collaboration in my secondment and onwards, **Arvid Kappas**, **Ginevra Castellano** and **Kerstin Dautenhahn** for their thoughtful feedback in my mid-term review, and **Mohamed Chetouani** for hosting me in Paris, and initiating and upholding this spectacular network.

I would like to thank international schools in Switzerland that opened their doors for our studies: I am grateful to **Melissa Skewers** at EPFL, **Paul Magnuson**, **Kearon McNicol**, **Adrian Hirst**, **Felicia de Lucia**, **Jessica O'Neill Casas**, **Jonathan Snell**, **Danielle Allard**, **Deirdre Canavan**, **Hannah Saks**, and **Marie-France Labelle** for all their efforts that enabled this research.

I have been fortunate to become a part of the CHILI lab, and share a wonderful experience together with and thanks to **Jauwairia Nasir**, my companion in CHILI and at schools wherever they may be culminating in a unique set of journeys together, **Sina Shahmoradi** of fun and indomitable spirit, **Arzu Güneysu Özgür** and **Ayberk Özgür** for their big hearts and being with me and Gülseren whenever we needed it including but not limited to rakı and King nights and

[I] This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 765955 (ANIMATAS Project).

barbecues, **Wafa Johal** for taking me on to CHILI and onboarding me to the team, **Hala Khodr**, **Lucas Burget**, **Daniel Tozadore**, **Victor Borja**, **Richard Lee Davis**, **Chenyang Wang**, **Zhenyu Cai**, **Yi-Shiun Wu**, **Jie Gao**, **Kevin Gonyop Kim** (and the cutest **Yuna Kim**), **Aditi Kothiyal**, **Thibault Asselborn**, **Ramtin Yazdanian**, **Louis Faucon**, **Elmira Yadollahi**, **Sven Viquerat**, **Killian Viquerat**, **Teresa Yeo**, **Laila El-Hamamsy**, **Sruti Bhattacharjee**, **Su Xiaotian**, **Catharine Oertel**, **Jennifer Olsen**, **Jade Cock**, **Oğuzhan Fatih Kar**, **Stian Håklev**, **Soheil Kianzad**, **Giorgia Marchesi**, **Thomas Gargot**, **Mortadha Abderrahim**, **Melike Cezayirlioğlu**, **Dorsa Safaei**, **Jules Cortois**, **Corinne Lebourgeois**, **Anthony Peguet**, **Laurent Boatto**, **Patrick Jermann**, and **Florence Colomb**–the real boss of the lab.

I would like to thank our students for their enthusiasm in together exploring various directions of research and development: our summer intern **Alexandra Chin** who was generously sponsored by a Summer@EPFL fellowship, and semester project students **Romain Maure**, **Erik A. Wengle**, **Yiwen Ma**, **Elisa Bianchi**, **Yuanyuan Zheng**, **Anne Donnet**, and **Laura Mathex**.

I would like to thank friends I have been lucky to meet in Switzerland that made this journey an amazing experience: **Ezgi Yücetürk**, **Tuğrulcan Elmas**, **Onur Yürüten**, **Natalie Yéghen-Yürüten**, and their joyful **Omnia Evren Yürüten**, **Gyulay Dzhilen**, **Ali Ozan Kaynak**, **Kaan Aşıkoğlu**, **Elif Vardar**, **Ezgi Gürbüz**, **Ekin Emeksiz**, **Ayyüce Yeniçeri**, **Mustafa Mete**, **Elif Özen**, **Ahmet Alparslan Aktaş**, **Kyril Kaufmann**, **Alp Yurtsever**, **Serdar Cüneyt Akyüz**, **Tuğçe Akyüz**, **Eda Bayram**, **Sinem Sav** since Bilkent, and **Anıl Tuncel** for many nights filled with music.

I would like to express my gratitude to my friends that have supported me throughout: **Deniz Sargun**, **Onur Memioğlu**, **Gökhan Uruk**, **Mehmet Efe Tiryaki**, **Çağrı Şakiroğulları**, **Berker Peköz**, **Kaan Sel**, **Metin Dündar Özkan**, **Zeynep Büyükşalvarcı**, and **Serhan Yılmaz**. I would like to extend my thanks to **Zerrin Celebi Ghavami** for guiding us in Switzerland, **A. Ercüment Çiçek** for showing me the ways of academia, **Melike Batgıray Abboud**, **Mohammad Abboud**, and **Hakan Kırımlı** for sharing his love for gastronomy and wisdom in life.

I sincerely thank my parents **Serpil Norman** and **Yaşar Norman**, without whom I would have never reached this point, and my sister **Özge Norman** for keeping springing surprises on me. Thank you **Simi** for being with me and Gülseren through the difficult times, with COVID-19 etc., giving all of your love before leaving this world. Last, but certainly not least, I thank my dear wife, my companion in life, my soul, **Gülseren Norman**: thank you for being there for me.

Lausanne, 20 November 2022 U. N.

Abstract

There is a growing trend towards designing learning activities featuring robots as collaborative exercises where children work together to achieve the activity objectives, generating interactions that can trigger learning processes. Witnessing such activities allows for making an interesting observation: humans, unlike robots, are highly skilled in detecting and addressing misunderstandings, building a mutual understanding about the task, and converging to a shared solution. A social robot equipped with these abilities can monitor the interaction and contribute to it, promoting and supporting the building of a mutual understanding, which may trigger learning. To verify this hypothesis, in this thesis, we first develop abilities for social robots to assess how humans build a mutual understanding. To this end, we (i) propose automatic measures to reveal structures in how children "align" in their dialogue and actions, when engaging in a collaborative activity aiming to foster their computational thinking skills and (ii) study how these lead to their performance in the task and learning outcomes; on data we collected in a large user study involving 78 children at schools for a collaborative activity we designed that elicits dialogue. Then, we equip the robot with mutual modeling abilities to build a mutual understanding with a learner: we (iii) present a framework for the robot to build and maintain a mental model, with its own beliefs about the activity, the human, as well as the human's beliefs about the robot; and (iv) evaluate the effects of robot behaviors that are guided by different mental models via an experiment with 61 children at schools, in which a child and the robot collaborate on a variant of the problem solving activity used throughout the thesis.

Keywords: mutual understanding, socially interactive artificial intelligence, mutual modeling, dialogue alignment, dialogue grounding, situated dialogue processing, computational thinking, robots for learning, collaborative learning

Résumé

On constate une tendance croissante à concevoir les activités d'apprentissage faisant intervenir des robots comme des exercices collaboratifs où les enfants travaillent ensemble pour atteindre les objectifs de l'activité, générant des interactions qui peuvent déclencher des processus d'apprentissage. Le fait d'assister à de telles activités permet de faire une observation intéressante : les humains, contrairement aux robots, sont très compétents pour détecter et traiter les malentendus, pour établir une compréhension mutuelle de la tâche et pour converger vers une solution partagée. Un robot social doté de ces capacités peut suivre l'interaction et y contribuer, favorisant et soutenant la construction d'une compréhension mutuelle, ce qui peut déclencher l'apprentissage. Pour vérifier cette hypothèse, dans cette thèse, nous développons d'abord des capacités pour les robots sociaux à évaluer comment les humains construisent une compréhension mutuelle. À cette fin, nous (i) proposons des mesures automatiques pour montrer la structure sur la façon dont les enfants "s'alignent" dans leur dialogue et leurs actions, lorsqu'ils s'engagent dans une activité collaborative visant à stimuler leurs compétences en pensée computationnelle et (ii) nous étudions comment celles-ci mènent à leur performance dans la tâche et aux résultats de l'apprentissage ; sur des données que nous avons recueillies dans une grande étude d'utilisateurs impliquant 78 enfants dans des écoles pour une activité collaborative que nous avons conçue et qui suscite le dialogue. Ensuite, nous dotons le robot de capacités de modélisation mutuelle pour construire une compréhension mutuelle avec un apprenant : nous (iii) présentons un cadre permettant au robot de construire et de maintenir un modèle mental, avec ses propres croyances sur l'activité, l'humain, ainsi que les croyances de l'humain sur le robot ; et (iv) nous évaluons les effets des comportements du robot qui sont guidés par différents modèles mentaux via une expérience avec 61 enfants dans des écoles dans laquelle un enfant et le robot collaborent sur une variante de l'activité de résolution de problèmes utilisée tout au long de la thèse.

Mots clefs : compréhension mutuelle, intelligence artificielle socialement interactive, modélisation mutuelle, alignement du dialogue, ancrage du dialogue, traitement du dialogue situé, pensée computationnelle, robots pour l'apprentissage, apprentissage collaboratif

Contents

List of Figures

List of Tables

List of Algorithms

List of Equations

List of Abbreviations

1 Introduction

1.1 Motivation

Learning activities are often designed as collaborative exercises where two or more children work together to achieve the activity objectives, generating interactions that can trigger learning processes. Witnessing such activities allows for making an interesting observation: although miscommunications and misunderstandings occur frequently, humans, even at young ages, are very good at overcoming them. Humans, *unlike robots*, are highly skilled in detecting and addressing misunderstandings, building a *mutual understanding* about the task at hand, and converging to a shared solution. Humans represent whether the other understood what they said or did by using the complex cognitive abilities of *mutual modeling*[I], i.e. a set of reciprocal abilities to build a *mental model* of the other: humans perform this by attributing beliefs, desires and other mental states to the other (Premack & Woodruff, 1978). These abilities are critical for humans to comprehend each other and react appropriately in their interactions: therefore, these skills are also fundamental for artificial agents (and therefore robots) to interact effectively with humans (Kopp & Krämer, 2021).

Social robots have a potential to become a part of educational environments by undertaking a unique position that extends their functional purpose to support human learning with personal and social dimensions of interaction (Belpaeme et al., 2018). In this thesis, we are interested in developing mutual modeling abilities for humanoid social robots as a way to leverage the social dimension: so that they can assess the mutual understanding between learners, and build a mutual understanding with a learner in order to support human learning. Thus, our ultimate research objective is to equip a robot with mutual modeling abilities in an educational scenario and investigate its potential benefits on the quality of the interactions

[I] Similar terms include: *Theory of Mind (ToM)* (Premack and Woodruff, 1978; Wimmer and Perner, 1983; Baron-Cohen et al., 1985); as a skill for a robot by e.g. Scassellati, 2002); as well as adopting the *intentional stance* (Dennett, 1987), that we may adopt also towards artificial agents by *mentalizing* (e.g. in C. L. Baker et al. 2017) or *social attunement* (Perez-Osorio & Wykowska, 2019), and *social cognition*. For instance, social cognition can be preferred over ToM, as the latter may bring in undesired theoretical commitments (Verbrugge, 2009). We prefer mutual modeling as it highlights the mutuality and modeling aspects; while linking work in collaborative learning as in Dillenbourg (1999), to research in Human-Robot Interaction (HRI) as in Lemaignan and Dillenbourg (2015).

and the learning outcomes.

1.2 Background

1.2.1 Collaborative Learning as the Effort to Build a Mutual Understanding

Collaboration occurs in a situation in which individuals work together as members of a group in order to solve a problem (Roschelle & Teasley, 1995). In a collaborative activity, members build shared, abstract representations; a *mutual understanding* that is shared in common among the members and is about the problem at hand (Schwartz, 1995). When a collaborative activity is in an educational setting, it is termed as *collaborative learning*; where individuals *learn* together. While there is no guarantee that the interactions between learners will induce learning (Dillenbourg, 1999), they can be *designed* to make it more likely that these learning processes will in fact occur, as they are shaped by the activity and the environment: it is the shared experience and the required *effort* to construct a mutual understanding together that may trigger the learning (Dillenbourg et al., 2009; A. King, 2007). The hypothesis that this thesis seeks to verify is that a robot equipped with mutual modeling abilities can act as an adaptive, intelligent component to complement a designed collaborative learning activity environment, and help trigger learning mechanisms by supporting this effort to build a mutual understanding.

Theories of individual cognitive development, on how we acquire knowledge as well as think about and understand the world, could give insight into the learning processes that can occur in collaborative learning situations. Linking cognitive development with *social interaction* goes back to the works by Mead (1934), Vygotsky (1930–1934/1978) and Piaget (1970/1971). The causal direction of cognitive development for socio-cultural theories based on Vygotsky's ideas is "from outside in": it is a result of social and cultural experience, that is socially guided and constructed through interactions between individuals (Gauvain, 2020). This takes the form "what children can do with others today, they can do alone tomorrow" (Vygotsky, 1934/1986). On the contrary, for Piaget, it proceeds with the child adapting and hence it is more "from the inside out" (DeVries, 2000). Alternatively, the causality may rather be bi-directional, or reciprocal: development "progresses in both a circular and spiral fashion", with the social interaction improving the individual, who can in return take part in more sophisticated interactions, and so on (Doise & Mugny, 1981/1984). Thus, an account for further development is *socio-cognitive conflict* (Doise & Mugny, 1981/1984; Mugny & Doise, 1978), and its *regulation* (Butera et al., 2019). One of the mechanisms that can bring about development in a conflict is *appropriation*, which allows the learner to consider the experienced social interaction to reexamine his/her own understanding (Rogoff, 1990). In our opinion, a robot with mutual modeling abilities can deliberately create a socio-cognitive conflict, so that the learner can be nudged towards the effort of appropriating the interaction, to the end of building a mutual understanding.

1.2.2 Grounding and Alignment in Dialogue as How We Understand Each Other

A major perspective for equipping robots with mutual modeling abilities builds on research in *psycholinguistics*, i.e. the psychology of language (Lemaignan & Dillenbourg, 2015). These studies do not focus on educational goals or how humans learn, but on how humans understand each other by analyzing human-human communication. They examine the *dialogue*, i.e. the series of utterances humans produce as they communicate with each other. Similar to understanding group processes in collaborative learning, studying a dialogue is challenging; as rather than an individual effort, a dialogue is a *joint activity* between two or more people, i.e. the *interlocutors* (Clark & Wilkes-Gibbs, 1986). Interlocutors ideally take turns "in an orderly way" to try to reach a mutual understanding (Clark & Schaefer, 1989). Clark and Schaefer refer to this collective process as *grounding*: interlocutors "add" to their *common ground*, i.e. what is already mutually understood, by making *contributions* that are *presented* by an interlocutor, and *accepted* by the other through some *evidence of understanding* in response.

Instead of grounding that "involves a good deal of modeling of one's interlocutor's mental state", Pickering and Garrod (2004) argue for an automatic *priming* mechanism as the basis of successful communication: when a listener encounters an utterance from the speaker, it is more likely that subsequently the listener will produce an utterance by using concepts and terms first used by the speaker. Here, the listener need not explicitly assent to the speaker's contribution as in grounding; instead, there can simply be an *implicit common ground* without explicitly modeling each other's beliefs, unless a misunderstanding requires repairing the representation. Like this, *routines*, i.e. expressions that are "fixed", become shared or *established* among interlocutors. Thus rather than grounding, they define *alignment* of representations, to mean the "development of similar representations in the interlocutors" (Pickering & Garrod, 2006). Therefore, the interlocutors succeed in understanding each other when there is *alignment* between them, or shared representations, at different linguistic levels. Pickering and Garrod (2004, 2006) in a cooperative maze game found that at first glance, the dialogue seemed unstructured, but then found underlying alignment structures that supported collaboration among the interlocutors: they argued that this is what makes the dialogue and hence a collaboration successful.

Works that study mutual understanding largely consider dialogue among adult interlocutors solving a problem together, e.g. in Fusaroli and Tylén (2016) and Garrod and Anderson (1987), without the added dimension of a learning goal. However, task performance is not always positively correlated with learning outcomes, as a student can fail in the task but learn from it (or even learn from failure as in the *productive failure* hypothesis by Kapur, 2008) or perform well in the task but not learn from it, as unproductive performers (Kuhn, 2015). Thus, works on collaborative learning study behaviors related to learning outcomes such as gaze and gestures, but can lack in-depth dialogue analysis. At the same time, several works on dialogue, particularly on mutual understanding, have extensively studied the dialogue, but not always with the added depth of learning outcomes. In Part I of this thesis, we bring together the two perspectives, to contribute to the exploration of the deep, complex and tangled relationship between what humans say and what humans do, and the outcomes of this in a learning context. For this purpose, in Chapter 3, we design a human-human collaborative learning activity mediated by a robot, in which what the participants say and do is not only tied to how they perform in the task, but also what they ultimately learn from the activity. In Chapter 4 and Chapter 5, we study how children are engaged in this activity by examining the data we collected at schools: we propose *automatic* measures that a robot can use to reveal structures in dialogue and actions on how humans build a mutual understanding, by assessing the *alignment* between the children resulting from the activity.

1.2.3 Acting in Dialogue to Enable a Robot to Build a Mutual Understanding

An artificial agent (and hence a robot) needs to process a dialogue and determine its "state" on-line to be able to track and perform grounding, and thus build a mutual understanding. Yet, the aforementioned *contributions model* of grounding by Clark and Schaefer (1989) is not very suitable for on-line processing, as it is hard to determine if a presentation is completed without observing the next utterance or action (Traum, 1999). Furthermore, the notion of grounding does not immediately apply to collaborative learning, because what is "shared" is at a different scale (Dillenbourg & Traum, 2006).[II] Dillenbourg and Traum note that psycholinguistics considers short dialogue episodes of seconds to minutes in a conversation, where the goal is to ground a single referent or a piece of information. They contrast this with collaborative learning, where there are longer episodes that can take hours in a problem-solving task, to co-construct concepts or laws about a domain: the goal is rather to trigger *conceptual change*, where the learners move from no conceptions or misconceptions to correct conceptions. Thus, for an agent to effectively support collaborative learning, a computational approach that can process a dialogue on-line and account for the different scales of grounding is needed.

As a computational model of grounding that can be used by an agent to track and perform grounding, Traum and Allen (1992) characterize grounding in terms of actions performed by the utterances of the interlocutors. The model serves as a *speech act theory* for grounding, where *speech acts* go back to Austin (1962) who observed utterances as performing actions. In this model, it is the performance of *grounding acts* like initiate, repair and acknowledge actions that leads to grounding: this model was directly used as part of a collaborative dialogue agent in Allen et al. (1995). Traum and Hinkelman (1992) extend this model to a *conversation acts model* of grounding, by generalizing speech acts to dialogue with a multi-level theory of action in dialogue that has four levels. At the lowest level, there are *turn-taking acts* like take-turn, keep-turn and release-turn that can be sublexical. At the second level are the grounding acts. At the next level, there are *core speech acts* such as inform, suggest, accept and reject actions that

[II] Similarly for *alignment*, as the complexity of what is being co-constructed for collaborative learning (concepts or laws) tends to be at a larger scale. Pickering and Garrod (2004, 2006) consider alignment of *linguistic representations* such as syntactic structure, meaning, and pragmatic use: less abstract than a concept. They consider alignment of *situation models*, i.e. mental representations of the particular state of affairs (Zwaan & Radvansky, 1998): more complex than linguistic representations, but maybe less than concepts or laws. The different scales for Dillenbourg and Traum (2006) are rather continuous, and separations arise by the focus of different theories.

provide the content to be grounded to the grounding acts: for instance, given a suggestion, the grounding acts initiate and acknowledge would be about the presentation and acceptance of the content that is being suggested. Finally, at the highest level, there are *argumentation acts* like elaborate, clarify and convince actions that can extend to the whole dialogue. This is a level closer to the scale of analysis for collaborative learning, where elaborations and clarifications contribute to e.g. conflict regulation that may lead to learning. Thus, the conversation acts model presents a suitable computational approach a robot can employ to perform grounding and hence build a mutual understanding in an educational scenario, by bridging the scales of grounding referents and co-constructing concepts.

In Part II of this thesis, for a robot to manifest its mutual modeling abilities, we develop dialogic elements that are inspired by the conversation acts model of grounding. We focus on the level of core speech acts, and the content to be grounded as the joint decisions that build up a problem solving process.[III] We operationalize a set of core speech acts as the unit (dialogic) elements of the interaction, such as suggest action for a decision, accept or reject actions for a suggested decision by the other. With these elements, the robot is able to work together with a human learner to construct a shared solution to a problem. As a test bench, in Chapter 6, we adapt the human-human collaborative learning activity from the first part to a human-robot version, in which the robot makes suggestions on what to do, that may be in line with what the human thinks or not. A joint decision is made only if it is suggested by one and agreed by the other. Thus, the human and the robot need to build a mutual understanding on what a correct solution is, and converge to a shared solution that they agree upon in order to solve the problem. We achieve this by incorporating the suggest, accept and reject actions into the activity as physical actions, as affordances available to the learner and the robot: a robot equipped with mutual modeling abilities interacts with a learner primarily through the activity in order to build a mutual understanding and converge to a shared solution, by taking actions and keeping track of actions taken by the learner.[IV] This activity and interaction mechanism allows us to compare the effects of different mental models that can be maintained by the robot, and investigate the potential benefits of the mutual modeling abilities for the robot on the interaction and the learning outcomes.

[III] This level of description in effect describes the interaction at a scale between that of psycholinguistics and collaborative learning.

[IV] We complement the robot's physical actions with speech. From the human side, we did not focus on the recognition of human's verbal input given the limitations of the state-of-the-art tools for natural language understanding that can be used by a robot (Marge et al., 2022). Indeed, another approach would be to primarily focus on spoken dialogue: for instance, a model to ground the intentions of a human user, by incrementally processing the dialogue acts that are detected from a voice interface, has been developed and evaluated on a pick-and-place robot (Hough & Schlangen, 2017a, 2017b). Instead, we focus on an action understanding approach, where dialogue acts are incorporated into the task itself as physical task actions, so that certain aspects of the social interaction happen through the task. With this simplification, we can specifically investigate the effects of the robot having different mental models (i.e. the mutual modeling abilities) on the interaction at a higher, decision-making level.

1.2.4 Modeling Other Agents to Build a Mutual Understanding

Humans have a strong tendency to attribute social meanings to objects, and *anthropomorphize* by treating them as living objects, even for geometric shapes that move around as in the classic animation experiment by Heider and Simmel (1944). To understand an object's behavior, Dennett (1971, 1987) distinguishes between three basic stances or strategies, that we adopt to *predict* its behavior and *explain* why it does what it does. By adopting the *physical stance*, for example we predict that a stone will fall to the ground if we release it, by its physical constitution and laws of physics. In the *design stance*, we predict that an alarm clock will sound at set time, by ignoring details about its physical constitution, and assuming it is designed that way and that it will behave according to its design under various circumstances. From the *intentional stance*[V], we treat the object as an *intentional system*; as a *rational agent* with beliefs, desires, and other mental states to do what it ought to do given those beliefs and desires. An alarm clock is too simple that this fanciful anthropomorphism is unnecessary: no further *predictive power* is gained by adopting the intentional stance and e.g. attributing it a “belief" about current time, or a “desire" to sound at set time. Meanwhile, towards people, the intentional stance often is the only practical strategy. It is also intuitive and very practical for us to consider a chess-playing computer as a rational agent who “wants" to win, who “knows" the rules etc. Artificial agents, such as humanoid robots, can be designed to induce our adoption of the intentional stance towards them, which might facilitate our interaction with them (Perez-Osorio & Wykowska, 2020; Schellen & Wykowska, 2019). The context and the robot's behavior might contribute to our adoption of the intentional stance over the design stance (Marchesi et al., 2019); while our motivation to e.g. interact effectively would determine if we are more or less likely to have these anthropomorphic representations (Epley et al., 2007).

When we treat agents as intentional systems, we observe their actions not simply as motions, but rather as goal-directed actions with an *intention*; as part of their plan to attain their *desires* given their *beliefs*, or acting to revise their beliefs through *perception* so that they can behave more effectively: the agent is *rational* if the action and the mental states are mutually consistent (Shoham, 1993). *Mentalizing*, i.e. mental state *inference*, is a key skill for social cognition, where an agent describes and predicts another agent's observable actions, by adopting the intentional stance and referring to unobservable mental states such as beliefs and desires that the modeling agent ascribes to the other agent. Mentalizing develops into a richer *Theory of Mind (ToM)* as a skill to attribute mental states to agents, where the agent understands that the other has a different perspective (Baron-Cohen, 1995). In order to find a computational basis for how humans do and an agent can perform mental state inferences,

[V]To be *intentional* means to be about, directed at or referring to an object as the content: this “contentfulness" (Dennett, 1988) is indicated by Brentano (1874/2009) and others as *intentionality*. Brentano argued that all and only mental states have intentionality: it is a “characteristic exclusively of mental phenomena. No physical phenomenon exhibits anything like it" (Brentano, 1874/2009). Specific types of sensations and emotions have been given as counterexamples to reject this thesis, e.g. by stating “(m)y beliefs and desires must always be about something. But my nervousness and undirected anxiety need not in that way be about anything." (Searle, 1983). Yet, for instance emotions might allow an intentional characterization (Crane, 1998). No matter if intentionality is constitutive for all mental states, in this thesis, we focus on *beliefs* that are mostly accepted to be intentional.

works by C. L. Baker et al. (2017) and Bartlett et al. (2019) and so on map the actions of an agent back to mental states that cause them. They explain the inference of an agent's beliefs, desires and perceptions by observing how it behaves in the state of the *world* that the agent is situated in. In this thesis, for the robot to maintain its mental model about the human, we also map the observed actions to mental states about the world: the human's actions gives cues to the robot on which the robot can make assumptions and form beliefs.

A family of computational methods for recursive reasoning in which agents can model other agents, as surveyed by Albrecht and Stone (2018), builds on *Markov Decision Processes (MDPs)*: MDP is a general formal framework for sequential decision making under uncertainty that has been studied since the 1950s (Bellman, 1957). It has been generalized to partially-observable contexts (Partially-Observable Markov Decision Processes (POMDPs) (Kaelbling et al., 1998)), that could be used by a robot to represent the beliefs of a learner e.g. about the activity, since they are not directly observable by the robot. The framework has also been extended to multi-agent contexts (e.g. Interactive-POMDPs (I-POMDPs) (P. Gmytrasiewicz & Doshi, 2005)) with which a robot can explicitly account for the presence of multiple agents, which is exactly the case in collaborative activities. It has also been extended for recursive, higher-order reasoning (finitely-nested I-POMDPs (P. Gmytrasiewicz & Doshi, 2005)), via which a robot can represent beliefs of the form "I believe that the learner believes that I believe ... " that would allow the robot to have a higher-order mental model. Within the context of an activity, the robot can be said to *have social cognition* by representing, updating, and reasoning with higher-order beliefs: these can include beliefs ascribed by the robot to the learner about the activity (*1^{st}-order mental model*), as well as those ascribed to the learner about the robot's beliefs about the activity (*2^{nd}-order mental model*). This contrasts to a robot *without social cognition*, i.e. that has a *0^{th}-order mental model* about the activity itself. There is also an extension for communicative actions (Communicative I-POMDPs (P. Gmytrasiewicz & Adhikari, 2019)) that can incorporate the dialogue acts, e.g. grounding acts performed by interacting agents.

The common use of MDPs is planning for optimal action, by assuming rationality and maximizing the cumulative reward, i.e. the sum of discounted immediate rewards. We believe that in a learning activity, this framework could be re-purposed to plan not for completing the task in a shorter time or fewer steps by rewarding these goals; but rather to try to improve learning, by automatically selecting actions that e.g. could reveal or challenge the inferred misconceptions held by the learner, and induce a socio-cognitive conflict. Thus, via automated planning within an MDP-based framework, a robot can choose its actions from its state-based model, where the state contains its higher-order beliefs about the activity and the learner: it can consider the future consequences of an action, and exhibit adaptive/learning behavior; instead of reflex-based behavior with possibly fixed rules which might rather be inferior with regards to the quality of interactions, and learning. In the second part of this thesis, we employ a state representation akin to I-POMDPs to recursively represent and maintain the robot's beliefs about its environment and the human learner. We use this representation to inform the robot's decisions and design robot behaviors that are guided by different mental models that are maintained by the robot. We evaluate the effects of the robot's model on the quality of the interactions and learning outcomes with a user study.

1.3 Research Objectives

The main objectives of this thesis are outlined as follows:

- **To assess how humans build a mutual understanding through dialogue and actions:** How humans build a mutual understanding in a learning activity and converge on a shared solution can inform the design of mutual modeling abilities for a robot to support learning. Thus, we develop methods a robot can use to (i) automatically extract patterns in how humans build a mutual understanding, and (ii) study how it leads to their performance in the task and the learning outcomes. With this in mind, we design a robot-mediated human-human collaborative learning activity, in which what children *say* and *do* is strongly tied to how they perform, and what they ultimately learn from the activity (Nasir, Norman, et al., 2020). On data we collected in a user study at schools, we investigate how children naturally and spontaneously understood each other and addressed the problem at hand. We focus on the children's dialogue and actions, and explore how they *align* within their dialogue, and from their dialogue to actions. We observe how these local level alignment contexts build to dialogue level phenomena, i.e. success in the task in terms of task performance and learning outcomes (Norman, Dinkar, et al., 2022).
- **To equip a robot with mutual modeling abilities to build mutual understanding**: We design and implement a mutual modeling framework for a robot to build and maintain a mental model with beliefs about the *world* (the activity), the human, and the human's beliefs about the robot (Norman, et al., 2021). On an example domain based on the problem in the robot-mediated human-human collaborative learning activity from Part I, we illustrate (i) what and how the robot can "think" about the world and the human, (ii) how it can decide on an action and (iii) generate an explanation for it. Based on these principles, we adapt the human-human collaborative learning activity from the first part to a human-robot version and evaluate the efficacy of the designed activity (Norman, et al., 2022).
- **To investigate the potential benefits of mutual modeling on the interaction**: In the adapted activity, a humanoid robot works together with a human learner to construct a shared solution to a problem. In this activity, the human and the robot need to build a mutual understanding on what a correct solution is, and converge to a shared solution that they agreed upon in order to solve the problem. With a user study, we explore the effects of the mutual modeling abilities we developed for the robot on the interaction and the learning outcomes measured by a pre-test and a post-test. We investigate the impact of using a correct vs. incorrect model of the activity by the robot, and the effects of using a mental model of the human maintained by the robot to drive the robot's behavior, compared to having a model of the activity only.

1.4 Thesis Structure

In the next chapter (**Chapter 2**), we present a literature review for our research context. Firstly, we highlight the unique position of social robots in educational scenarios, and their potential benefits to facilitate human learning. Then, we discuss the state-of-the-art approaches to analyze dialogue in order to reveal the learning processes. Finally, we consider the approaches for a robot to model other agents and their environment.

In **Part I**, we study human-human mutual understanding: we develop methods for a robot to automatically assess the dialogue and actions of humans, to reveal structures in how humans build a mutual understanding. In **Chapter 3**, we present a collaborative learning scenario we designed to elicit dialogue in how humans build a mutual understanding. Then, we propose novel rule-based algorithms to automatically and empirically assess collaboration, by studying how children align their representations of the activity in the data from **User Study 1** we conducted at schools. In **Chapter 4**, we focus on *lexical alignment* (what was said), i.e. alignment at a lexical level. Then, in **Chapter 5** we propose a measure of *behavioral alignment* (what was done), i.e. the alignment occurring when instructions provided by one interlocutor are followed (or not) by actions of the other interlocutor. We then observe how these *local level* alignment contexts build into a function of *global level* phenomena, i.e. success in the task, in terms of both task performance and learning outcomes. We made the dataset collected in User Study 1 and tools to study alignment in children's dialogues publicly available.[VI]

In **Part II**, we explore mental modeling for a robot, and develop skills for a robot to build a human-robot mutual understanding. In **Chapter 6**, we adapt the design of the aforementioned collaborative learning activity, so that one human learner and the robot work together. The design is motivated by the mechanisms of how humans build a mutual understanding discussed in Part I. We evaluate the efficacy of the activity in **User Study 2** we conducted remotely with children at schools. In **Chapter 7**, we formalize a mutual modeling framework for a robot in an I-POMDP-like manner, to explore representations a robot can have about its environment and the agent, i.e. the human learner. In **Chapter 8**, we design three robot behaviors that are guided by different mental models that are maintained by the robot: (i) a correct model of the activity, (ii) an incorrect model of the activity, and (iii) an incorrect model of the activity with a model of the human. We evaluate the effects of the robot's model on the quality of the interactions and learning outcomes on data from **User Study 3** we conducted with children in schools.

Lastly, in **Chapter 9**, we present the main findings and contributions in this thesis, and discuss limitations and directions for future research. This is a first step towards understanding what to model, how to model and how to use a model about the environment and the human(s).

[VI] The anonymized *dataset* consisting of transcripts, logs, and responses to the pre-test and the post-test, and *tools* for our analyses are publicly available online, from the Zenodo Repositories DOI: 10.5281/zenodo.4627104 for the dataset, and DOI: 10.5281/zenodo.6974562 for the tools.

2 Literature Review

2.1 Introduction

The ultimate research objective in this thesis is to develop mutual modeling abilities for humanoid social robots in order to support human learning. To address this objective, we focus on the relevant skills for a robot to (i) assess the mutual understanding between learners, and (ii) build a mutual understanding with a learner, as described in Section 1.3. This set of objectives lies in the intersection of three research fields as applied to educational settings: Human-Robot Interaction (HRI), Discourse Analysis (DA) and Artificial Intelligence (AI).[I] The use of social robots in educational scenarios has been studied in HRI, assessing how humans understand each other through verbal interaction has been the subject of DA, and designing agents that can build a model of other agents has been investigated in AI.

This chapter is organized as follows. Firstly, for HRI, we consider the potential benefits and possible roles of social robots in education, the need to investigate robot behaviors to facilitate learning, and how we employ the robot in this thesis to achieve our objectives in Section 2.2. Then, we look into how previous work in DA studied verbal behavior and dialogue to extract patterns in how humans learn and understand each other; and how we can build on these to develop abilities the robot can use to assess the mutual understanding between learners, in Section 2.3. Next, we consider the approaches in AI to develop agents that can model other agents, that informs how we design our robot to build a mutual understanding with a learner, in Section 2.4. Finally, we present our overall perspective in Section 2.5.

2.2 Unique Position of Social Robots to Support Human Learning

In this thesis, we employ a physically embodied agent—a *social robot*—in educational scenarios, and investigate the effects of mutual modeling abilities for the robot on human learning. Social robotic agents have a potential to become a part of educational environments, by undertaking a unique position that extends their functional purpose with personal and social

[I] as fields of research within *Robotics*, *Sociolinguistics*, and *Computer Science*, respectively.

dimensions of interaction (Belpaeme et al., 2018). Virtual pedagogical agents and Intelligent Tutoring Systems (ITSs) have demonstrated their effectiveness on learning (Krämer & Bente, 2010; Kulik & Fletcher, 2016; Schroeder et al., 2013), where the social, emotional and motivational aspects could be the moderating factors that are crucial in how they facilitate learning (Krämer & Bente, 2010). Among agents that provide support with social interaction, physically embodied robots tend to have a higher impact on learning and be more effective for the desired changes in behavior, compared to virtual pedagogical agents (Belpaeme et al., 2018). To leverage the social dimension, Belpaeme et al. note that social robots can behave in ways to encourage greater effort by the learners, which can help them learn better. Thus, robots have been shown to produce positive effects on learning similar to pedagogical agents and ITSs, e.g. in Kennedy et al. (2015), Ramachandran et al. (2016), and Tanaka and Matsuzoe (2012), where the robots assume diverse roles in the interaction.

The robot's behavior, as well as its responsibilities and therefore contributions to the learning process are determined by its role: for example, a tutor robot needs to navigate the learner to gain knowledge about the content being taught (Rohlfing et al., 2022). Mubin et al. (2013) characterizes the types of roles for a robot in a learning scenario, in terms of how it is involved in the learning processes: (i) robot as a *learning tool* or teaching aid that has a passive function, e.g. to teach robotics or other science, technology, engineering, and math (STEM) content (Karim et al., 2015); (ii) robot as a *co-learner*, peer or companion, tutor or tutee, serving as an *agent* that actively and spontaneously participates in the interaction. As an example that uses a social robot as a peer/companion, Okita et al. (2009) showed that an ASIMO robot that cooperates with the child like a classmate can to promote learning. Social robots have been used as teachers or tutors to guide children's learning of skills such as a second language, where it is the robots that are *teaching* humans (Vogt et al., 2019). The robots have also been utilized as teachable agents, namely as "surrogate pupils" for children to teach (Werfel, 2013), where it is rather the humans that are *teaching* robots: to (i) the benefit of the robot itself to acquire new skills, akin to the *learning from demonstration* paradigm (Argall et al., 2009); or (ii) the learning of children, e.g. in Lemaignan, Jacq, et al. (2016). Several works on the latter role employed the *learning by teaching* paradigm (Chase et al., 2009): a NAO robot takes the role of a novice needing help from the human; e.g. to enhance their vocabulary via a robot controlled by *Wizard-of-Oz* (Tanaka & Matsuzoe, 2012), or to improve handwriting via an autonomous robot (Lemaignan, Jacq, et al., 2016). Such a role is not very straightforward for a human to take on, as it might not be convincing or even doable: this highlights a unique position for social robots in education.

In this thesis, we explore ways to benefit from this unique position and the added social dimension: we bring in the robot to mediate, as a companion, two related learning scenarios in Part I and Part II, with the functional purpose to support the learning activity by completely automating the scenarios. Meanwhile, the robot provides assistance through the social dimension, and (for us) serves as a test bed for developing and evaluating the mutual modeling abilities for the robot in educational settings. This implicit role of automating work with social robots, the robot as an assistant that provides support through social rather than physical interaction, has been studied as socially assistive robots (SARs) (Matarić & Scassellati, 2016). Robots have been successfully employed in various settings (Matarić, 2017), like service settings such as healthcare and therapy, for rehabilitation to “augment human care and existing robot-assisted hands-on therapy" (Matarić et al., 2007), elderly care (Pu et al., 2019; H. Robinson et al., 2014), and education (Belpaeme et al., 2018; Leite, 2013; Papadopoulos et al., 2020); as well as special education for children with neurodevelopmental disorders (Papakostas et al., 2021). For instance, robots have served this implicit role in the context of robot-assisted language learning (RALL), and acted as a tutor, tutee, or a peer to promote language development (Neumann, 2020; van den Berghe et al., 2019).

In Part I, we employ the robot to (i) motivate the two learners to work together by providing support through minimal verbal and non-verbal behaviors, and (ii) give minimal feedback on their solutions, as described in Chapter 3; while fully automating (as a SAR) a human-human collaborative learning scenario that we designed. This corresponds to supporting/assisting roles for the robot, where it gives social and emotional support and endorses engagement during the learning process. For instance, Caruana et al. (2022) explored how a robot can act as a “reading buddy" to support children's engagement in reading by using different robots (NAO, MiRo and Cozmo): although only one of the robots (NAO) could engage in social dialogue, they found that all the robots supported the engagement of children. This kind of a supportive role for the robot has been demonstrated to be effective in RALL contexts (Deublein et al., 2018; Engwall & Lopes, 2020; Hong et al., 2016; Hsiao et al., 2015), by e.g. (Hsiao et al., 2015) that observes a robot is advantageous over a tablet computer.

In our scenario, in Part I, we then develop perceptive capabilities for the robot, to automatically assess the mutual understanding between the learners based on their dialogue and actions in Chapter 4 and Chapter 5. Previous research also indicates that perceptive abilities can inform dialogic abilities of the robot, where it can adapt to the characteristics of the human can enhance the experience and form a supportive learning environment, e.g. adaptation to speech (Gulz et al., 2011; Kumar et al., 2010). For instance, Lubold et al. (2018) showed that a robot that speaks socially and adapts to the child's pitch has a higher positive effect on learning, than a robot that does not adapt or carry out social dialogue.

In Part II, we explore another way to benefit from the added social dimension: we design the robot to autonomously and more directly participate in the collaborative problem solving process by making decisions together with a learner and converging to a shared solution, in a human-robot version of the activity that we adapt (in Chapter 6) from the human-human activity in Part I. This use is similar to Okita et al. (2009), that compared an ASIMO robot that cooperates with the child like a classmate, against a lecturing robot that demonstrates a correct way to solve the problem at hand, and a parallel-playing robot that simply follows the human: they found that children that interacted with the cooperative robot learned more. As another example for a social robot acting as a peer that has been highly effective, Mazzoni and Benvenuti (2015) induced a socio-cognitive conflict between the child and the robot via the robot raising doubts and not providing the solution to the learner, to teach new words in a second language.

2.3 Analysis of Verbal Behavior to Reveal Learning Processes

Broadly speaking, both *verbal* and *non-verbal* behaviors could be indicative of the learning process (Trausan-Matu & Slotta, 2021). Non-verbal behaviors (gaze, gestures, laughter ...) have been linked with the *quality of interaction* (see Bangalore Kantharaju et al., 2020; Jermann et al., 2011; Jermann and Nüssli, 2012; Schneider et al., 2021). Verbal behaviors could consist of textual input such as manual/automatically obtained transcripts, or even acoustic features of the speech (pitch, loudness ...). Speech duration for instance was found to be longer in participant pairs that collaborated better (Jermann & Nüssli, 2012). While both acoustic and non-verbal behavior are important to study interaction, analysis can sometimes be limited without further context; for instance it may not always be the case that speech indicates better collaboration, as speech could even consist of negative interactions.

2.3.1 Analyzing Textual Representations of Dialogue

In Part I, we study how humans build a mutual understanding by mainly utilizing textual input, i.e. manually transcribed dialogues. Textual inputs can allow for a finer grained analysis of collaborative learning; as using text can give further insight into "the processes of development of collective thinking in and by dialogue" (M. J. Baker et al., 2021). Textual inputs then serve as basis for coding schemes that can be used to *represent* various types of verbal interactions. The goal is to label each segment in the dialogue, that can range from a word, to a sentence, to the complete dialogue itself (Strijbos et al., 2006). For example, to analyze the process of how learners construct arguments through dialogue, Weinberger and Fischer (2006) proposed a scheme that manually labeled online, written discussions. There were labels along multiple process axes, such as the extent to which learners contributed to the dialogue, and the content of their contribution (then automated in further works by Dönmez et al., 2005; Mu et al., 2012; Rosé et al., 2008). Thus, (both manual and automated) representation of dialogue can elucidate the learning process, by showing trends that happen in collaboration, and the effect of particular patterns in dialogue on the learning outcomes (Borge & Rosé, 2021).

The analysis of verbal interactions in educational settings has also been investigated in the domain of *ITSs*. These systems aim to adaptively facilitate learning by extracting content from the learners' contributions to dialogue and continuously modeling the evolution of their learning. An important component of dialogue-based ITSs is the representation of the input text into speech acts (D'Mello & Graesser, 2013). Speech acts serve to classify the discourse into communicative/pragmatic functions, such as backchannels (e.g. "uh-huh"), metacognitive statements (e.g. "I need help") and so on. This representation is required by the ITSs to model the learner's progress and generate an appropriate response. For instance, most speech acts used by learners consist of answers to questions asked by the tutor (D'Mello & Graesser, 2013). Following this, many works that focused on educational scenarios have developed methodologies to perform dialogue (speech) act classification using both supervised and unsupervised methods (e.g. Boyer et al., 2010; Ezen-Can and Boyer, 2014; Ezen-Can and Boyer, 2015a, 2015b; Goodman et al., 2005).

In the *intersection between collaborative learning and ITSs*, analyzing student interactions to see how they communicate and collaborate with each other has served as a basis to give adaptive feedback (Tchounikine et al., 2010). These interactions are more complex than conventional one-on-one tutoring, as there are added social dimensions (Howley & Rose, 2016). Here, ITSs are used to further the students' skills in a collaborative activity (Scheuer et al., 2010), or even facilitate learning. Walker et al. (2011a, 2011b) for instance, have implemented a system in such a scenario, where feedback is adaptively given depending on the interactions students are having in the classroom. Here, student collaboration was actively analyzed using the system put forth by Rosé et al. (2008): an automated text classification software used in educational contexts. Thus distinguishing types of interactions based on verbal behavior has been implemented with success to equip an agent acting as a tutor with better feedback capabilities. In Part I, we utilize a *dataset* we collected in User Study 1, where pairs of children are engaged in a robot-mediated collaborative learning activity that we designed to elicit dialogue: the robot uses minimal terminology to explain the task to the children. While the robot intervention is minimal in this dataset, our work also contributes towards research that could enhance the capabilities of the robot; enabling it to intervene to facilitate learning as a future research direction.

2.3.2 Complications of Working with Speech Data

Since our dataset from User Study 1 consists of *spoken* dialogues among children, in addition to studying the process of collaboration, we would like to observe *how* the speech modality contributes to collaboration. Previous research also indicates the importance of the speech modality in educational scenarios. Litman et al. (2004) for instance found that learning gain is positively impacted when the interaction modality includes speech, compared to solely written input for human-human tutoring data. Related to collaboration, the co-presence of interlocutors, visibility and audibility in the medium have been found to dramatically affect the process of mutual understanding (Clark & Brennan, 1991; Dillenbourg & Traum, 2006).

In ITSs, a first step in the pipeline is the "transformation of the input into a parsable format" (D'Mello & Graesser, 2013). This is not straightforward when the input is spoken, as the quality of transcription (i.e. the parsable format) can depend on the performance of the Automatic Speech Recognition (ASR) system. After the input is transcribed, and also for general dialogue systems, it is then collapsed into a semantic frame (see Louvan and Magnini, 2020; Tur and De Mori, 2011 for general task oriented-dialogue), consisting of speech acts and certain keywords/phrases to estimate *local* and *global* levels of learning. The performance of the dialogue act classifier depends heavily on pre-segmented input utterances, where the assumption is that each input utterance will correspond to an output speech act label. There is a missing step between transcription of the speech (which may include interlocutors interrupting each other and themselves, using multiple dialogue acts per speaker turn ...) and dialogue act classification; i.e. the segmentation of the utterances. However, the task of automatic dialogue act segmentation itself, remains an open issue (e.g. see Zhao and Kawahara, 2019), particularly when the input is spoken conversation. Thus while speech acts are important in the analysis of educational data, further methodologies are required to improve dialogue act segmentation before classification on completely spontaneous dialogues. In Part I of this thesis, we develop a different methodology that is not focused on dialogue acts, to analyze textual inputs arising from completely spontaneous dialogues. This avoids the requirement of pre-segmented utterances. However, we still depend on the good quality transcriptions, as discussed further in Chapter 4.

2.3.3 Analysis of Alignment in Educational Dialogue

The alignment of behavior has been investigated at varying linguistic levels, from acoustic/prosodic (e.g. Lubold and Pon-Barry, 2014; Thomason et al., 2013; Ward and Litman, 2007a, 2007b), to lexical (e.g. Ward and Litman, 2007a, 2007b), and syntactic (e.g. Reitter and Moore, 2014). Research on alignment in educational scenarios can shed light on learning outcomes. For instance, Sinclair and Schneider (2021) showed that linguistic and gestural alignment in dialogue correlate with learning and are indicative of success in collaborative problem solving. Similarly, Ward and Litman (2007b) found that the convergence of lexical and acoustic behaviors can predict learning outcomes. Convergence of behavior can be thought of as a variation of alignment. Even in an automated setting, Lubold et al. (2018) showed that a robot that aligns with the interlocutor and speaks socially has a positive effect on learning.

Research on alignment can also give information about the *process of communication, the dynamic between interlocutors*, etc. For instance, to build *rapport* (“the development of personal relationships between speakers over time" (Sinha & Cassell, 2015)), interlocutors become closer to each other in terms of acoustic/prosodic behavior (Lubold & Pon-Barry, 2014). Following this, Lubold et al. (2015) found that pitch alignment of a learning companion leads to higher perceptions of rapport. Sinha and Cassell (2015) showed behavioral convergence and rapport are linked to each other besides being correlated to learning gains; but in this case, rapport leads to convergence (of speech rate) in dyadic peer tutoring conversations.

Meanwhile, we consider automatic measures of alignment in terms of engagement, for completeness of the exposition of this literature survey. For instance, visual attention mechanisms have been studied as indicators of a human’s engagement in a conversation or the task (Oertel et al., 2020). In joint attention, two agents build a perceptual alignment to an object in terms of attentional behavior, which has been for instance measured by gaze cross-recurrence (Jermann & Nüssli, 2012). In contrast, *with-me-ness* is a gaze-based indicator proposed in education to measure how much a learner is “with the teacher" conceptually, through a joint focus on what they refer to (Sharma et al., 2014). This measure was successfully adapted to HRI for interactive tasks, as a metric of alignment to quantify to what extent the human is "with" the robot, by inferring the human's focus of attention and comparing it with the targets of attention that are elicited by the task (Lemaignan, Garcia, et al., 2016).

These works serve as a foundation to indicate the positive influence that alignment has on learning outcomes. In Chapter 4, we propose automatic measures of *verbal (lexical) alignment* to observe in our dataset whether local alignment is associated with global task success. In Chapter 5, we similarly study *behavioral alignment* (when instructions provided by one interlocutor are either followed or not followed with physical actions by the other interlocutor). As discussed, works have studied how alignment can build to other phenomena such as rapport (Lubold & Pon-Barry, 2014), and vice versa (Sinha & Cassell, 2015). However, to the best of our knowledge, we are the first to focus on how alignment arises from the timely occurrence of actions in a physical environment, thus concretely forming a link between what was said, and what was done as a result.

2.3.4 Analysis of Alignment in Other Settings

Also outside of educational scenarios, several methods have been proposed to automatically compute alignment, to gain further insight into the communication process and the dynamics between interlocutors; e.g. to study deceptive and truthful speech in interview dialogues (S. I. Levitan et al., 2018), and study the relationships between levels of alignment in multi-party dialogues for a cooperative board game (Rahimi et al., 2017). Since we focus on lexical and (novel) behavioral alignment, we discuss other works that have studied lexical alignment and proposed methodologies to automatically extract its occurrences.

Lexical alignment (also referred to as *entrainment*) has been studied by focusing on various features such as referring expressions (Brennan & Clark, 1996), repeated sequences in utterances (Dubuisson Duplessis et al., 2017, 2021), frequent words in the discourse (S. Levitan et al., 2018), hedge words (S. Levitan et al., 2018) and even expressions related to the task (referred to as "topic words" in Rahimi et al., 2017). Many of these works build from older linguistic works, to automatically study lexical alignment.

We "align" with the literature on lexical entrainment in some aspects, e.g. by focusing on expressions related to the task. However, there are drawbacks with the current methodologies to automatically compute entrainment. Firstly, many works that study entrainment approach it as a high-level process (e.g. Friedberg et al., 2012; Nenkova et al., 2008; Rahimi et al., 2017), where the focus is on quantifying the overall entrainment between the interlocutors to see how it can build in the discourse (by looking at proximity: a degree of similarity, and convergence: its evolution). However, while lexical alignment builds in the discourse, there is no consideration for the individual contributions of the interlocutors. Hence, we break down lexical alignment into these individual contributions, i.e. we treat it as a process that has different stages (compared to a holistic approach). We consider the first time an interlocutor introduces an expression (priming), and the time when the other interlocutor utilizes the same expression (establishment). Then, like other works, we see how alignment can build in the discourse. Specifically, while we also focus on expressions related to the task (e.g. Rahimi et al., 2017), we propose to focus on how such expressions are established: this is distinct from checking overall entertainment on a set of words related to the activity.

Dubuisson Duplessis et al. (2017, 2021) proposed automatic and generic measures to extract *lexical structures of alignment* (which they refer to as *verbal alignment*) in a task-oriented dialogue. The proposed method works on alignment based on surface matching of text and does not focus on other levels of linguistic alignment as envisioned by Pickering and Garrod (2004). However, it is done with the aim of automatically finding these text patterns in the dialogue, by sequentially processing a transcript in an unsupervised manner. In our case, we have a *situated dialogue* where the interlocutors share time and space, and the dialogue is about and happens within the context of their common environment. In Chapter 4, we provide a new tool/framework for studying situated dialogue that builds on the automatic and generic methodology by Dubuisson Duplessis et al. (2017, 2021): this implies to model how the interlocutors refer to their environment and to extend the tool based on this model. Then, in Chapter 5, we propose new measures that allow us to study behavioral alignment in situated dialogues; i.e. via automatically inferring instructions given by the interlocutors, and then linking those instructions to actions taken in the physical environment.

2.4 Approaches to Modeling Other Agents to Improve Interaction

In Part II of this thesis, we develop modeling abilities for the robot in order to support human learning. A central area of research in AI, as the "study of constructing machine intelligence from an agentive perspective" (Zambak, 2009), focuses on designing agents that can have an effective interaction with other agents, by being capable of reasoning about the actions, goals and beliefs of other agents (Albrecht & Stone, 2018). For this purpose, the agent can construct a *model* of the other agent, where in general a model is a *mapping* from the observed history of interaction between the modeling and the modeled agent to properties of interest to be the predicted about the modeled agent. Various scenarios necessitate effective modeling capabilities for agents: for instance, (i) ITSs need to do *learner modeling*, e.g. represent the learner's understanding and misconceptions about the activity in order to track and/or improve learning learning (J. R. Anderson et al., 1990; McCalla et al., 2000); (ii) dialogue systems need to do *user modeling* (Kass & Finin, 1988), e.g. represent the plans and intentions of the interlocutors in a task-oriented dialogue (Cohen & Perrault, 1979; Grosz & Sidner, 1986; Litman & Allen, 1984); (iii) autonomous cars need to model the other human drivers to effectively coordinate with them (Sadigh et al., 2018), and represent human intentions for effective on-road pedestrian avoidance (Rasouli et al., 2020); and (iv) industrial robots need to represent their human teammates coordinate and collaborate better (Tabrez et al., 2020). In this thesis we investigate whether, by modelling the human learner it interacts with, the robot can not only improve the quality of the interaction but also the learning outcomes: the robot is the modeling agent, the learner it interacts with is the modeled agent.

2.4.1 Modeling with Incomplete Information

In Part II, we investigate how a robot can build a model of a human in an educational setting, that involves collaboratively solving a given problem in their shared environment. The human may not know how to solve the problem, therefore, the robot has *incomplete information*: the parameters necessary to determine the human's approach are unknown to the robot. Computational modeling of other agents with incomplete information goes back to the beginnings of *game theory* that links how an agent should act to an equilibrium analysis from an *external perspective*: this involves finding steady strategy *equilibrium profiles* called the Nash equilibria (Nash, 1950), i.e. states where "no agent has incentive to deviate from its strategy given that others do not deviate from theirs" (Sandholm, 2010). Then, *decision theory* uses this analysis in order to calculate the action that maximizes the agent's expected utility, typically by Bayesian update of the individual prior beliefs (Aumann & Maschler, 1995; Harsanyi, 1967).

A typical example for the game-theoretic approach is to model the other agent by *fictitious play* (G. W. Brown, 1951; J. Robinson, 1951) where the agent computes the average empirical frequency of the other agent's actions from the history, assuming fixed behavior: this has been utilized by multi-agent reinforcement learning (MARL) (Claus & Boutilier, 1998). Another example by Kalai and Lehrer (1993) is *rational learning*, where agents maintain subjective, Bayesian beliefs about the possible strategies of the other agents: the agents' objective is to maximize their expected reward, with respect to their own subjective beliefs: Bayesian updating follows from this maximization of the expected utility. In competitive, perfect information games like chess, the leading solutions have been based on the *minimax* decision rule; that involves the agent optimizing against the possible loss at the worst case (Campbell & Marsland, 1983). This game-theoretic analysis has been extended to larger and more complex games, e.g. the (earlier) standard benchmark Poker (Rubin & Watson, 2011; Sandholm, 2010) to develop agents that bested top human players (Bowling et al., 2015; N. Brown & Sandholm, 2018, 2019).

Doshi et al. (2020) note that for every agent to act according to the equilibrium profile, the epistemic conditions are: (i) the agents are *rational*, i.e. maximizing utility, and (ii) the joint payoff/utility of the game is part of their *common knowledge*, i.e. it is mutual knowledge at all levels, such that "all know it, all know that all know it, all know that all know that all know it, and so on *ad infinitum*" (Aumann, 1999), leading to infinitely nested beliefs (Mertens & Zamir, 1985). Attaining common knowledge is impractical in general (Halpern & Moses, 1990): thus, equilibria could be impractical in some situations, as the game situation can call for assumptions about a *common knowledge of rationality* and *common priors*, in order to have "a well-defined expectation: the protagonist's expected payoff, as calculated from her strategy and her belief about the strategies of the others" (Aumann & Dreze, 2008). In the educational settings we are interested in, these assumptions would not be satisfied about the human, since the human might not a priori know but rather *learn* through the interaction how to solve the given problem.

As a step towards learning about the human, that can also be learning over time, inverse reinforcement learning (IRL) approach infers the reward function of the other agent, by assuming an unknown but fixed reward function that the agent optimally follows (Ng & Russell, 2000). Alternatively, the agent can learn about and imitate the human as a collaborator, assuming fixed behavior (Chan et al., 2019). To learn from agents that are also learning, the agent can infer the changing reward from its observations, by assuming that the agent is also learning the task and accordingly improving its sub-optimal behavior over time (A. Jacq et al., 2019; Ramponi et al., 2020), or learn to improve its performance by observing the behavior of other agents in a shared environment with a common reward function (Ndousse et al., 2021). The agent can anticipate the learning of the other agents to shape their learning process in the next few steps (Foerster et al., 2018), and can do so symmetrically in the longer term (Lu et al., 2022). A unifying perspective on inverse decision modeling by Jarrett et al. (2021) subsumes these approaches that characterize the behavior of an agent behaving sub-optimally etc., by generalizing the "the idea of 'boundedness' in sequential decision-making".

2.4.2 Intentional and Recursive Modeling to Represent Higher-order Beliefs

In Part II, we design the robot to build an *intentional model* of the human, by attributing to the human beliefs about the environment. There exists a large variety of modeling approaches in the literature, that differ considerably in their assumptions and methodology, and serve different purposes in connection to the needs of their communities (Albrecht & Stone, 2018). Dennett (1987)'s intentional stance allows for a useful characterization of these modeling methods: modeling is either *intentional*, i.e. it attributes meaningful mental states to the modeled agents like beliefs and desires in order to explain their behavior, or not (Doshi et al., 2020). Intention modeling brings about the modeled agent to also model others, that results in *recursive modeling*: the agent can explicitly describe nested beliefs in a way that simulates the reasoning process of an agent in the form "I believe that you believe that I believe (...)".

An approach that represents an agent's finite nested knowledge is recursive modeling method (RMM), that can incorporate recursive reasoning into the decision making of the agent in a single time step: RMM differs from the above-mentioned game-theoretic approach as it takes the *subjective perspective* of the individual interacting agent, by modeling its current belief state (P. J. Gmytrasiewicz & Durfee, 2000). The recursive model structure of RMM is expressive in the sense that it also includes at its levels *sub-intentional models* that do not have ascription of beliefs[II], as well as *no-information models* that can represent the knowledge limits of the modeling agent. RMM has been extended to Interactive-POMDP (I-POMDP) by adding elements of belief update and sequential planning (P. Gmytrasiewicz & Doshi, 2005) and following the paradigm of Partially-Observable Markov Decision Process (POMDP) (Kaelbling

[II] P. J. Gmytrasiewicz and Durfee (2000) note that the design and physical stances of Dennett (1987) are two sub-intentional stances used to predict behavior: (i) by the functionality, e.g. the combined behavior of a digital circuit is a result of the functions of its components, implementing "the intentions of its designer" (Hamscher, 1991), and (ii) by a physical/dynamics description of the object, e.g. motion of a bouncing ball (Forbus, 1980), or representing the strategy of an agent as deterministic finite automata (Carmel & Markovitch, 1996), respectively.

et al., 1998), so that I-POMDPs account for the presence of other agents by including their models in the state space: an agent decides sequentially according to its beliefs about the environment and about the other agents, where the belief of an agent about the state is represented as a probability distribution over the possible states. Doshi et al. (2014, 2010) used I-POMDPs to successfully model recursive reasoning behavior of humans, on data from user studies by Goodie et al. (2012): they made appropriate simplifications; for instance, instead of choosing the action that maximizes the expected utility, they use *quantal response* (McKelvey & Palfrey, 1995) i.e. selecting actions proportionally to the utilities.

Milliez et al. (2014) present a spatio-temporal belief system for an agent to build and maintain a 3D geometric model of the world: it can generate an independent model of the human with symbolic facts that do perspective taking, and enable divergent belief management. In the same manner, Breazeal et al. (2009) maintain separate, parallel belief bases, where the proposed approach can attribute beliefs that are possibly false to other agents and drive the robot's behavior accordingly: they found that the model produces robot behavior that is consistent with the human's behavior, and that the robot can infer the same conclusions as the human do. In Part II of this thesis, we also follow this distinct belief bases approach.

Alternative to quantitative/probabilistic representations as in I-POMDPs are qualitative representations of belief through epistemic logics, beginning with Aumann (1999) that model logic of knowledge and belief in multi-agent settings: what each agent believes about what the other agent believes. In ordinary epistemic logic, modal operator K (or B) is used to represent knowledge (or belief) such that the formula $K_i\phi$ reads "Agent i knows that ϕ.", and $K_jK_i\phi$ corresponds to "Agent j knows that agent i knows that ϕ.". Dynamic Epistemic Logic (DEL) (Bolander & Andersen, 2011; van Ditmarsch et al., 2007) extends this by allowing actions in the form of event operators that modify facts in the world, similar to situation calculus as in STRIPS (Fikes & Nilsson, 1971). For instance, Dissing and Bolander (2020) used DEL to implement ToM for a robot and demonstrated that it allows the robot to pass false belief tasks.

Lemaignan and Dillenbourg (2015) highlight the potential benefits of adopting logic-based approaches to ToM and mutual modeling. To equip robots with these skills, for instance, Buisan et al. (2022) focus on task planning for the robot to generate a joint plan, while considering the human's decisions and actions: it uses a symbolic knowledge and reasoning framework to semantically represent the state of the environment (Sarthou et al., 2019), and a physics reasoner to make inferences (Sallami et al., 2019). In this thesis, we use subsymbolic approaches like I-POMDP over logic-based approaches, while still keeping the close link between by using a sparse state representation that keeps a direct semantic link to linguistic descriptions.

POMDPs and their extensions are particularly suited for uncertainty, by incorporating partial observability of the state and any non-determinism in the state changes into the planning, with a "distinct Markovian mark imprinted on automated planning" (Jensen, 2014)[III]. Partial

[III] by assuming a *Markov* model: "the model is Markov if the state transitions are independent of any previous environment states or agent actions" (Kaelbling et al., 1996).

observability leads to the notion of a belief state; yet this state space quickly becomes too large, making POMDPs impractical for complex tasks. In contrast, logic-based approaches like DEL offer compact and transparent representations for belief states: these can be fused together with POMDPs as in Herzig et al. (2003) and C. Wang and Schmolze (2005), by providing logical descriptions to the value functions, e.g. as in first-order MDPs (Boutilier et al., 2001)[IV]. In this thesis, we adopt a practical perspective, and represent the robot's interaction with its environment in an abstraction of the finitely-nested I-POMDP, that links directly to logical descriptions on the belief state to guide the robot's behavior in a rule-based manner. Our focus is on what kind of mental models the robot can have about the activity (e.g. a correct vs. incorrect model and conception of the problem), how the human learner can be represented (with belief attributions from the actions the human takes), and how this mental model about the human can be used to drive the robot's behavior (e.g. to adapt to the human, by preferring choices in their common ground that consists of their aligned, agreed-upon choices).

2.5 Conclusion

The purpose of this chapter is to review the trends in HRI, DA and AI around the use of social robots in educational settings, and the ways in which such robots can be endowed with mutual modelling abilities, to the purpose of supporting human learning. We surveyed approaches, methods and relevant research for a robot to automatically (i) assess the mutual understanding between learners, and (ii) build a mutual understanding with a learner, along the objectives of this thesis as described in Section 1.3. It is indeed not straightforward to design robot capabilities that can effectively help improve learning as it comes as a deeper goal than the typically studied immediate, performance goals that machines tend to monitor to drive their behavior.

It is clear from the research reviewed that in the literature there exists a large variety of roles and behaviors for the robot, modalities or representations of dialogue it can assess that relate to the mutual understanding between humans, and modeling approaches to represent the human to build a mutual understanding. However it is not clear at all which of these roles, behaviors, modalities etc. are effective in an educational setting and how. Adapting formal methods (i) to the abilities a robot can use, (ii) to perceive humans, (iii) within an interaction aiming at supporting the human's acquisition of specific learning goals presents interesting challenges, and asks for bringing together different perspectives, to verify the findings and intuitions from the literature: paving a way for fresh perspectives.

[IV] Or by probabilistic extensions logic, see De Raedt and Kimmig (2015) for a survey.

Part I

Human-Human Mutual Understanding

3 Collaborative Learning Activity Design to Elicit Dialogue

3.1 Introduction

In this chapter, we present the design of a robot-mediated human-human collaborative problem solving activity that aims to improve the Computational Thinking (CT) skills of children. The activity is designed in a way that elicits dialogue on how humans build a mutual understanding as they solve a problem together. We describe a dataset of 39 teams of two children (i.e. 78 children in total), resulting from User Study 1 that we conducted with 96 children aged 9–12 in two schools. Then, we present an overall evaluation of the activity in terms of the participants' performance in the task, learning outcomes, and subjective perspectives about their own competence, engagement, mutual understanding, and the robot.

Results highlight that the link between learning and performance is not straightforward; as we observe no evidence for a correlation between performance and learning. Meanwhile, performance correlated with participants' self-assessment of their competence and mutual understanding. Surprisingly, despite its rudimentary behavior, participants tended to perceive the robot as highly competent, intelligent, friendly, likeable, not distracting, and report not feeling a need for more feedback from the robot; regardless of their performance in the task.

This work corresponds to the following publication:

> Nasir*, J., **Norman*, U.**, Bruno, B., & Dillenbourg, P. (2020). When positive perception of the robot has no effect on learning. *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 313–320. https://doi.org/10.1109/RO-MAN47096.2020.9223343. **Contributed equally to this work.* (Nasir, Norman, et al., 2020)

Within this work, the author of this thesis implemented the learning activity components and their interaction with the robot, co-directed User Study 1 at schools together with colleague Jauwairia Nasir, and contributed to the design of the activity, evaluation tests and data analysis.

In the two remaining chapters of Part I, we use a subset of this dataset from User Study 1 to

Table 3.1 – Research questions and hypotheses for the evaluation of the JUSThink activity

No.	Research Question / Hypothesis
RQ1	*Is the activity successful in achieving its goal to elicit dialogue? Is the amount of dialogue associated with the performance in the task or learning outcomes?*
H1.1	The speech activity is high.
H1.2	Higher the speech activity, the better is the performance in the task.
H1.3	Higher the speech activity, the better are the learning outcomes.
RQ2	*How are the learning outcomes after collaboration? Is this associated with the performance in the task?*
H2.1	Participants perform better in the post-test than the pre-test.
H2.2	The better the performance in the task, the better are the learning outcomes.
RQ3	*How do participants perceive their engagement, mutual understanding, and the robot?*
H3.1	Participants' self-assessment of their engagement and mutual understanding lies more on the positive side of the spectrum than on the negative side.
H3.2	Participants' self-assessment of the robot lies more on the negative side of the spectrum than on the positive side.
RQ4	*How do the performance in the task and the learning outcomes relate to participants' self-assessment?*
H4.1	Teams that perform well perceive their mutual understanding, engagement, and self-competence more on the positive side than teams that perform badly, and have a more positive assessment of the robot.
H4.2	Teams that learn perceive their mutual understanding, engagement, and self-competence more on the positive side than teams that do not learn, and have a more positive assessment of the robot.

develop methods a robot can use to automatically assess the mutual understanding between the learners: we focus on their dialogue in Chapter 4, and the relationship between their dialogue and actions in Chapter 5. Then, in Part II, we adapt the activity described here to a human-robot version, to develop skills for the robot to build a mutual understanding.

The described activity and the dataset from User Study 1 has also served as a baseline for other lines of research: for instance, Nasir et al. (2021) studied the dataset to identify behavioral profiles from multi-modal features about speech, gaze, facial expressions, and actions.[I] Based on these profiles, Nasir et al. (2022a, 2022b) proposed methods for the robot to automatically assess the engagement of children in the learning processes: their goal is to equip the robot with the skills to intervene effectively and support the engagement in the direction of learning. These skills can complement the mutual modeling abilities we develop in this thesis.

[I] Nasir et al. made public two anonymized datasets that are derived from the dataset of User Study 1: the datasets contain engagement related features on the teams' multi-modal behaviors based on their speech activity, action logs, affective states, and gaze patterns as well as the learning outcomes. The team level dataset is available from the Zenodo Repository DOI: 10.5281/zenodo.4288832, and time series dataset is at DOI: 10.5281/zenodo.5576057.

3.1.1 Research Questions and Hypotheses

Table 3.1 presents our research questions and hypotheses for the overall evaluation of the designed activity in this chapter: we refer to this collaborative activity as the *JUSThink (collaborative) activity*, inside the *JUSThink (pedagogical) scenario* that also includes evaluation tests etc. In RQ1, we consider to what extend the activity is effective in eliciting dialogue, and whether the amount of dialogue is related to the performance in the task or the learning outcomes. In RQ2, we evaluate how the teams' performance in the task relates to their learning outcomes; as the performance is what the robot can manipulate more easily, e.g. by giving hints. In RQ3, we look into the participants' self-assessment of their engagement and mutual understanding, as well as how they perceived their competence, stress, and the robot, in order to gain a general view on their subjective experience. In RQ4, we check whether and how much the participants' self-assessment reflects their performance and learning.

For RQ1, we postulate that to collaborate effectively, the participants speak to each other to produce a dialogue in which they convey information and collaboratively decide on actions throughout the activity: we hypothesize in H1.1 that this amounts to a considerable portion of the activity. In this manner, we hypothesize in H1.2 that the more they speak, the higher is their performance in the task, as the activity is designed in such a way that the best decisions are made together, and this decision making is governed by the dialogue. In H1.3, we hypothesize that the more they speak, the better are learning outcomes, since by dialogue, they likely have argued and resolved misconceptions they had about the problem.

For RQ2, we postulate that the collaboration to solve the problem together would have a positive impact on the learning outcomes: the effort to build a shared solution can help the participants realize their misconceptions, regardless of their success in finding a correct solution. Hence, we hypothesize in H2.1 that the participants perform better after the activity (in the *post-test*) than before the activity (in the *pre-test*). Without a control group, we can only check if they improved but we cannot claim this the “impact" of collaboration. Thereby, we design the tests as individual exercises, that aim to measure the participants' understanding of the underlying concepts needed to correctly reason about the problem. Furthermore, whether the participants succeed in finding a correct solution (or how close they got to one) can reflect if they learn or not (or how much they learn): thus, we hypothesize in H2.2 that the learning outcomes correlate with the performance in the task.

For RQ3, we expect the participants to be highly engaged with the activity as a game and a challenge given by the robot, and thus endeavor to collaborate effectively towards solving the problem; regardless of their success at the end. Therefore, in H3.1, we hypothesize that the participants tend to perceive themselves as highly engaged and good at understanding each other, at the end of the activity. Furthermore, since the robot's support and interventions are limited, e.g. contains only generic hints that do not guide them towards a correct solution, we hypothesize in H3.2 that the participants have a negative assessment about the robot.

For RQ4, we expect that the participants' performance in the activity and learning outcomes correlate with how they perceive their engagement, mutual understanding, and the robot. Thus, we hypothesize in H4.1 that the participants that perform better (and in H4.2, the participants that learn more) tend to have a more positive assessment in these regards.

3.2 Design

We design the learning experiences that can shape the participants' interactions by standard in instructional design through the *backward design* approach (Wiggins & McTighe, 2005): we (i) identify the desired results, i.e. the *learning goals*, in Section 3.2.1, (ii) determine acceptable evidence that indicate the desired results are achieved, as the *evaluation tests*, in Section 3.2.2, and (iii) design the activities that can lead to learning as a *pedagogical scenario* in Section 3.2.3.

3.2.1 Learning Goals

Recent research on educational curricula stresses the need for learning *Computational Thinking (CT)* skills in schools, as going beyond simple digital literacy to developing these CT skills becomes crucial (Menon et al., 2020). With this in mind, we design an activity to expose school children aged 9–14 years old to an optimization problem on networks; where we ask them to reason on a graph, and reasoning on a graph is a general skill of CT.

We choose solving the *minimum spanning tree (MST) problem* as the underlying objective of the task, that the children are not expected to be familiar with. An *instance* of the MST problem is defined on a *network* $G = (V, E)$ with a cost function on its edges, where V is the set of nodes, and $E \subseteq V \times V$ is the set of edges that connects pairs of nodes. A *solution* to this problem is a subset of edges $T \subseteq E$; and the goal of the problem is to find a T that connects all the nodes in the network to each other by some path, such that the total cost of the edges in T is minimized. This corresponds to finding a minimum-cost sub-network connecting all the nodes of a network. There exist algorithms like Jarník's algorithm and Kruskal's algorithm that are simple, and are guaranteed to find a minimum-cost solution (see Cormen et al., 2009).

As per the MST problem, our desired result is that after completing this task, a participant will be able to *correctly choose a subset of connections on a given network*, so that (i) all nodes are connected to each other by some path, and (ii) the total cost on these connections is minimized. We decompose this target skill into subskills by identifying the concepts underlying the MST problem, that build to a correct understanding for a participant about this problem, as follows:

Concept 1: (*existence* of a solution) Given an instance of the MST problem, there *exists* a solution if and only if there is a path between every pair of nodes in the network.

Concept 2: (*validity* of a solution) Given an MST instance and a solution, the solution is *valid* or *feasible* if and only if it connects every node pair in the network by some path.

Concept 3: (*correctness* of a solution) Given an MST instance and a solution, the solution is *correct* or *optimal* if and only if it is a feasible solution with the minimum cost.[II]

With an understanding of these concepts, a participant will be able to:

LG1: identify if there *exists* a solution for a given problem instance or not

LG2: identify if a given solution is *valid* for a (given) problem instance or not

LG3: identify if a given solution is *correct* for a (given) problem instance or not

These goals target the *understand* level of learning objectives in the revised Bloom's taxonomy (L. W. Anderson & Krathwohl, 2001).

3.2.2 Assessment of Learning

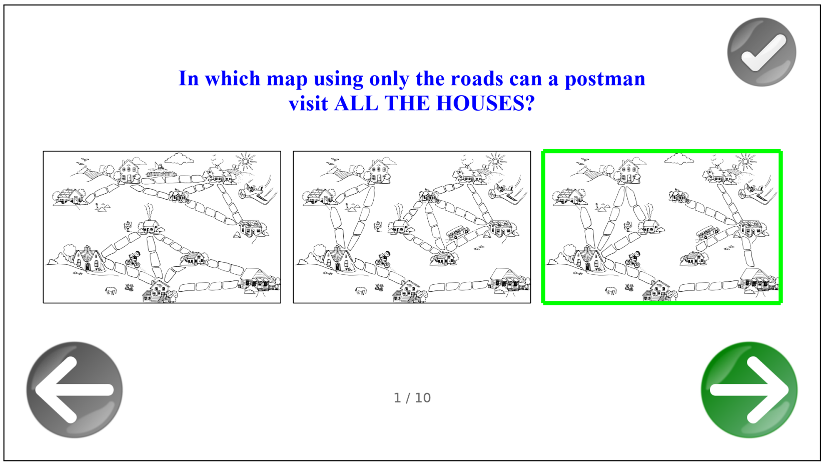
As evidence about the desired results, we measure the learning outcomes by comparing the responses of the participants in the pre-test and the post-test, to be solved individually by each participant. We design the tests as a sequence of multiple-choice questions with a single correct answer that assess the participants' understanding of the Concepts 1–3, by testing on the related subskills LG1–LG3.

We assess a participant's current understanding for the learning goal of identifying the existence of a solution (LG1) by checking whether the participant can distinguish the connected network from disconnected ones (as in Figure 3.1a). We evaluate the learning goal of identifying the validity of a solution (LG2), by checking if the participant can choose the network that has a valid solution marked on it (as in Figure 3.1b). Similarly, for the learning goal of identifying the correctness of a solution (LG3), we check if the participant can choose the network with the correct solution marked on it (as in Figure 3.1c).

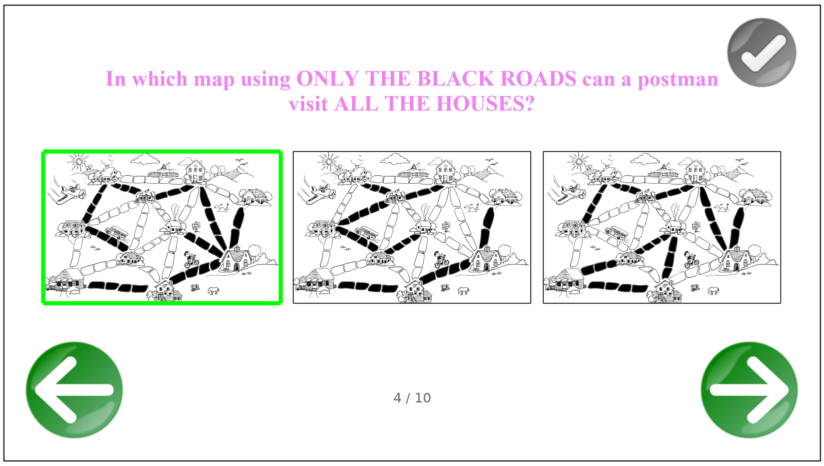
The pre-test and the post-test consist of 10 items in total: they contain three, three, and four items that measure a participants' understanding of the Concepts 1, 2, and 3, respectively. Figure 3.1 shows an example item for each concept and the associated test prompt (i.e. question), that is the same for the items of a concept. The tests are defined in a context other than the context of the collaborative activity: in the tests, the goal is to connect houses for a postman to travel between them; whereas it is connecting rare metal mines in Switzerland for the collaborative activity. The visuals are based on variants of the graphics in the *muddy city*[III] problem. The post-test is obtained by randomly shuffling the pre-test items and their response choices that belong to the same concept, and then mirroring the images given in the response choices vertically. The participants need to answer all items, and wait for the other to complete before submitting. All items were validated by experts of the domain and experts in education.

[II] Depending on the network, an optimal solution may not be unique; however, the minimum cost is unique.

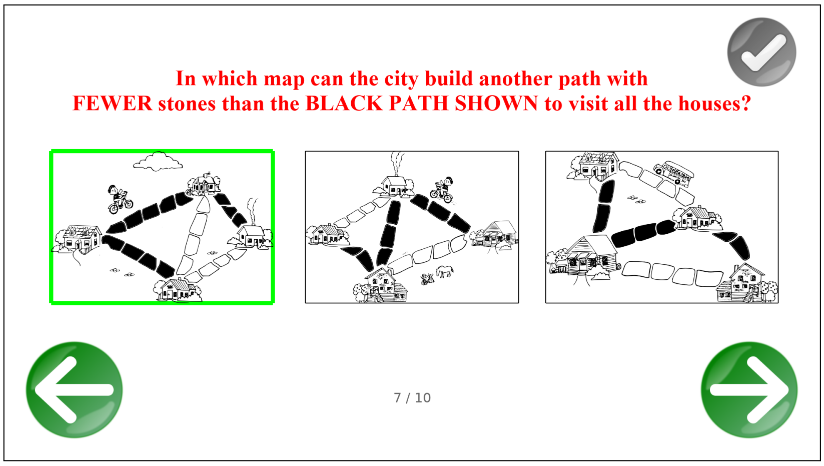
[III] https://classic.csunplugged.org/activities/minimal-spanning-trees, accessed October 2022, appeared in Bell et al. (2015).



(a) LG1—Example test item for Concept 1 on the existence of a solution



(b) LG2—Example test item for Concept 2 on the validity of a solution



(c) LG3—Example test item for Concept 3 on the correctness of a solution

Figure 3.1 – Example test items in the JUSThink scenario. The selected choices marked with a green border are the correct answers.

Table 3.2 – Sequence of activities in the JUSThink scenario

Activity	**What are the humans supposed to do?**	**What does the robot do?**	**Time (min)**
Welcome	Enter their name, age and gender on the screen	Welcome the participants, ask them for personal details	2
Tutorial	Listen to the robot	Introduce the task goal: connecting rare metal mines with railway tracks while spending as little money as possible	2
Pre-test	Answer a list of multiple-choice questions on the screen	Ask the participants to answer the pre-test items	≤ 10
Demo	Listen to the robot and follow the illustrations on the screen	Explain the game, the two complementary game views and their functionalities	3
Collaborative Activity	Work together to find a cheapest railway network connecting all the rare metal mines by: (i) drawing or erasing tracks that connect pairs of mines in turns (ii) submitting any agreed-upon solution to the robot for evaluation and feedback	At the submission of a solution: If the submitted solution is optimal, congratulate the participants and move to the post-test. Otherwise, reveal the cost difference between the submitted solution and an optimal one and motivate the participants to try harder. Point out the availability of the history of submitted solutions if the participants are not successful after several attempts.	≤ 25
Post-test	Answer a list of multiple-choice questions on the screen	Ask the participants to answer the post-test items	≤ 10
Questionnaire	Rate on a 5-point Likert scale a set of items about engagement, mutual understanding and the robot	Ask the participants to respond to the questionnaire items	≤ 5
Goodbye	See the robot wave goodbye	Thank the participants for their help, say goodbye	≤ 1

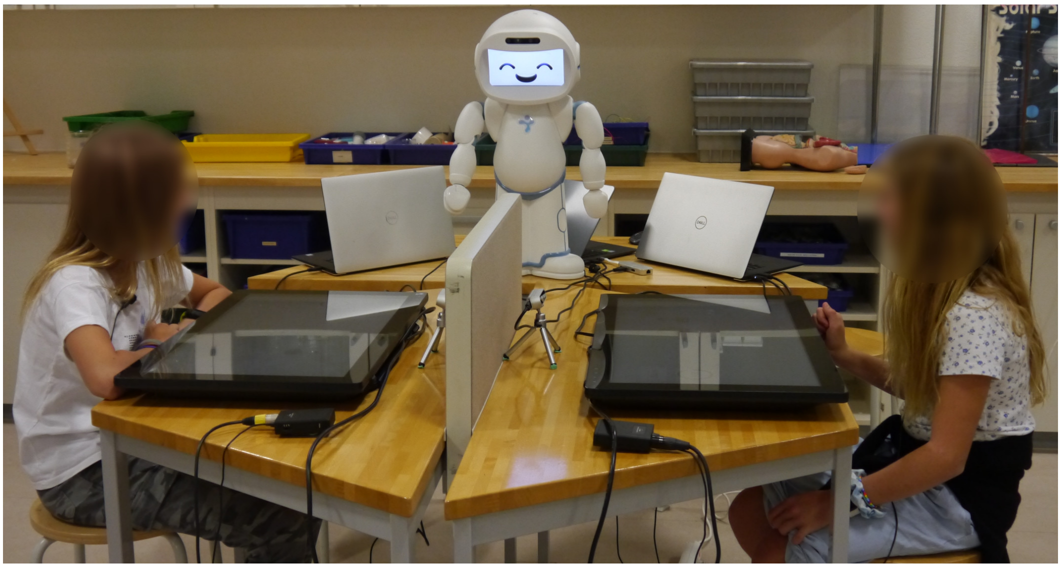


Figure 3.2 – An example scene of two children engaged in the JUSThink scenario in Study 1

3.2.3 Pedagogical Scenario

As an experience that can make the desired results happen, we design a human-human-robot interaction scenario where a humanoid robot orchestrates a sequence of activities as described in Table 3.2. The robot first introduces the scenario in the context of a game, which takes place on a fictional map of Switzerland with rare metal mines located in the mountains. The goal of the game is to build a railway network to help the miners go from any mine to any other, and spend as little money as possible to build these railways. In the *JUSThink (collaborative) activity*, a pair of children solve a problem instance together. Before and after the collaborative activity there are the *tests*, where the participants individually reply a series of test items that assess their understanding of the problem and its underlying concepts. At the end of the interaction, the participants individually answer a *questionnaire*.

The JUSThink Collaborative Activity

Setting A humanoid robot (QTrobot[IV]), plays the pretend role of the "CEO" for a rare metal mining company. It presents the activity to the children as a game, asking them to help it build a network to collect rare metals, by connecting rare metal mines one another with railway tracks. They are told to spend as little money as possible to build the tracks, which change in cost according to how they connect the rare metal mines. Thus, the underlying goal is to find a solution that minimizes the overall cost, i.e. an optimal solution for the given network.[V]

[IV] QTrobot by LuxAI SA, Luxembourg, https://luxai.com

[V] In the network used in the JUSThink activity, there exist 10 nodes: { 'Luzern', 'Interlaken', 'Montreux', 'Davos', 'Zermatt', 'Neuchâtel', 'Gallen', 'Bern', 'Zurich', 'Basel' } and 20 edges—see Figure 3.3. A description of the network, with the node labels (e.g. "Mount Luzern"), x, y position of a node, possible edges between the nodes, and edge costs, is publicly available online with the dataset, from the Zenodo Repository DOI: 10.5281/zenodo.4627104.

Table 3.3 – Questionnaire items in the JUSThink scenario

No	Item	Group	Category
1	I was trying very hard to find the best solution.	Cognitive at Task Level (IMI)	Task Engagement
2	It was important to me to do well at this task.		
3	I thought this activity was quite enjoyable.	Affective at Task Level (IMI)	
4	I enjoyed trying to find the best solution.		
5	I was trying very hard while discussing with my friend about the activity.	Cognitive at Social Level (IMI)	Social Engagement
6	It was important for me to discuss with my friend while finding the best solution.		
7	Discussions with my friend were quite interesting.	Affective at Social Level (IMI)	
8	I enjoyed discussing with my friend about the activity.		
9	I think I did pretty well at this activity.	Perceived Competence (IMI)	Own Competence
10	I am satisfied with my performance at this task.		
11	I felt tense while doing this activity.	Pressure/Tension (IMI)	Stress
12	I think my friend understood my instructions very well.	Cognitive (IMI-like)	Mutual Understanding
13	I think my friend understood my emotions very well.	Affective (IMI-like)	
14	I think the robot is competent (capable).	Robot (Godspeed)	Robot
15	I think the robot is intelligent.		
16	I think the robot is friendly.		
17	I think the robot is likeable.		
18	I think the robot is distracting.	Robot (Godspeed-like)	Robot Behavior
19	I think the robot should give more useful feedback.		
20	I liked the robot.		
21	I would like to play the same game with the same friend.	Game and Friend	
22	I would like to play the same game with another friend.		
23	I knew my friend well.	Known Friend	
24	How many minutes do you think you spent on the part where you played with your friend to find the best solution?	Perception of Time	

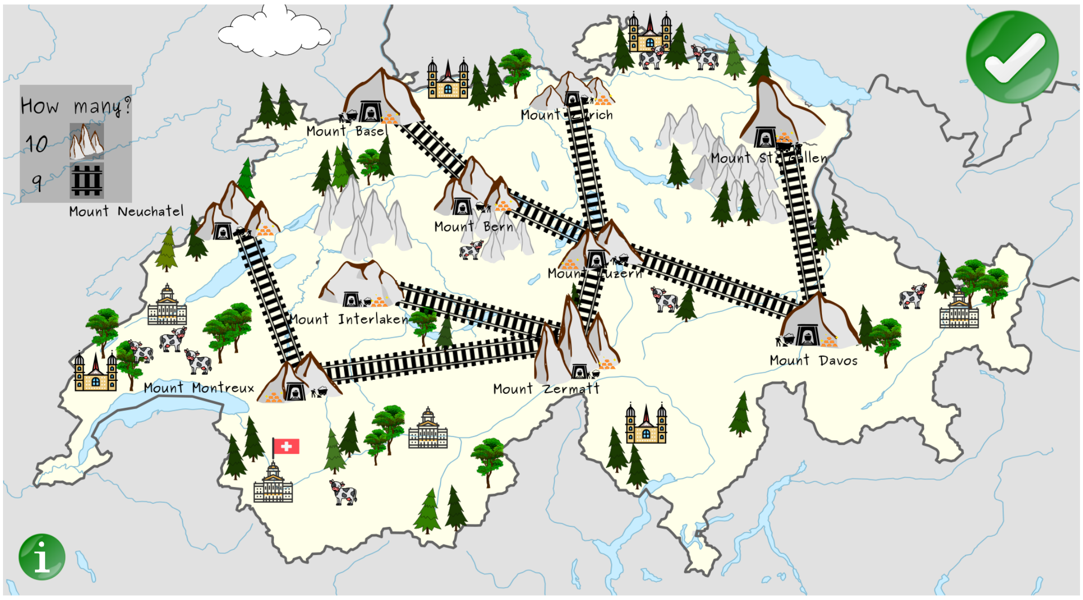
Children participate in teams of two to collaboratively construct a solution, by drawing and erasing tracks. Once all rare metals mines are connected to each other, they can submit their solution to the robot for evaluation. They must submit their solution together, the robot then reveals whether their solution is an optimal solution or, if not, how far it is from an optimal solution (in terms of its cost). In the latter case, children are also encouraged by the robot to try again. They can submit a feasible/valid solution as many times as they want until the allotted time for the activity is over.

Setup Two children sit across each other, separated by a barrier. A touchscreen is placed horizontally in front of each child. Children can see each other, but cannot see the other's screen, as in Figure 3.2. They are encouraged by the robot to verbally interact with each other, and work together to construct a solution to the activity. The screens display two different views of the current solution to the children. One view is an *abstract view*, where the rare metal mines are represented as nodes, and the railway tracks that connect them as edges (see Figure 3.3b). The other view, or the *visual view*, represents the rare metal mines and railway tracks with images (see Figure 3.3a). A child in the abstract view can see the cost of built edges, but cannot act upon the network. That is, after an edge is built, its cost is shown in this view regardless of whether it was built or removed. Conversely, in the visual view, a child can add or delete an edge, which is a railway track, but cannot see its cost. The views of the children are swapped every two *edit actions*, which is any addition or deletion of an edge. Hence, after every two edit actions, the child that was in the abstract view moves to the visual view and vice versa. A *turn* is thus the time interval between two view swaps, i.e. in which one child is in the visual view and the other child is in the abstract view. This design serves as a *collaboration script*, that aims at encouraging the children to collaborate.

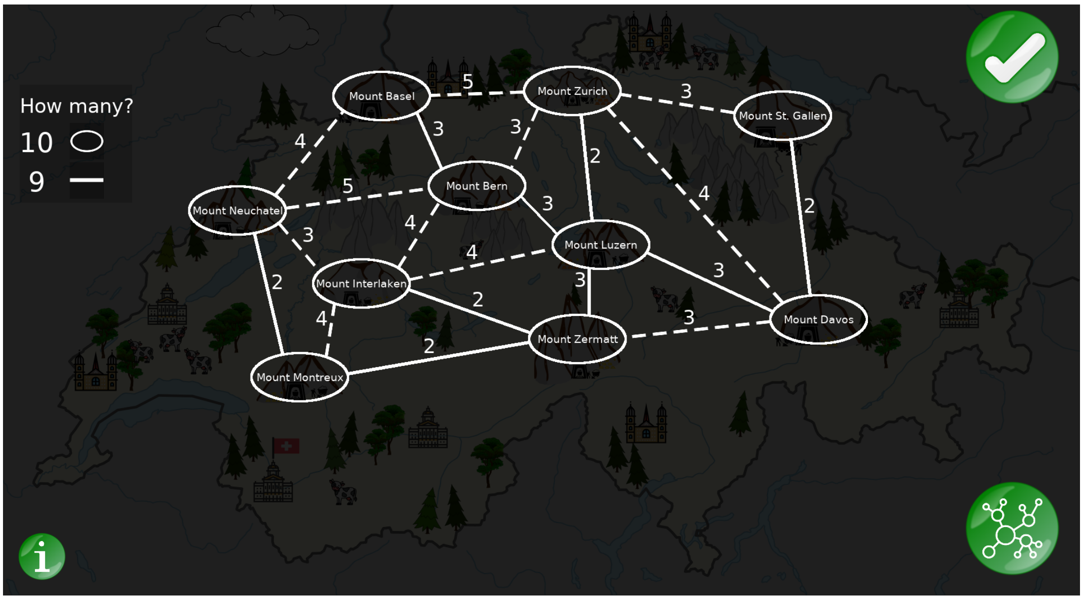
Questionnaire

In order to get an insight on the impressions of the participants regarding their experience, at the end of the scenario, each participant individually answers 24 items, listed in Table 3.3. 11 items belong to or are derived from the *Intrinsic Motivation Inventory (IMI)* (Ryan & Deci, 2000), which "is a multidimensional measurement device intended to assess participants' subjective experience related to a target activity in laboratory experiments"[VI]. There are eight items on engagement (Items 1–8); two items on own competence and one on stress (9–11). Two items refer to mutual understanding (12–13), that are complemented by three other on the relationship with the team partner (21–23). Four items are from the *Godspeed questionnaire* that is widely used in HRI (Bartneck, Kulić, et al., 2009), and assess the perception of the robot, its behavior and helpfulness. Item 24 is on the perception of time elapsed (with options), that can be indicative of how the participants perceived their engagement; as "time flies" when you are highly engaged and have been enjoying an activity. Items 1–23 are on a 5-point Likert scale (Likert, 1932); and Item 24 is on an 8-point scale with options 5, 10, ..., 40 min.

[VI]From http://selfdeterminationtheory.org/intrinsic-motivation-inventory, accessed October 2022



(a) An example visual view: only the selected connections (i.e. the current solution) can be seen. In this view, new connections can be selected (by a press-drag-release gesture from a mine to another mine, if there is a possible connection between them), and the existing connections can be removed (by a long-press gesture; as long as it does not result in two paths with tracks that are disconnected from each other). The possible connections that can be selected are revealed with a press gesture on a mine.



(b) The corresponding abstract view, where all possible connections (edges) were built and hence can be seen along with their costs. In this view, edges cannot be selected or removed. The selected edges are indicated as solid line segments, while those that are no longer selected are dashed. These built edges with their costs are shown until the solution is submitted. The built edges and the total cost of the selection for each submitted solution can be displayed via the history button on the bottom right. Also, a previous solution can be brought back, erasing the current solution, through this history window.

Figure 3.3 – An example game scene from the JUSThink activity, showing the two complementary views. The selected solution is an optimal solution.

Figure 3.4 – Example scenes illustrating the robot's behavior in the JUSThink scenario

Items concerning the perception of the robot refer to its competence, intelligence, friendliness and like-ability, and are complemented by behavioral items on being distracting and giving useful feedback. Engagement here entails the effort put in for solving the task (cognitive engagement at task level) as well as for discussions with the partner to solve the given problem (cognitive engagement at social level). It also includes the enjoyment that the participants had with regard to the task (affective engagement at task level) as well as discussions with their partner (affective engagement at social level). Similarly, mutual understanding is measured both in terms of understanding of their instructions to each other for solving the task (cognitive) and understanding of each others' emotions (affective).

3.2.4 Robot's Role and Behavior

The robot's role in the JUSThink scenario is twofold:

1. mediate and automate the entire interaction (see Table 3.2); pause the participants' applications, give instructions, and move from an activity to the next upon its completion,
2. give feedback on the submitted solutions, provide basic hints (as mentioned in Table 3.2), and support the interaction through minimal expressive behaviors. The behaviors include verbal support, using participants' names, and making facial expressions and supportive gestures. Some of the behaviors can be seen in Figure 3.4. They are pre-programmed and the same for all the teams. For instance, the robot made observing and thinking gestures that are triggered at regular intervals, and gave the same type of feedback: "you are away from the minimum by x Francs. Let's try again.".

3.2.5 How the JUSThink Activity Elicits Dialogue

This activity is particularly suited to study mutual understanding, as it is designed in such a way to create *interdependence*, i.e. a mutual reliance to further the task, between the participants. To achieve the task's objectives, the participants need to build and converge on a shared solution via taking *joint actions*, that generates a *dialogue* on how they build a mutual understanding about what is a correct solution and how to construct this solution. This interdependence requires participants to align with each other on multiple levels, e.g. how to refer to the environment and how to represent the activity, in order to succeed.

Concretely, the activity enables:

1. *(Turn-taking) Swapping and visual view control*: Since a turn changes every two edit actions, if a participant has a particular action they want to take, they have to either wait for their turn in the visual view to implement the desired change, or instruct the other participant. Here, we can consider an idealized perspective on the activity: at a given time, the participant in the abstract view is an Instruction Giver (IG) who describes their instructions for the task by using specific referring expressions, and the other (in visual view) is the Instruction Follower (IF) who executes the action (this is akin to the Map Task (A. H. Anderson et al., 1991)). The activity design creates a frequent swapping of views. This aims to discourage participants from working in isolation or in fixed roles of IG and IF, which could potentially happen in collaborative tasks.

2. *Routine expressions and alignment in the task*: Since the robot uses brief and general instructions to present the activity and its goal, the participants must figure out for themselves the way to approach the activity. The task-specific referents, which are the names of the rare metal mines (named after cities in Switzerland, a multi-lingual country), are potentially unfamiliar to participants. Participants must refer to the task, and then align with the other to form routine expressions. By aligning, they establish a shared lexicon, and thus align their representations of the activity with each other.

3. *Submission of solutions*: Since the participants have to submit their solution together by pressing the "submit" button that is present in both views, they have to, at least, align in terms of their intent to submit. Alternatively, one participant must convince the other to reach a common intent.

3.2.6 Implementation and Setup for Data Collection

Equipment Figure 3.5 shows the physical layout of for the JUSThink scenario. Table 3.4 presents the list of equipment used in Study 1. The interaction is recorded by three cameras: one environment camera filming the whole scene and two cameras each focused on a child's face. Two computers, connected to the two touchscreens and to the robot's local network, manage the activity. The face cameras are connected to a third computer to alleviate the

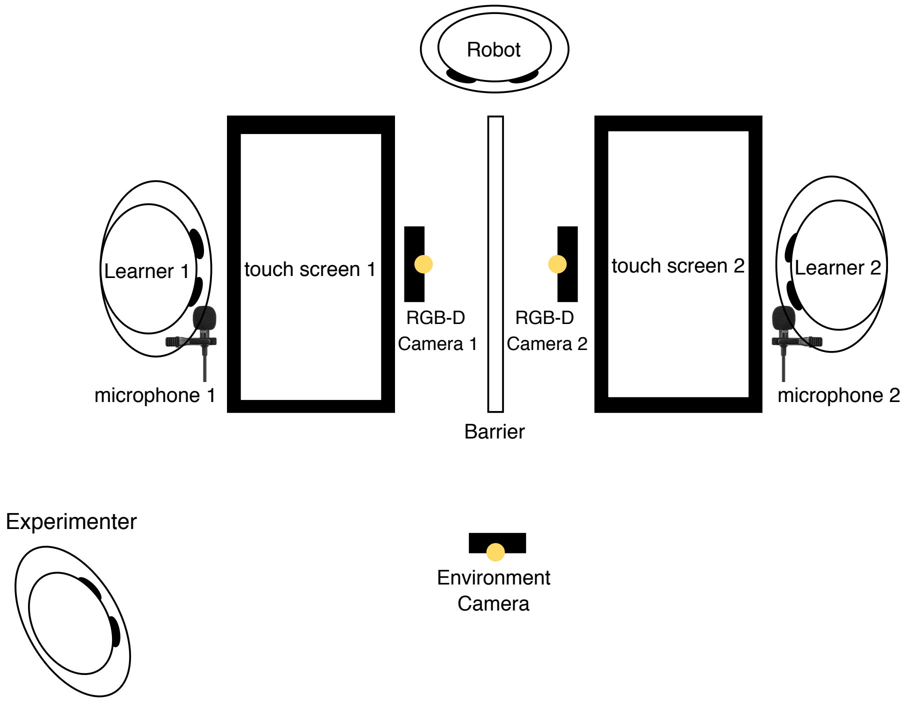


Figure 3.5 – Physical setup of the JUSThink scenario

burden of bandwidth in the local network. Audio is recorded as two mono audio channels synchronized to each other by the environment camera, with one lavalier microphone per channel. The microphones are clipped onto the children's shirts.

Software Each participant interacts with an instance of the participant application that is written in Python and uses the `pyglet` windowing and multimedia library for visualization. A separate instance of the application is run for each participant in a team. The robot behavior application is also developed in Python and governs what the robot does and when. The applications communicate via the Robot Operating System (ROS). We use the `rosbag` package of ROS for logging. The applications are implemented and tested on 64-bit Ubuntu 18.04 (Bionic Beaver) with Python 2.7 and ROS Melodic Morenia.

3.3 Materials

3.3.1 User Study 1: JUSThink Human-Human Observational Study

The study was conducted with 96 children aged 9 to 12 years. The children participated in teams of two: the teams were formed randomly, without considering the gender, nationality, or the mother tongue. They include mixed and same gender pairs, and this information is available but not used in our analyses. Due to technical issues during the experiment, 18

Table 3.4 – List of equipment used in User Study 1 for the JUSThink scenario

Equipment	Model and Description	Quantity
Humanoid Robot	*QTrobot* by LuxAI, 2019 version that has a Raspberry Pi 4. It creates a local network by with Wi-Fi Access Point with the Raspberry Pi. We use its ROS API by LuxAI to control its speech, gestures and facial expressions. The text-to-speech service is based on software by Acapela: we use a child voice "Ella".	1
Touchscreens	*IIYAMA ProLite T2435MSC-B2 24-inch monitor*; with capacitive 10 point multi-touch technology. Its hinged stand allows laying it down completely flat for use as a giant table. It works as a USB touch input device via Human Interface Device (HID) protocol, on Silicon Integrated Systems (SiS) HID Touch Controller by USBest Technology.	2
Environment Camera	*Canon XA25 Professional Camcoder*. It has two phantom-powered XLR audio inputs, and synchronizes the audio and video signals. We use a Libec TH-650DV Tripod System to adjust the camera's height, pan and tilt.	1
Microphones	*Audio-Technica PRO 70 Cardioid Condenser Lavalier Microphone*. They are connected to the XLR inputs of the camcoder.	2
Face Cameras	*Intel RealSense Depth Camera D435*. It captures RGB-D images. We use its ROS API maintained by Intel to access the images.	2

participants are omitted from the analysis, resulting in a dataset of 78 children in 39 teams (41 females: $M = 10.3, SD = 0.75$; and 37 males: $M = 10.4, SD = 0.60$).

The study was conducted in two international schools in Switzerland over the span of two weeks, where the medium of education is in English, and hence students are proficient in English. A session lasted approximately 50 minutes. The children are asked to speak in English. There were always two experimenters available in the room, but the system was fully automated to require the least intervention by the experimenters. While the participants reported to be familiar with robots as a part of their curriculum and science, technology, engineering, and math (STEM) activities, they did not have a prior experience with the robot used in this study. Ethical approval was granted by the EPFL Human Research Ethics Committee (HREC), No. 051-2019/05.09.2019.

3.3.2 The JUSThink Dataset

The *JUSThink Dataset* contains data of $N = 39$ teams participating in the JUSThink scenario, for about 15.9 hours of data in the collaborative activity: the duration of the collaborative activity has $M = 24.5$ min ($SD = 4.9$ min; from 11.2 to 35.9 min). The dataset consists of:

1. *Recorded audio files*: as separate channels with one channel per child

2. *Event log files*: Timestamped events that include touch events, edit actions, submitted solutions, and commands to the robot
3. *Pre-test and post-test responses*: Individual responses for 10 items per test (as in Figure 3.1)
4. *Questionnaire responses*: Individual responses of each child for 24 items (as in Table 3.3)
5. *Recorded video files from the environment camera*: of the overall scene (see Figure 3.2)
6. *Recorded video files from the face cameras*: Video focused on the face of each child

In Part I of this thesis, we analyze and derive from parts of this dataset; in Part II, we do not refer to this dataset and continue on with other studies. Specifically, in this chapter, we use the event log files to measure the teams' performance in the task, pre- and post-test responses to measure the learning outcomes, and (only in this chapter) questionnaire responses for the participants' self-assessment. In Chapter 4, we additionally use the recorded audio files to obtain transcripts. In Chapter 5, we complement these transcripts by linking them with the edit actions from the event logs. Other lines of research have also used this dataset, e.g. Nasir et al. (2022a, 2022b, 2021) utilized the face camera recordings to infer emotional expressions of the participants, and audio recordings to extract features on voice activity.

3.3.3 Speech Activity Levels from the PE-HRI Dataset

For speech activity levels, we made use of the *PE-HRI Dataset* that is available from the Zenodo Repository DOI: 10.5281/zenodo.4633092. It is a derived dataset that processed the original JUSThink Dataset described above, for the purposes of extracting engagement related multi-modal features for the teams' behaviors. The number of teams in the PE-HRI Dataset is 34, while the original dataset has 39 teams: five teams had been removed due to incomplete data. See Nasir et al. (2022b), Nasir et al. (2021) for further details on the PE-HRI Dataset.

The speech activity (i.e. the feature `voice_activity` in the dataset) is measured as the percentage of time that a team is speaking over the entire duration of the task, as averaged per participant. The speaking or not was determined for audio frames of size 30 ms by the Google WebRTC Voice Activity Detector (VAD) for each audio channel of the two channels per team (associated with each participant of the team), via using the Python package `py-webrtcvad`. The aggressiveness mode to filter out non-speech was set to the highest level (i.e. 3, the most aggressive): i.e. the false positives for speech activity are minimized, to the extent that this tool permitted.

3.4 Methods

3.4.1 Measuring Performance in the Collaborative Activity

For task performance, we consider relative error (error_S) of a submitted solution S, which is the scaled difference of the cost of the solution compared to an optimal solution, i.e.:

$$\mathsf{error}_S \triangleq \frac{cost - optimal\ cost}{optimal\ cost} \tag{3.1}$$

Equation 3.1 – Error formula used to measure task performance in the JUSThink activity

It indicates how close a solution is to an optimal solution in terms of its cost: $\mathsf{error} = 0\%$ means the solution is an optimal solution. We use error, the lowest error a team achieves throughout their collaboration, to measure the overall performance of the team in the task.

3.4.2 Measuring Learning Outcomes

Learning measures commonly build upon the difference between the post-test and pre-test results, e.g. in Sangin et al. (2011); which indicate how much a participant's knowledge on the subject has changed due to the activity. We measure the learning outcomes on the basis of the relative learning gain (learn_P) of a participant P, which essentially is the difference between pre-test and post-test, normalized by the margin of improvement or decline (as in Sangin et al., 2011). It is computed from *pre-test* and *post-test scores* (i.e. *pre* and *post*) as:

$$\mathsf{learn}_P \triangleq \begin{cases} \frac{post - pre}{max\ score - pre} & , post > pre \\ \frac{post - pre}{pre} & , post \leq pre \end{cases} \tag{3.2}$$

Equation 3.2 – Learning gain formula to measure learning outcomes for the JUSThink activity

It indicates how much the participant learned as a fraction of how much the participant could have learned, and ranges from -100% that indicates the maximum decline (with a 0% score in the post-test), to 100% that means maximum possible improvement (by achieving a 100% score in the post-test, from any score in the pre-test). We use learn, the average relative learning gain of both participants in a team, to measure the team's learning outcomes.

3.5 Results and Discussion

3.5.1 RQ1 on Eliciting Dialogue

The speech activity levels are quite high ($M = 44.6\%$, $SD = 9.11$, ranging from 28.2% to 67.9%), i.e. there was a participant that was speaking on average about half of the time during the activity. Therefore, the data supports our hypothesis H1.1: the activity successfully elicited

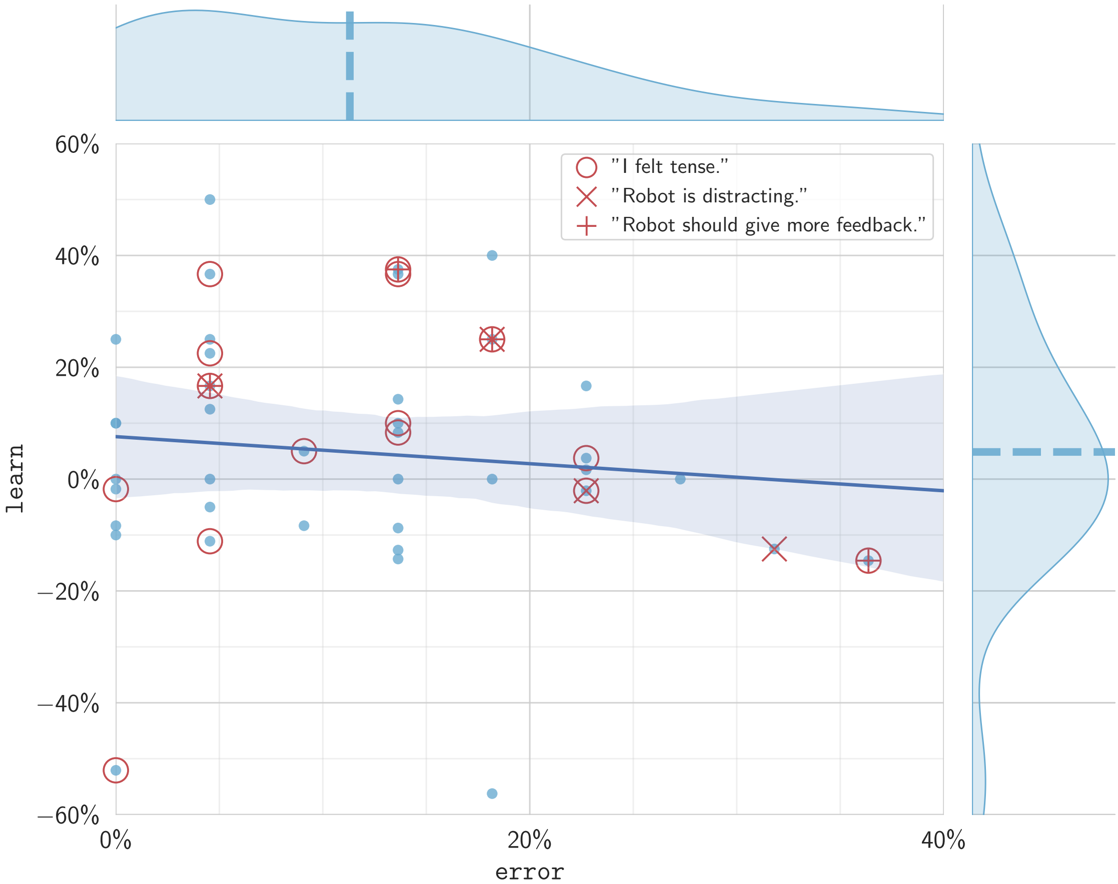


Figure 3.6 – Performance vs. learning plot of the JUSThink Dataset ($N = 39$ teams). The mean of the measures are shown as dashed lines, with the fit of a univariate kernel density estimate. Fit of a linear regression model is indicated with a solid line and a 95% confidence interval. The teams that reported feeling tense, found the robot distracting, or believed the robot should give more useful feedback are marked; by annotating the teams with the team-average rating ≥ 4 for the Items 11, 18, and 19 in the questionnaire, respectively.

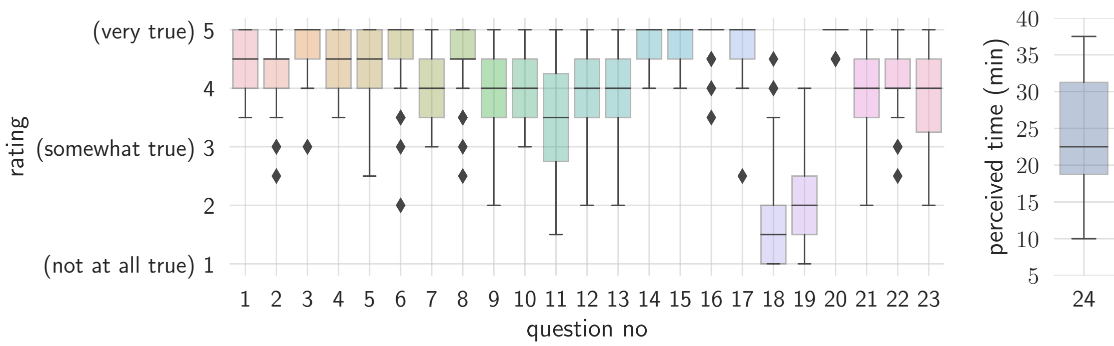


Figure 3.7 – The teams' responses to the questionnaire items in the JUSThink Dataset ($N = 39$ teams). The boxplots with whiskers with maximum 1.5 IQR show the distributions of the ratings averaged per team for each item. The items are listed in Table 3.3.

dialogue in terms of speech activity levels.

Performance in the task as measured by error has a weak negative correlation coefficient with the speech activity levels (Spearman's $\rho = -0.26, p = 0.14$).[VII] Furthermore, the learning outcomes measured by learn has a very weak correlation coefficient with the speech activity levels (Spearman's $\rho = 0.16, p = 0.37$). Therefore, the data does not support a monotonic relationship between voice activity levels and performance in the task (H1.1) or learning (H1.2). Then, the amount of verbal interaction by itself is not sufficient to determine the performance or learning; the content of what is said and other factors influence these end results.

3.5.2 RQ2 on Learning and its Link with Performance

There is a spectrum of learning outcomes for the participants, and their average is low (learn_P has $M(78) = 4.9\%, SD = 28.8\%$, and ranges from -75% to 100%). A Wilcoxon signed-rank test shows that the participants' scores in the post-test are not statistically significantly different than in the pre-test ($W(78) = 1237.0$, $p = .44$)[VIII], and the effect size is negligible (Cliff's $\delta(78) = -.08$)[IX]. Thus the data does not support H2.1 for positive learning outcomes. We believe that the context in our tests being different from the task may have lead to low learning gains.

Figure 3.6 shows how the teams are distributed in the learning outcomes vs. task performance space (as measured by the team-averaged learn, and the team's error, respectively). The (averaged) learn is low ($M(39) = 4.9\%, SD = 21.8\%$; ranging from -56.3% to 50.0%). Only 8 teams out of the 39 ($\approx 21\%$) found an optimal solution in the collaborative activity (i.e. achieved error $= 0\%$): the error values have average 11.3% ($SD = 9.6\%$, from 0% to 36.4%). Learning outcomes measured by learn has a very weak correlation coefficient with error (Spearman's $\rho = -0.09, p = 0.60$). Therefore, the data does not support H2.2: we can not conclude that there is a significant relationship between how well a team performs and how much they learn.

A possible explanation for the low learning gains, as well as the lack of evidence for a correlation between performance and learning, is that our tests rely on a *far transfer* from the task to the test. Instead of a *near transfer* to closely related contexts, far transfer requires abstraction and adaptation between different contexts, which is not spontaneous (Perkins & Salomon, 1992).

[VII] The magnitude of Spearman's ρ can be interpreted by: $.00-.19$ "very weak", $.20-.39$ "weak", $.40-.59$ "moderate", $.60-.79$ "strong", and $\geq .80$ "very strong" (Evans, 1996).

[VIII] Scores in the pre-test and the pre-test are not normally distributed (Shapiro-Wilk's $W(78) = .94$, $p = .0013$, and $W(78) = .93$, $p = .0002$, respectively).

[IX] The magnitude of Cliff's Delta (δ) can be interpreted via thresholds $|\delta| < .147$ "negligible", $|\delta| < .33$ "small", $|\delta| < .474$ "medium", and otherwise "large" (Romano et al., 2006).

3.5.3 RQ3 on Participants' Self-Assessment

Figure 3.7 shows the distribution of the team-averaged ratings for the items in the questionnaire. Participants (at a team level) perceived themselves as highly engaged at both task and social levels (*Mdn*(39) = 4 to 5 for all Items 1–8). Furthermore, they gave high ratings to the mutual understanding of their instructions and emotions by their partners (with *Mdn* = 4 for both Items 12 and 13). Thus, the data supports H3.1 for an overall positive perception of engagement and mutual understanding.

The participants rated the robot very highly on its competence, intelligence, friendliness, and like-ability; despite its simple behavior and limited support (*Mdn*(39) = 5 for all Items 14–17). Since the interaction lasted about 50 min; there was ample time for the participants to form an opinion on the characteristics of the robot and its limitations. Although most of the teams were not successful in finding an optimal solution, we see that the majority of them still think that the robot does not need to give more useful feedback (*Mdn* = 2 for Item 19). Furthermore, very few participants found the robot distracting (*Mdn* = 1.5 for Item 18; only 4 teams rated ≥ 4, i.e. $\approx 10\%$ of the teams). Hence, contrary to our expectations, the data presents evidence against H3.2, that expected a rather negative assessment about the robot in these regards.

The participants reported feeling somewhat tense during the activity (*Mdn*(39) = 3.5 for Item 11), and tended to perceive themselves as competent (*Mdn* = 4 for Items 9 and 10). The latter is interesting because many teams were not able to find an optimal solution, and even then they tended to rate their competence highly. This perception may be then derived due to high cognitive engagement, rather than success in the task. Furthermore, participants expressed a desire to play the game with the same or a new partner (*Mdn* = 4 for Items 21 and 22). Lastly, the perception of time spent in the activity as a measure of engagement is comparable to the actual time the game ($M = 24.6, SD = 7.9$ min for Item 24, while activity duration has $M = 24.5, SD = 4.9$ min).

The positive perception of the robot about its competence and intelligence can be interpreted as follows: in addition to social desirability bias (e.g. see Nederhof, 1985; Richman et al., 1999, especially in our case with the robot asking to fill the questionnaire that has questions about itself), the humanoid form of the robot, and its human-like behavior with speech, gestures and facial expressions promote anthropomorphic representations about the robot: this can make it more likely to be perceived as having a high level of agency and intelligence (Epley et al., 2007; W. J. King and Ohya, 1996; also see Bartneck, Kanda, et al., 2009; Moro et al., 2019).

3.5.4 RQ4 on the Link Between Performance and Learning with Assessment

Task Performance and Self-Assessment (H4.1)

The task performance measure error has moderate negative correlations with the participants' perceived mutual understanding of their instructions and emotions by their partners (Spearman's $\rho(39) = -0.47, p = .002$ and $\rho = -0.44, p = .006$; for Items 12 and 13, respectively): that is, the better the teams performed (i.e. lower the error), the higher they tended to rate their mutual understanding. We obtain a similar result when we compare the teams that performed well (19 teams with error $< M(\text{error}) = 11.3\%$) with those that performed badly (the remaining 20 teams): Kruskal-Wallis tests show that ratings of well-performing teams are statistically significantly different from badly-performing teams ($H(19, 20) = 8.07, p = .005$ and $H = 4.49, p = .03$ for Items 12 and 13): well-performing teams rated their mutual understanding higher, with large and medium effect sizes (Cliff's $\delta = -.52$ and $-.39$).

There is a weak negative correlation between the error and finding the activity enjoyable, as an indicator of affective engagement at task level (Spearman's $\rho = -0.37, p = .02$ for Item 3). Similarly, error weakly correlated with giving importance to the discussion with the partner while finding an optimal solution, that indicates cognitive engagement at a social level ($\rho = -0.35, p = .03$ for Item 6). In addition, the participants that achieved lower error also tended to believe that they did well, with a moderate correlation ($\rho = -0.40, p = .01$ for Item 9): this indicates a higher perceived competence for the better performing teams. Indeed, as shown in Figure 3.6, teams that reported high levels of stress (15 teams with rating ≥ 4, i.e. $\approx 38\%$ of the teams), the robot being distracting (4 teams), or wished for more useful feedback (4 teams) are dispersed throughout the plot, regardless of their performance (or learning outcomes). For this reason, we deem H4.1 to be only partially supported by our findings, and specifically to be rejected concerning the perception of the robot.

Learning Outcomes and Self-Assessment (H4.2)

There are only weak to negligible correlation coefficients and no evidence for any significant correlations between the learning outcomes and the self-assessment ratings (e.g. Spearman's $\rho(39) = 0.28, p = .09$ and $\rho = 0.03, p = .86$ for Items 12 and 13 on mutual understanding; and $\rho = 0.2, p = .18$ for Item 9 on competence). By comparing the teams' ratings with respect to their learn, we observe that the ratings are not statistically significantly different for teams with positive learning outcomes (19 teams with learn $> 0\%$) vs. others (20 teams) by Kruskal-Wallis H tests (e.g. $H = 0.81, p = .37$ and $H = 0.007, p = .93$ for Items 12 and 13 on mutual understanding; and $H = 3.23, p = .07$ for Item 9 on competence). Thus the data does not support H4.2.

Participants seem to have based their assessment of self-competence on apparent representations of learning and achievement, e.g. success-failure in the game, rather than the tests which we used to measure their learning. Similar results were reported e.g. in Fry (1976), who observed "subjects who experienced success made significantly greater gains in positive self-assessments, and failure subjects made significantly greater gains in negative self-assessments". Since the participants did not receive feedback on their scores in the tests, they did not experience their success in the tests: therefore their ratings did not reflect how well they did in the tests, but rather how they did in the collaborative activity.

3.6 Conclusion

In Part I of this thesis, we are interested in developing abilities for humanoid robots to assess how humans collaborate as they solve a problem together; where they build a mutual understanding of each other about the problem, in order to converge on a joint solution. Our context is learning activities, in which what the humans say and do is strongly tied to how they perform, and subsequently what they will ultimately learn from the activity. As our first step towards investigating this relationship, in this chapter, we described a novel robot-mediated human-human collaborative problem solving activity, that is designed to elicit dialogue on how humans build a mutual understanding as they solve a problem together with joint actions. We detailed our data collection with a user study in two schools, and the dataset we created from this user study that involves 39 teams of two children aged 9–12 (78 children in total) participating in this activity. Then, we presented an overall evaluation of this dataset in terms of the participants' performance in the task, learning outcomes, and subjective perspectives.

Results show that the activity was successful in eliciting dialogue, as indicated by the high levels of speech activity. Results about the participants' self-assessment indicate that they perceived the robot very positively, as highly competent etc.; regardless of their performance in the task, and despite the robots simple behavior and limited support. Furthermore, although performance correlated with participants' self-assessment of their competence and mutual understanding, we saw no evidence for correlations of these assessments with learning. Therefore, in the next chapters, we shift our focus from subjective perspectives to the objective phenomena, specifically to what was said and done. These can be complemented by the subjective assessments in order to support conclusions or explain certain attitudes of the participants towards each other, the activity and the robot. However, subjective assessments by themselves need not be sufficient to reflect how and how well the participants build a mutual understanding, how they perform and whether they learn or not; as also seen by the results in this chapter.

Results show that the how children perform in the task does not directly translate to whether or how much they learn from the activity. Thus, we look upon performance and learning as "multi-layered" goals, that are realized as the outcomes of the deep, complex and tangled relationship between what the children say and what they do: the immediate, performance goals, and the deeper, learning goals that stem from it (or not). Therefore, next, we look into the *process* of how the interaction evolves into success or failure in the task, and whether it results in learning or not. For this purpose, in the next chapter (Chapter 4), we examine what was said in their dialogue: in terms of whether and how they align within their dialogue, and how this relates to the outcomes about their performance and learning. Then, in Chapter 5, we look into how their dialogue relates to what was done i.e. their actions: in terms of whether and how their actions align with their dialogue, and whether this indicates the outcomes.

4 What We Say as Cues for Human-Human Mutual Understanding

4.1 Introduction

In this chapter, we propose novel rule-based algorithms to automatically and empirically measure collaboration and mutual understanding, by studying the *alignment* (i.e. the development of shared representations of interlocutors at different linguistic levels (Pickering & Garrod, 2004, 2006)) in the dialogue between the children resulting from the activity we described in Chapter 3: the JUSThink activity. Results of Chapter 3 showed that the activity was successful in eliciting dialogue, as indicated by the high levels of *speech activity*: but, what kind of dialogue is it? Can the robot do more than monitoring the speech activity? This activity results in a *situated dialogue*, where the participants share the time and space, and the interaction is about their common environment: the dialogue has an *interdependence* on the immediate environment. Thus, in this chapter, we delve into the *speech content*, i.e. what is said. To study alignment, we focus on the formation of *expressions related to the activity* that the robot can specifically aim to detect, as focusing on these expressions allows us to target information very specific to making progress in the task. Using these expressions we consider two alignment contexts: in this chapter, we study (i) *lexical alignment* (what was said), i.e. alignment at a lexical level. Then, in Chapter 5, we study (ii) *behavioral alignment* (what was done), i.e. a new alignment context we propose to mean when instructions provided by one interlocutor are either followed or not followed with physical actions by the other interlocutor.[I]

Results highlight that all teams were (lexically) aligned to some degree, regardless of their performance in the task or the learning outcomes. The better-performing teams tended to align earlier, while this pattern is inconclusive for teams with positive learning outcomes. Inspecting the distributions of the establishment times (i.e. *when* alignment occurs) together with their proposed solutions gave further insight into the collaboration processes, specifically revealing a *collaborative period* that most of the teams came up with their best solutions. This is important because this period is automatically inferred by our alignment measure that pro-

[I]We make this dataset and tools to study alignment in children's dialogues publicly available (containing transcripts, action logs and alignment tools), from the Zenodo Repositories DOI: 10.5281/zenodo.4627104 for the dataset, and DOI: 10.5281/zenodo.6974562 for the tools.

cesses only the dialogue content, without consulting to the actions taken by the participants. This highlights that in a situated educational activity, focusing simply on expressions related to the task can still give good insights into the nuances of collaboration.

This work corresponds to the following publication:

> **Norman***, **U.**, Dinkar*, T., Bruno, B., & Clavel, C. (2022). Studying alignment in a collaborative learning activity via automatic methods: the link between what we say and do. *Dialogue & Discourse*, *13*(2), 1–48. https://doi.org/10.5210/dad.2022.201. **Contributed equally to this work.* (Norman, Dinkar, et al., 2022)

The current work stemmed from a secondment of the author of this thesis to Télécom Paris in France, under the supervision of Prof. Chloé Clavel and in collaboration with Tanvi Dinkar, within the scope of the H2020 ANIMATAS Project; while being hosted at Sorbonne University. Within this work, the author of this thesis implemented the tools, prepared a derived dataset that involved transcribing audio from the JUSThink dataset, and together with Tanvi Dinkar, contributed to the data analysis and the interpretation of the results.

Other lines of research cultivated on the derived dataset and analysis described here: e.g. Norman, Dinkar, et al. (2022) and Dinkar (2022) also investigated spontaneous speech phenomena (e.g. “um", “uh" ...) as informative cues in the dialogue, and how these were used in the process of alignment. For instance, the authors observed that for most of the teams, the fillers tended to occur around the establishment times. Then, building on the methodology in Chapter 5, they found that well-performing teams verbalized the marker “oh" more when they were behaviorally aligned, compared to other times in the dialogue.

4.1.1 Research Question and Hypothesis

We focus on the *task-specific referents*, that interlocutors minimally require to succeed in the task. Therefore, we restrict the possible referring expressions to ones that contain task-specific referents, in particular, to the objects that the interlocutors are explicitly given on the map (see Figure 3.3). Interlocutors only need this terminology with certain function words to progress in the task (e.g. “*Montreux* to *Basel*"). We believe that this design choice is particularly suited to study alignment in this type of activity, as the frequent swapping of views encourages the interlocutors to communicate with the other their intents using these referents. This allows us to focus on verbal contributions that are explicitly linked to the *situatedness* of the task and its association to a final measure of *task success*: the performance of the teams as well as their learning outcomes as two separate measures. While there are certainly other referring expressions to consider (e.g. “That mountain there") not containing task-specific referents, it would require some degree of manual annotation, or other inference mechanisms. We thus focus on task/domain specific referents that can be automatically extracted.

Therefore, our research question is: **“How do the interlocutors *use* expressions related to the task? Is this associated with task success?"**. This considers lexical alignment, i.e. the *use* of task-specific referents by focusing on *what did the interlocutors say*. We investigate how these local alignment contexts could build to a function of dialogue level task success. We specifically consider the link between expressions related to the task and task success through the *routines*', i.e. the shared expressions', temporality.

Our hypothesis is: **"Task-specific referents become routine early for more successful teams."**. We expect more successful teams to establish routine expressions earlier in the dialogue. Ideally by quicker establishment of routine expressions, teams will understand each other faster and thus have greater task success.

4.2 Materials

In this chapter, from the JUSThink Dataset (described in Section 3.3.2), we utilize:

- *Recorded audio files*: Audio was recorded as two mono audio channels synchronized to each other, with one lavalier microphone per channel. The interlocutors were asked to speak in English. The microphones were clipped onto the interlocutors' shirts. At a local level, to study lexical alignment (and also behavioral alignment in Chapter 5), we transcribe a representative subset of the audio files.
- *Event log files*: Event log entries consist of timestamped edit actions and submitted solutions. At a dialogue level to measure *performance* in the task, we use the teams' best score calculated from all the submitted solutions, using the same measure error defined in Section 3.4.1 (i.e. as in Chapter 3).
- *Pre-test and post-test responses*: At a dialogue level to measure *learning outcomes*, we use interlocutors' scores in the pre-test and the post-test, using the same measure learn defined in Section 3.4.2 (i.e. as in Chapter 3).

4.2.1 Choosing a Representative Subset of the JUSThink Dataset

We sample the data because the state-of-the-art Automatic Speech Recognition (ASR) tools were not sufficiently good on the dialogues (see Section 4.2.3 for a comparison). Therefore, in order to study lexical alignment in this chapter (and also behavioral alignment in Chapter 5), we need to manually transcribe the data: for this purpose, we develop a selection process so that the set is as representative as possible of the whole corpus. We firstly select a *subset* of 10 teams of the dataset. The teams are selected randomly according to the task success distribution (see Figure 4.1); which we measure through performance in the task and learning outcomes measured from the pre-test and the post-test. We thus choose a representation of the percentage of successful teams (30% compared to 21% of the whole dataset).

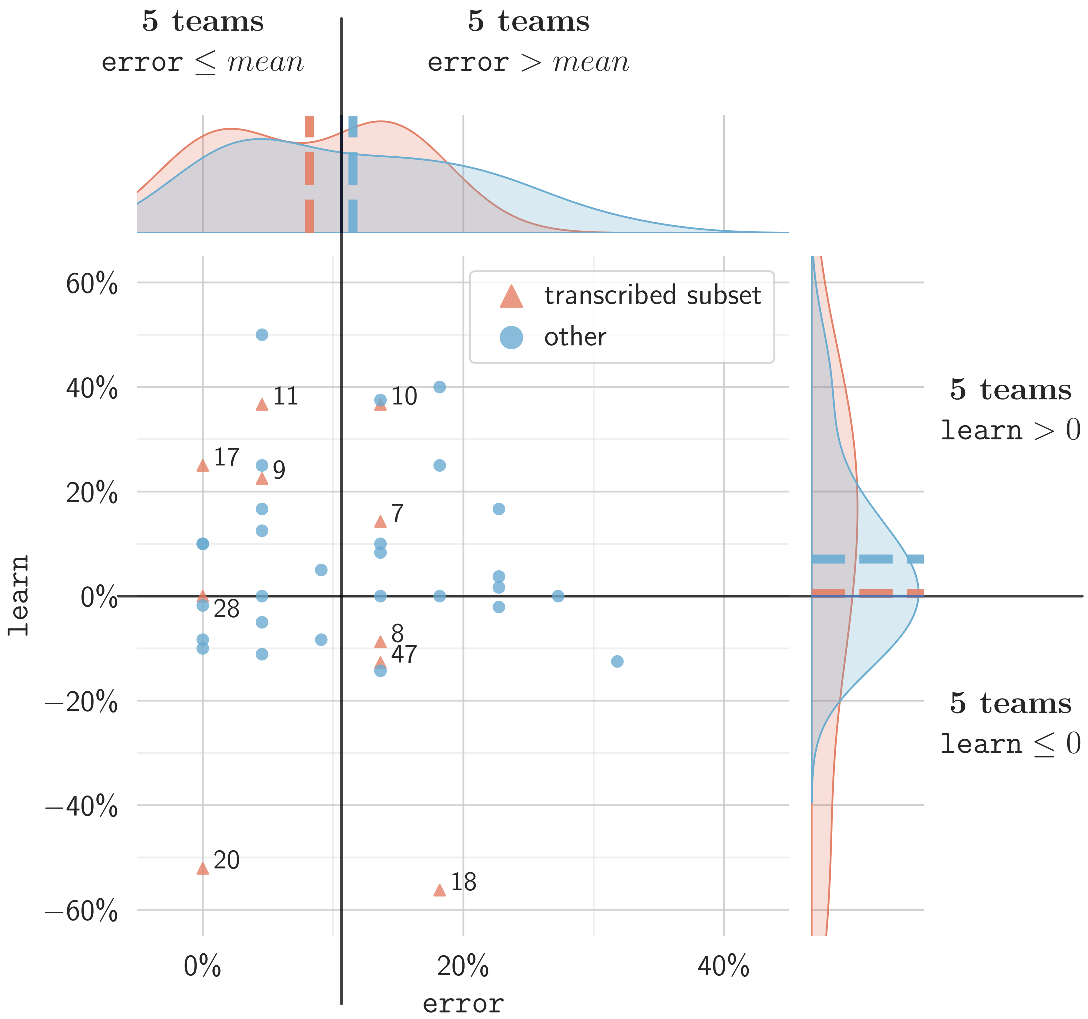


Figure 4.1 – Learning vs. performance plot for sampling of the JUSThink Dialogue Dataset ($N = 10$ teams), showing the transcribed subset and the non-transcribed teams. The black lines indicate the criteria for choosing the samples to transcribe (5 teams with learn > 0 etc.). The mean of a set (transcribed or other) is shown as a dashed line, with the fit of a univariate kernel density estimate for the set. Numbers denote the index of the transcribed teams.

Table 4.1 – Descriptive statistics for the JUSThink Dialogue Dataset ($N = 10$)

Variable	Mean	SD	Min	Max
number of utterances	478.7	157.0	297	773
length of utterances (in tokens)	6.6	5.4	1	32
total duration (min)	23.7	7.1	11.2	36.0
time per submission (min)	2.6	2.3	0.7	12.8
duration of a turn (sec)	25.0	30.3	1.1	240.2
number of submitted solutions	9.4	4.7	4	19
number of turns in task	50.4	21.4	28	98

4.2.2 The JUSThink Dialogue Dataset

The *JUSThink Dialogue Dataset* contains gold-standard/manual transcripts for the selected subset of 10 teams (out of 39) in the JUSThink Dataset (see Section 3.3.2); obtained manually due to the poor performance of state-of-the-art ASR systems on this dataset (which consists of children's speech with music playing in the background). The accuracy of ASR on the dataset is given in Section 4.2.3. The transcripts account for ≈ 4 hours of data, where the mean duration of the task is ≈ 23 min. Table 4.1 provides further details about the transcribed subset.

The transcripts report which interlocutor is speaking (either A or B) and the start and end timestamps for each utterance, beside the utterance content. Utterance segmentation is based on the definition of an Inter-Pausal Unit (IPU) by Koiso et al. (1998), defined as "a stretch of a single interlocutor's speech bounded by pauses longer than 100 ms". We also annotated punctuation markers, such as commas, full stops, exclamation points and question marks. Transcription included incomplete elements, such as "Mount Neuchat-" in "Mount Neuchat-um Mount Interlaken". Pronunciation differed among and within interlocutors (for example, for the word "Montreux", pronouncing the ending as /ks/ or /ø/), due to the unfamiliarity of the interlocutors with the referents, and individual accents. As our methodology is dependent on matching surface forms, we standardize variations of pronunciation in the transcripts, and we do not account e.g. variations in accent. A graduate student (the author of this thesis) completed two passes on each transcript, which were then checked by another native English speaking graduate student with experience in transcription/annotation tasks.

The JUSThink Dialogue Dataset as well as the data used to extract the learning outcomes and performance in the task (transcripts, logs, and responses to the pre-test and the post-test, and a description of the network in the activity) are available from the Zenodo Repository, DOI: 10.5281/zenodo.4627104. In addition, the code that reproduces the results and figures given in this chapter are available from the Zenodo Repository, DOI: 10.5281/zenodo.6974562.

4.2.3 Automatic Speech Recognition (ASR) Comparison

An accurate transcription of the task-specific referents is crucial for both our lexical, and (in Chapter 5) behavioral alignment measures. To evaluate how well automatic speech recognition would perform in obtaining transcripts, we used the Google Cloud Speech-to-Text services.

To configure the speech recognition system, we extracted a 15 seconds long audio sample (which is a part of excerpt in Section 5.3.1) that contains several task-specific referents, as well as a filler "um": see the reference transcript in Table 4.2. We need to assist the system towards improving its accuracy for the task-specific referents in the JUSThink activity, as these are context specific words (e.g. "Montreux") that do not occur frequently and thus would be difficult to recognize. To do so, we supplied the task-specific referents that are available a priori from the definition/context of the activity as "hints" to the ASR system. Then, we used the additional boost feature of the system to increase the probability that these specific phrases will be recognized, as described in the system's documentation[II] (boost = 200, by experimentation). Furthermore, we set the system to use the enhanced "video" model[III] that is particularly suitable for audio "that have multiple speakers", "that was recorded with a high-quality microphone or that has lots of background noise". This is the case for our data that contains background music of the game and some audio spill where one interlocutor's microphone could pick up sound from the other interlocutor.

For the audio sample, Table 4.2 presents our manually obtained reference transcript, the automatic transcript with the default model (without our configuration), and the transcript with a model adapted to our dataset. The transcript with the adapted model seems promising, as the task-specific referents are correctly transcribed. With the same adapted configuration, we automatically transcribed the complete audio files, for the subset of data.

Speech recognition results are typically evaluated by using word error rate (WER), defined as the number of errors in the prediction that has $N_{prediction}$ words (the Levenshtein distance or minimum edit distance for words), normalized by the number of words in the input ground truth N_{input} (Errattahi et al., 2018):[IV]

$$\mathrm{WER} \triangleq \frac{S+I+D}{N_{input}}, \tag{4.1}$$

Equation 4.1 – Word error rate (WER) formula to evaluate speech recognition performance

where S, I and D denote the total number of substitutions, insertions, and deletions. WER is not input/prediction symmetric; and the upper bound of WER is not 100%, but:

$$\max\{N_{input}, N_{prediction}\} / N_{input}.$$

[II]See the ASR's configuration documentation at https://cloud.google.com/python/docs/reference/speech/latest/google.cloud.speech_v1.types.SpeechContext, accessed August 2022

[III]Available models: https://cloud.google.com/speech-to-text/docs/basics#select-model, accessed August 2022

[IV]We use the `jiwer` Python library to compute WER.

Source	Start (s)	End (s)	Int.	Utterance
reference	552.5	556.7	A	what about Mount Davos to Mount , Saint Gallen ?
	558.6	562.6	B	because what if , you did if we could do it ?
	562.5	566.1	A	what about Mount um Davos to Mount Gallen ?
default	552.2	556.1	A	What about Mount Davis to Mount Saint Helen ?
	558.5	562.3	B	Cuz what if you did , if we could do it .
	558.6	560.9	A	What if you did ?
	561.7	565.9	A	Could you it ? What about Mount Davis to mount gallon ?
	566.0	566.7	B	Mount .
adapted	552.6	556.5	A	What about Mount Davos to mount St . Gallen ?
	558.7	562.8	B	Cuz what if you did , if we could do it .
	562.5	565.8	A	What about Mount Davos to mount Gallen ?

Table 4.2 – Automatic transcription results for an audio sample from JUSThink Dataset

For instance, a recognition system that outputs a different word for each input word has $S = N_{input}, I = D = 0$ and therefore, $WER = N_{input}/N_{input} = 100\%$. If the system predicts two incorrect words for each input word, then $S = I = N_{input}, D = 0$ and thus $WER = 200\%$ (Morris et al., 2004). The lower bound 0% is achieved when the prediction is the same as the input.

In addition to WER, we use the error rate for task-specific referents only as a list of domain-specific keywords (since only task specific referents are required for the methodology of this and the next chapter in Section 4.3 and Section 5.3). WER and keyword error rate are tend to be very similar for high accuracy transcripts; with the keyword rate increasing more rapidly than WER as the recognition system's performance decreases (Park et al., 2008).

The error rates for the automatic transcripts are given in Figure 4.2. The comparison shows that the word error rates are very high ($Mdn = 62.6\%$), and varied between interlocutors (ranging from 37.7% to 253.4%).[V] When we filter for the task-specific referents, we see that the referent error rates are high as well ($Mdn = 47.3\%$). Thus, it is infeasible to use automatic transcripts for this work, and therefore use gold-standard transcripts. We compute WER over the complete transcripts; averaging over utterances would require a robust segmentation length of the utterances vary in length and when averaged over utterances.

4.3 Methods

In the JUSThink activity, the children are initially prompted by the robot to work with each other, and later simply given the cost difference for their sub-optimal solution(s). However, almost all of the exchanges are between the two participants. After careful observation of the dialogues in the dataset, We observe the tendency to ignore the robot unless submitting a solution. Thus, we treat this triadic activity as a dyadic dialogue.

[V] The WERs have been reported to be as low as 2 – 3% for some of the benchmarks; however, they were shown to be much higher than the reported results especially in real-life datasets, up to 25% (Szymański et al., 2020). A low WER does not necessarily mean better language understanding accuracy (Y.-Y. Wang et al., 2003).

4.3.1 Studying How Referents Contribute to Lexical Alignment

A routine is formed when a referring expression is commonly used by both interlocutors. For example, “Montreux" is a task-specific referent, and one interlocutor might *prime* the referring expression “Mount Montreux". If the other interlocutor *reuses* this referring expression, it becomes a *routine*, i.e. a common part of the dialogue.

For our methodology, a routine is specific to the exact matching of token sequences in two utterance strings. We thus formally define a *routine* expression (adapted from Dubuisson Duplessis et al., 2017, 2021, Pickering and Garrod, 2004) as a referring expression shared by two interlocutors if:

1. the referring expression is produced by both interlocutors, and
2. it is produced at least once without being part of a larger routine.

In particular, we define the utterance at which a referring expression becomes routine as the *establishment* of that routine. We extract the utterances at which the routines are primed and established from the transcripts as in Dubuisson Duplessis et al. (2017, 2021). Then, we filter for the routines that contain a task-specific referent.

To investigate when the routine expressions become established, we study (i) the *establishment time* of a routine, i.e. the end time of the utterance at which the expression is established, and (ii) the *collaborative period* of a team, i.e. the duration between first quartile (Q1) to third quartile (Q3), i.e. the interquartile range (IQR) of the establishment, where half of the establishments occur.

To then study our hypothesis, for task performance, we check if the median establishment times are significantly earlier for better-performing teams by Spearman's rank correlation and its statistical significance (between the median and the error). For the learning outcomes, we compare the distribution of the median establishment times of teams with positive learning outcomes (Teams 7, 9, 10, 11, 17) and others (Teams 8, 18, 20, 28, 47) by Kruskal-Wallis H test.[VI]

To study the process of lexical alignment, we consider:

1. the establishment times of all the routines in real time (as teams took different durations to complete the task),
2. the establishment times of the routines that are established in the common time duration (i.e. the first ≈ 11 min for all the teams, which was the time taken by the quickest team), and

[VI] Since we have groups of 5 teams for learning, we use Kruskal-Wallis (SciPy's implementation of Kruskal-Wallis works with ≥ 5 samples). It can not be used for performance that have 3 teams in the high-performing group.

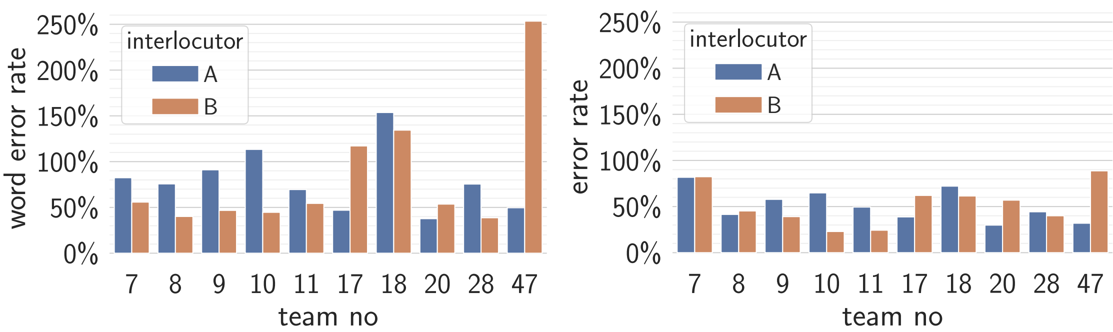
3. normalized establishment times, that are scaled by the duration of each team itself to reflect the *progress* of a team's interaction, from 0% progress at the beginning of the activity, to 100% when the interaction ends (either by finding an optimal solution, or by being intervened by the experimenters to end the task).

To gain further insight, we compare the distribution of the establishment times of the teams, through inspecting the box plots of all teams that are compared side-by-side and sorted by decreasing task success.

4.3.2 Accuracy of the Routine Extraction Algorithm

The algorithm extracts routine expressions by the *exact matching of token sequences*: thus, the accuracy of the inference depends only on having accurate transcripts. We have gold-standard transcriptions with standardized variations of pronunciation (the details of the transcription are given Section 4.2.2). Thus the extraction of routine expressions, and subsequently determining the priming and establishment times are not sources of error.

However, this exact matching of token sequences is exhaustive. The original work by Dubuisson Duplessis et al. (2021) is intended to measure alignment by enumerating all existing matches (for example, even if interlocutors primed and established the token "what", this would be counted towards a routine expression formed). This is why we filter for referents that are specific to the task. Therefore, we would like to highlight that the alignment is *lexical*, based on a surface level matching of token sequences. Furthermore, we keep in mind the issues of transcribing disfluent speech (Le Grézause, 2017; Zayats et al., 2019), and thus use transcription software (Praat) to ensure e.g. there are no unnecessary insertion, substitution or deletion of disfluencies.



(a) Word error rates

(b) Task-specific referent error rates

Figure 4.2 – Error rates for the transcripts obtained by automatic speech recognition

4.4 Results

Figure 4.3 shows for each team the distribution of the establishment times of the routines, sorted by task performance. The median establishment times for all of the routines is strongly positively correlated with the task performance measure error (Spearman's $\rho = 0.69$, $p = .026$), that is, **better-performing teams establish routines earlier**.[VII] We see from the figure that the results are influenced by the variation in the duration of the activity: the interaction ends for the well-performing teams when they find a correct solution, whereas the badly-performing teams continue their interaction until the experimenters intervene and stop the activity.

Figure 4.4 shows the normalized establishment times. We see that the **establishment occurs around the middle of the dialogue**: establishment times have mean of the medians = 63.0% (combined SD = 22.2%). While we hypothesized that establishment will happen early in the dialogue, this is the case for an ideal dialogue; people will "share" expressions earlier. However, we observe that there is an *exploratory* period (the period before Q1), where the interlocutors take the time to understand the task, followed by a *collaborative* period that corresponds to the establishment period (the period between Q1 and Q3). We expect that if the interlocutors had to complete the task again, the establishment/collaborative period would be closer to the start of the dialogue. While it is expected that better-performing teams have established routines, from the figures we observe even badly-performing teams still successfully established routines. Thus it is to be noted that **all teams, regardless of performance, have aligned to some degree**.

We observe that five teams (7, 8, 10, 18, and 47) that had their error > *mean* (see Figure 4.1) started collaborating later in their dialogue, in terms of when they establish most of their routines (median establishment time > 60%). Though the measures of task performance would only reflect that these teams performed badly with a final overall score of performance, our alignment measures reflect details of the performance. Interlocutors are gently reminded a few minutes before the end of the task the remaining time, but were not rushed to find a solution: which could have changed the way they aligned by forcing an establishment period. Thus, we observe through our alignment measures that **badly performing teams were simply slower to collaborate and establish routines**.

The distribution of the median establishment times are not statistically significantly different for teams with positive learning outcomes (Teams 7, 9, 10, 11, 17) vs. others (Teams 8, 18, 20, 28, 47) by Kruskal-Wallis H test ($H = 0.88$, $p = .35$ for times in absolute values, and $H = 0.10$, $p = .75$ for normalized times).

[VII] The magnitude of Spearman's correlation coefficient (ρ) can be interpreted by using the thresholds from Evans (1996), i.e. 0.00 – 0.19 "very weak", 0.20 – 0.39 "weak", 0.40 – 0.59 "moderate", 0.60 – 0.79 "strong", and 0.80 – 1.00 "very strong" .

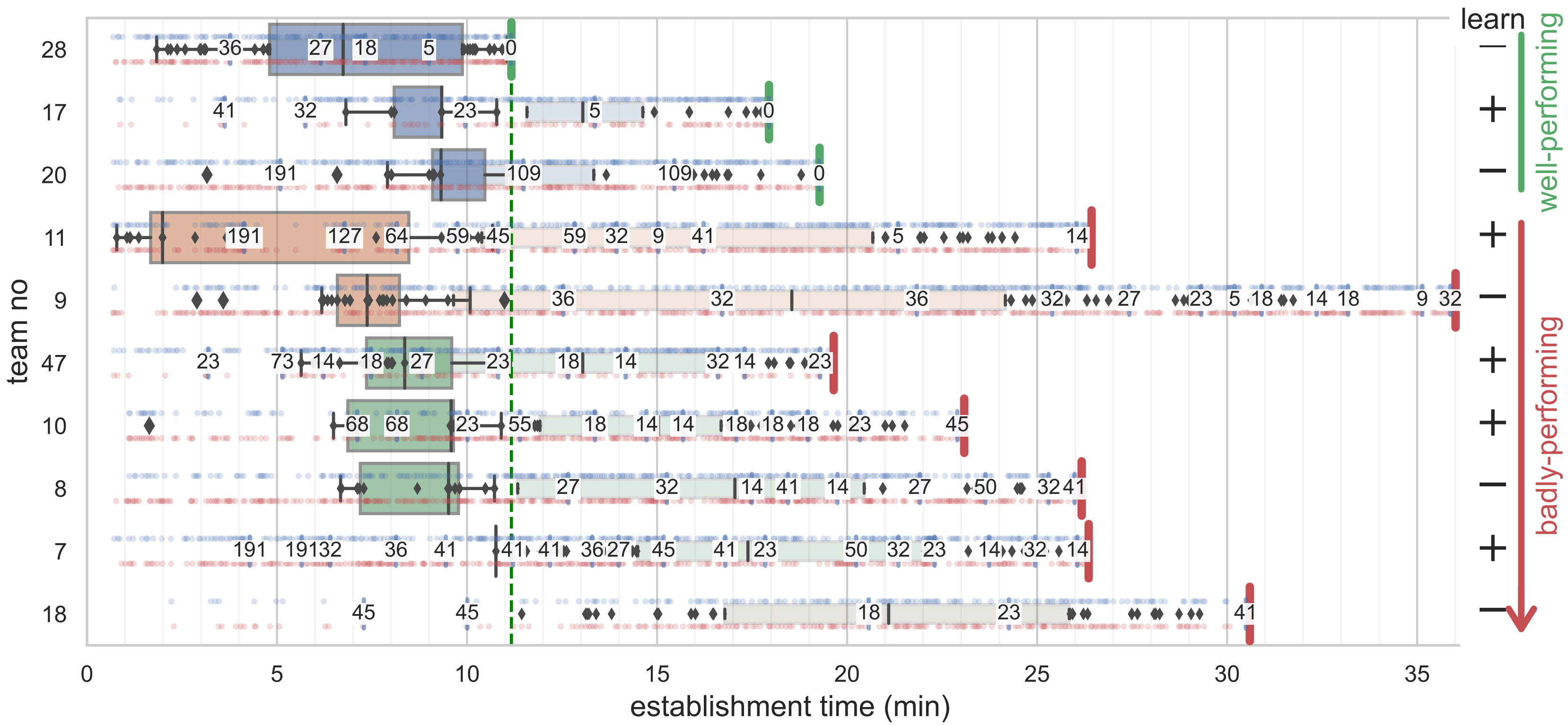


Figure 4.3 – Establishment times in the JUSThink Dialogue Dataset. The teams are sorted by decreasing task performance (i.e. increasing error). The boxplots, with the same color have the same error (sorted by increasing duration for the ties). The thick, bold boxplots with whiskers with maximum 1.5 IQR show the distributions of the establishments that occurred in the common duration (i.e. as marked by the dashed green line), while the thin boxplots show the distributions through the total duration of interaction. The learning outcome of each team is indicated with a plus (‘+’) for learn > 0, or a minus (‘−’) otherwise. Solid lines indicate the end of the interaction, by submitting a correct solution (in green) or timing out (in red). The thin blue lines indicate submission of a solution, with the number showing the error quantifying how far the submitted solution is from an optimal solution in terms of its cost (e.g. error = 0% means the team has found an optimal solution). The red and blue dots indicate the utterance times of the interlocutors, to give an idea of when the interlocutors are speaking.

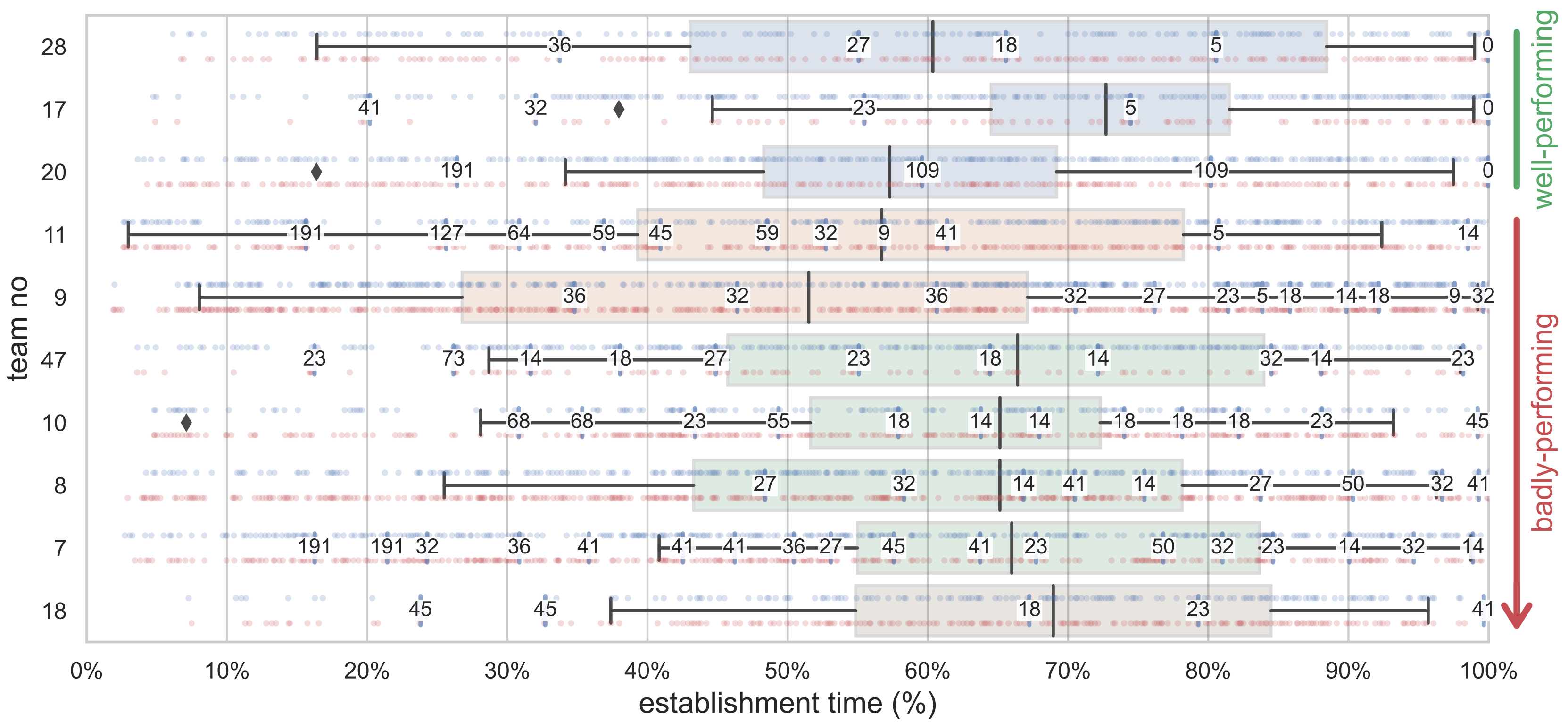


Figure 4.4 – Establishment times normalized by the duration for each team separately, i.e. 100% indicates the end of that team's interaction. Sorting and annotations are as in Figure 4.3.

4.5 Discussion

We would like to discuss the notions of *collaborative period* and *exploratory period*, by which we refer to periods that are automatically inferred by our lexical alignment measure and the distribution of its detections over the dialogue: we highlight that this measure processes only the dialogue content, without consulting to the actions taken by the participants. Interestingly, these periods correspond to meaningful time intervals as also suggested by the names we refer to them, when checked together with the teams' submissions. Next, we discuss and expand on these observations.

To interpret and understand what is happening in these intervals, we examine the *error progression* of teams' submitted solutions during the task (Figure 4.4). During the exploratory period (which is inferred from dialogue only), nine out of ten teams had their highest error. For teams that performed well, their collaborative period had their lowest error before finding an optimal solution (thus ending the task), and for other teams; their closest solution to the optimal solution. There are 8 such teams that exhibit this behavior. The teams that did not achieve a correct solution, never regressed to their largest error from the exploratory period.

Looking at this through the *number of attempts*, for the well-performing teams, their collaborative period was productively used, with their next solution reaching an optimal cost (Teams 17 and 28), or one more attempt before their optimal cost (Team 20). Several teams that did not perform well have a greater number of attempts submitted after their collaborative period (Teams 9, 7, 8, 10, 47). Their submission pattern of attempts indicate a "trial-and-error" strategy on how to solve the task, for example, Teams 8 and 9 increasing their submissions after their collaborative period, or Teams 7, 10 and 47 submitting throughout.

Lastly, we compare the team that had the most task success (Team 17, with positive performance and learning gains) with the team that had the lowest task success (Team 18). The two teams have established routines later in their dialogue: Team 17 that found a correct solution, and Team 18 that could not (their median establishment times are around 70%). Yet, Team 17 had a focused establishment period (with a smaller IQR = 18% vs. 30% of the time, respectively, see Figure 4.4). We interpret that Team 17 was able to turn it around and find a correct solution, while Team 18 did not; ending up as the worst performing team. 17 has positive learning gain ($\mathsf{learn} = 25\%$), while it is negative for Team 18 ($\mathsf{learn} = -56\%$, the highest decrease among the transcribed teams, see Figure 4.1). While bad performance could be reflected through a collaborative period starting later in their dialogue (such as Team 18; with later establishment, and not finding a solution in time), Team 17 shows that there are exceptions to this. Team 18 also possibly got confused, reflected in the high and negative learn.

4.6 Conclusion

Collaborative learning activities are a particularly interesting type of collaborative task, due to their “multi-layered goals"; typically including immediate, performance goals (e.g. finding the solution to a math problem) and deeper, learning goals (e.g. understanding the notion of equation). Collaboration often involves a dialogue among interlocutors, where the dialogue contains structures on how humans build a mutual understanding of each other about the activity: these patterns can be “revealed" and quantified by studying the dialogue. It has been shown that a dialogue is successful when there is alignment between the interlocutors, at different linguistic levels. Thus, in this chapter, we focused on *lexical alignment*, i.e. alignment at the lexical level, and proposed novel rule-based algorithms to automatically and empirically measure this this process. We prepared and studied a corpus of data (dialogue transcripts from audio files) generated by teams of two children engaged in the collaborative learning activity from the previous chapter (Chapter 3), which aims at providing an intuitive understanding of graphs and spanning trees.

A finding of this chapter is the discovery that the measures we propose are capable of capturing elements of lexical alignment in such a context. We see that all teams establish routines, regardless of task success. An assumption that might be commonly made, is that the more aligned interlocutors are, the more are chances of their task success (usually, measured by performance alone). Indeed, our results from this chapter indicate that this is not necessarily the case for lexical alignment: we rather observed that better performing teams were earlier than badly performing teams to align by our measures. This pattern is inconclusive for teams with positive learning outcomes. Thus, inspecting the distributions of the establishment times together with their proposed solutions gave further insight into the collaboration processes, specifically revealing a *collaborative period*, that turns out to be the period where most of the teams came up with their best solutions. This period is inferred only from the dialogue content, on task-related expressions that the robot can specifically target to detect. Interestingly, it corresponds to a very productive time interval for the better performing teams, in terms of achieving their lowest cost so far; this we discovered only after checking the submissions of the teams in these intervals. We believe that this captures some local level alignment patterns that might have otherwise been overlooked when only considering success in the task.

5 Linking What We Say and Do to Model Mutual Understanding

5.1 Introduction

In the previous chapter (Chapter 4), we studied the dialogue between children, resulting from solving a problem together in the JUSThink activity: we examined the formation of the expressions that refer to the activity in what was said, and automatically quantified children's collaboration and mutual understanding in terms of *lexical alignment*. In this chapter, we study their dialogue together with the actions that they take in the task: we present algorithms to automatically measure the *behavioral alignment* (what was done in relation to what was said), i.e. a new alignment context we propose to mean when instructions provided by one interlocutor are either followed or not followed with physical actions by the other interlocutor.[I] These algorithms serve as automatic methods a robot can use to detect the alignment: the method for inferring the instructions is based on the task-related expressions that the robot can specifically target to detect (as in robot's use of lexical alignment in Chapter 4), and then, the robot can execute the alignment method with its observations of the next actions as inputs.

Results highlight that all teams were behaviorally aligned, irrespective of their task success. In terms of task performance, we observe a general trend that better performing teams tended to follow up the instructions with actions (as matches) early overall (as absolute times), as well as in the progress of their interaction (normalized times). For learning, although the general trend is not statistically confirmed, we gained insight into the nuances of the dynamics of interaction, by inspecting the normalized (mis)match time plots alongside with the submission costs: there seems to be *conflicts* that may have been collaboratively resolved and resulted in learning, or remained unresolved and had adverse effects on the learning outcomes.

This work corresponds to the following publication:

Norman*, **U.**, Dinkar*, T., Bruno, B., & Clavel, C. (2022). Studying alignment in a collab-

[I] We make this dataset and tools to study alignment in children's dialogues publicly available (containing transcripts, action logs and alignment tools), from the Zenodo Repositories DOI: 10.5281/zenodo.4627104 for the dataset, and DOI: 10.5281/zenodo.6974562 for the tools.

orative learning activity via automatic methods: the link between what we say and do. *Dialogue & Discourse*, *13*(2), 1–48. https://doi.org/10.5210/dad.2022.201. **Contributed equally to this work.* (Norman, Dinkar, et al., 2022)

Within this work, the author of this thesis implemented the tools, and together with collaborator Tanvi Dinkar, contributed to the data analysis and interpretation of the results.

5.1.1 Research Question and Hypothesis

Our research question in this chapter is: **"How do the interlocutors *follow up* the expressions related to the task with actions? Is this associated with task success?"**. We investigate how the use of these expressions manifests in the interlocutors' actions within the task, and whether this is associated with their task success. For this purpose, we consider the *instruction*s of an interlocutor, as the verbalized instructions one interlocutor gives to the other, which we extract through their use of task-specific referents. A physical manifestation of this instruction could result in a *corresponding* edit action, or a *different* edit action. We investigate the follow-up actions of the task-specific referents which its effect on task success, through the follow-up actions' temporality.

We hypothesize: **"Instructions are more likely to be followed by a corresponding action early in the dialogue for more successful teams."**. We expect that the earlier interlocutors align with each other in terms of instructions and follow-up actions, the better they progress in the task, and the greater the chance of success in the task. This idea of verbalized instructions being followed up by corresponding actions is in line with previous research on alignment (i.e. interlocutors being in alignment in a successful dialogue). However, work on collaborative learning suggests that individual cognitive development (in our case, positive learning outcomes) happens via socio-cognitive conflict (Doise & Mugny, 1981/1984), and its regulation (Butera et al., 2019). In our task, this means a verbalized instruction could be followed by a corresponding or a different action; as a different action could result in collaboratively resolving conflicts and together building a solution—resulting in task success.

5.2 Materials

5.2.1 Combining JUSThink Dialogue Dataset with the Actions

To analyze the teams' dialogue together with their actions as a process, we combine the transcripts from JUSThink Dialogue Dataset (Section 4.2.2) with the actions from the (original) JUSThink Dataset (Section 3.3.2) to a series of timestamped *events* in a subject-verb-object(-turn-attempt) format, to obtain a verbal and physical actions list A for each team:

- Each utterance in the transcript is added as an action with the verb 'says'.
- Each edit action from the logs is added with the verb 'adds' or 'removes', according to

whether it is an add action or a remove action, respectively.

Each event is represented like an action $a \in A$ with fields:

- a.subject $\in$ {'A', 'B'}, the two participants that are collaborating to solve the problem.
- a.verb $\in$ {'says', 'adds', 'removes', ...}, the edit actions and utterance action for matching instructions with edit actions. There are also other actions like button presses; but they are not used in our analysis.
- a.object $\in$ {Utterances} $\cup$ $\{(u, v) : (u, v) \in$ Edges$\}$. The available edges that can be added or removed can be seen in Figure 3.3a.
- a.turn $\in \{1, 2, \ldots, n\}$ indicating the turn number of the period the action belongs to (where for utterances, the start time of the utterance belongs to). After every two edits, the turn number incremented by one.
- a.attempt $\in \{1, 2, \ldots, m\}$ indicating the attempt number of the period the action belongs to (where for utterances, the start time of the utterance belongs to). After every submission, the attempt number is incremented by one.

5.2.2 The JUSThink Dialogue and Actions Dataset

The *JUSThink Dialogue and Actions Dataset* contains the dialogue (utterances of the participants) and actions (physical actions in the activity, e.g. to add a connection, submit a solution) as a series of timestamped events for $N = 10$ teams of the JUSThink Dataset (see Section 3.3.2). Table 5.1 presents descriptive statistics for this derived dataset: the table expands on the entries of Table 4.1 (that are replicated here for convenience) with action information.

The JUSThink Dialogue and Actions Dataset is made available with the code that derives it from the Zenodo Repository, DOI: 10.5281/zenodo.6974562. All relevant data to prepare this dataset as well as extract the learning outcomes (transcripts, logs, and responses to the pre-test and the post-test, and a description of the network in the activity) are available from the Zenodo Repository, DOI: 10.5281/zenodo.4627104.

5.3 Methods

5.3.1 Studying Behavioral Alignment

We follow an idealized perspective on the activity: at a given time, the participant in the abstract view is an Instruction Giver (IG) who describes their instructions for the task by using specific referring expressions, and the other (in visual view) is the Instruction Follower (IF)

Table 5.1 – Descriptive statistics for the JUSThink Dialogue and Actions Dataset ($N = 10$)

Variable	Mean	SD	Min	Max
number of events (utterances and physical actions)	748.8	235.3	451	1076
number of utterances	478.7	157.0	297	773
number of edit actions (add and remove edge)	101.9	46.1	56	206
length of utterances (in tokens)	6.6	5.4	1	32
total duration (min)	23.7	7.1	11.2	36.0
time per submission (min)	2.6	2.3	0.7	12.8
duration of a turn (sec)	25.0	30.3	1.1	240.2
number of submitted solutions	9.4	4.7	4	19
number of turns	50.4	21.4	28	98

who executes the action (this is akin to the Map Task (A. H. Anderson et al., 1991)).[II] The activity design creates a frequent swapping of views. This aims to discourage participants from working in isolation or in fixed roles of IG and IF, which could potentially happen in collaborative tasks.

If an instruction is verbalized (e.g. to connect "Mount Basel to Montreux") by an interlocutor (IG), it could result in an action of connecting the two. We hence say an instruction *match*es an action when the instruction is executed by the other interlocutor (IF) via an action in the situated environment, and within the period of a turn of views before they are swapped. We study the discrepancy created when the IF does not follow the IG, which we call a *mismatch* of instructions-to-actions. In the following dialogue excerpt, the instruction (to connect Gallen to Davos) matches the action (connecting these two):

> ***A*:** Now should I do from Mount Neuchatel to Mount Basel?
> ***B*:** No. Do mount . . .
> ***A*:** Look . . .
> ***B*:** To Interlaken maybe, because we did Interlaken at the last, so maybe if we do it before.
> ***A*:** Oh yeah. This one here?
> ***B*:** Yeah mount, turn to Mount Interlaken. *[Instruction to add an edge to Mount Interlaken.]*
> ***A*:** Wait so I'm doing Mount Montreux to Interlaken?
> ***B*:** Yeah.
> ***A*:** Okay. It's your turn.
> ⟨*A* connects Mount Montreux to Mount Interlaken ⟩ *[Match!]*
> ***B*:** Because it's, it's not it's at 4 francs, oh Jesus! It's expensive here!
> ***A*:** Oh i never knew that this was what it's, oh now i get it. Yeah.

[II] In fact, the IF could also be following their own intuitions and ignoring the IG. Yet, ultimately we think this is a justified assumption, as only the IF is in control of the actions.

Table 5.2 – Example output for recognizing instructions and detecting follow-up actions (by RECOGNIZE-INSTRUCTIONS (Algorithm A.2) and MATCH-INSTRUCTIONS-TO-ACTIONS (Algorithm A.4)), for Team 10. *View* denotes which view the interlocutor, either *abstract* (Ab) or *visual* (V). *Annotations* denotes the automatically inferred instructions and follow-up actions in the activity. For example, *Instruct$_A$* indicates that interlocutor A has given an instruction to add two nodes (inferred from referents), which can be partially recognized *(Gallen,?)*. As shown, the algorithm builds up (or "caches") instructions until an edit action is performed ('–' in Utt.). Since B is in the visual view, their inferred instruction is deliberately not matched.

	Utt.	View	Verb	Utterance	Annotations
A	198	Ab	says	Maybe we start from, Mount Zermatt ?	Instruct_A(Add(Zermatt,?))
B	199	V	says	No lets do Mount Davos to, where do you wanna go?	Instruct_A(Add(Zermatt,?), Instruct_B(Add(Davos,?))
A	200	Ab	says	... to Mount, St Gallen.	Instruct_A(Add(Zermatt,?), Instruct_B(Add(Davos,?) , Instruct_A(Add(Gallen,?))
B	201	V	says	Okay.	*As previous*
B	–	V	adds	Gallen-Davos	Instruct_A(Add(Zermatt,?), Match_B (Instruct_A(Add(Gallen,?)))

The following dialogue excerpt illustrates a mismatch:

***B*:** Go to Mount Basel. *[Instruction to add an edge to Basel.]*
***A*:** That's, it's expensive.
***B*:** Just do it.
***A*:** You can't, you can't, I can't because there's a mountain there ...
***A*:** So I'm going, so I'm going here.
⟨*A* connects Mount Interlaken to Mount Bern⟩ *[Mismatch!]*

5.3.2 Recognizing Instructions

We firstly extract instructions from the utterances through the interlocutors' use of task-specific referents. In the schema as shown in Figure 5.1, our input consists timestamped dialogue transcripts and action logs. When the input is an utterance, it is processed to infer instructions using entity patterns. To check these entity patters, we employ Named Entity Recognition (NER) feature of the `spaCy` Python library that performs this entity recognition. We add the node names of the mountains (e.g. "Montreux"), and also verbs; i.e. "add", "remove" ... Then, if the input utterance contains our custom entity patters, then we automatically infer instructions from the utterance by joining these entities together. For example, the result may be Add(Node_1,Node_2), because the interlocutor explicitly said the verb "add" and also the names of the mountains.

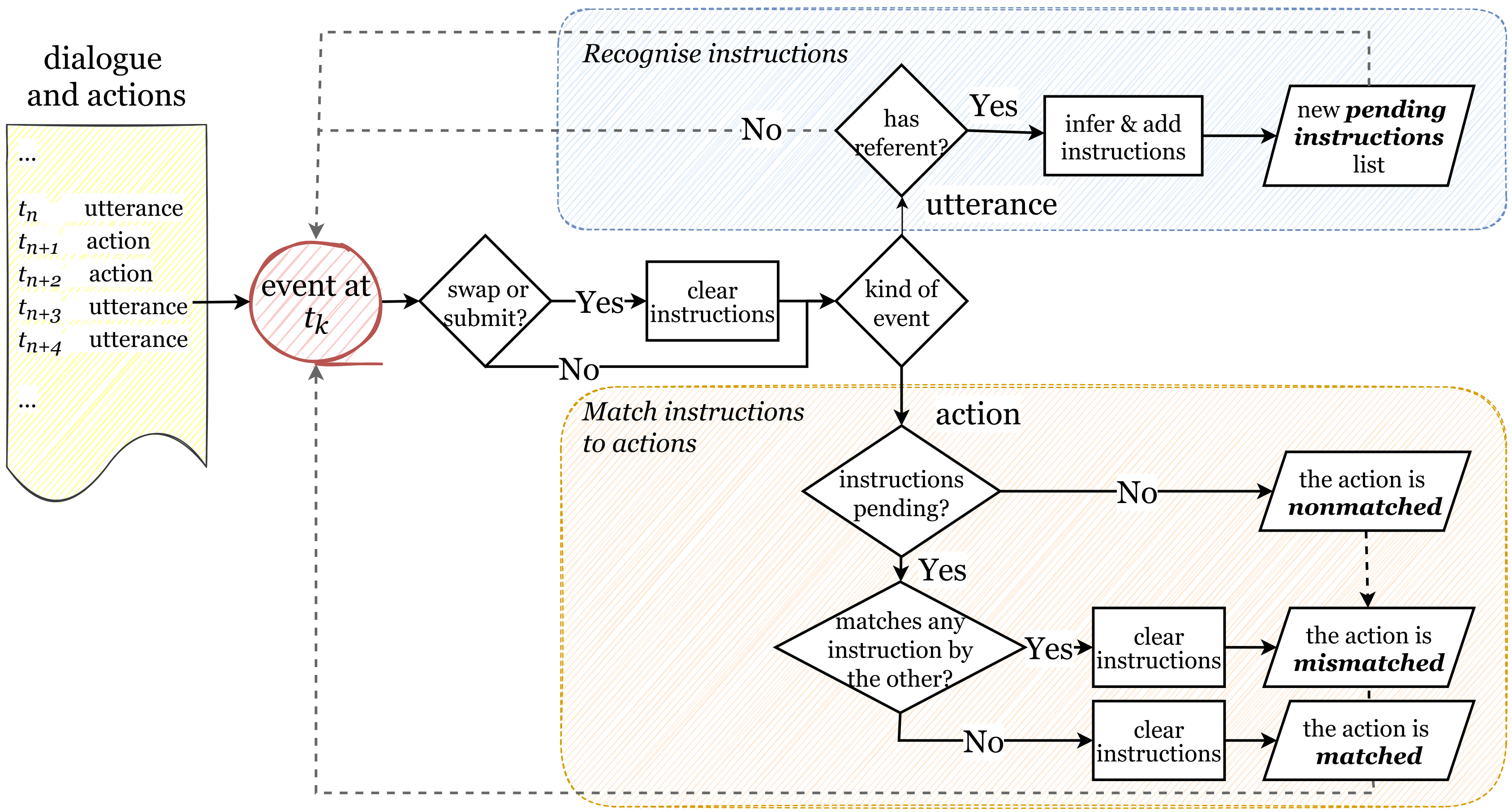


Figure 5.1 – Flowchart for automatic annotation of behavioral alignment (e.g. the parallelogram shows the annotated output)

5.3.3 Matching Instructions-to-Actions

To find (mis)matches of instructions-to-actions, we then follow the logic as given in the schema to determine whether there is an inferred instruction in the instructions list at the time an action was taken. Then, we check whether the action matches or mismatches the inferred instruction.

In Section 5.3.1, Gallen-Davos was the result of a negotiation, rather than a complete given instruction by the IG. B says in one utterance "oh Mount Davos" and then in another "to Mount Gallen yeah do that", resulting in two cached inferred instructions; (Davos,?) and (Gallen,?) respectively. This accounts for some amount of multiple speaker turn[III] negotiations, and possible other ways of referring to task-specific referents (e.g. "Now go from *there* to Davos"). Though an instruction could be carried out after the views swap again, i.e. in the following turn, in our methodology, the pending instructions are cleared at every swap (if "swap or submit", then "clear instructions"), resulting sometimes in a *nonmatched* action, or when an action occurred but there was no inferred instruction. A full and concrete example of the added annotations of our automatically inferred instructions-to-actions is given in Table 5.2.

To investigate our hypothesis, we compare the distributions of the match and mismatch times of the teams, by following the same methodology stated in Chapter 4 (which investigated establishment and priming times). In particular, we check if the median match or mismatch times are significantly earlier for better-performing teams by Spearman's rank correlation, and for the two groups of learning, by Kruskal-Wallis H test.

5.3.4 Accuracy of the Algorithms

There are two possible sources of error: while (i) inferring instructions, and then while (ii) matching them with actions. To infer instructions, we allow for a build up of instructions over a period of time (caching instructions), and then clear these instructions. This ensures that an instruction given at the start of the interaction is not matched with an action at the end of the interaction, i.e. there is a *temporal constraint* of when an instruction is a valid instruction.

The main source of error in inferring instructions could be in *anaphora resolution*, e.g. if an interlocutor says "Connect that node *there*". This is why we allow partially inferred instructions (i.e. only one of the node names is mentioned) and (partial) matching of these instructions with actions. For example, from the utterance "Maybe we start from Mount Zermatt.", with only one node being explicitly stated, we infer the partial instruction *Add(Zermatt,?)*. Then, we consider the follow-up action as matched, if it is an add action, with 'Zermatt' as one of the nodes of the added connection. Thus, the inferring of instructions could be considered *greedy*, and suffer from inferences at each iteration without considering broader context.

[III] We explicitly use *speaker turn* to distinguish from a turn in the collaborative activity; i.e. every swapping of views.

With regard to matching instructions to actions, problems arise when considering the *formulaic definition of behavioral alignment*. The algorithm is explicit in considering which interlocutor is in which view (i.e. always characterizing interlocutors into IF and IG). Matching instructions-to-actions are defined in a computationally strict way, and may not consider matched instructions resulting from negotiations, or consequently the build up of matches over larger periods of time (as the instructions are cleared frequently). Furthermore, given this formulaic definition, this may be a *difficult task to manually annotate* by human annotator, i.e. constantly considering who is in the position to give instructions versus follow (which changes throughout the interaction).

The annotation of high-level constructs, such as engagement, emotion and in our work, behavioral alignment, all have a perceptual component that is intrinsically hard to define, and thus guide for annotation purposes (e.g. see Nasir et al., 2022b). Therefore, given the way behavioral alignment is defined, our measures *may not capture the whole picture*, of what an annotator might *perceptually* label the interaction to be.

5.4 Results

The median match times has a moderate positive correlation coefficient with the task performance measure error (Spearman's $\rho = 0.59$, $p = .08$). This means that better-performing teams tend to have matches earlier. While close to significant (however, $p > .05$), this follows our results regarding lexical alignment; i.e. **better performing teams align both verbally and behaviorally earlier** (in terms of matched instructions-to-actions) than badly performing teams. **Our hypothesis is weakly supported** by these results. Better-performing teams tend to have mismatches earlier as well (Spearman's $\rho = 0.70$, $p = .024$).

Figure 5.2 shows the the distribution of the (mis)match times for each team in a common time frame. When the times are normalized by the duration for each team separately, we still observe a positive, moderate correlation coefficient (Spearman's $\rho = 0.54$, $p = .11$), that indicates a general trend that matches happen earlier for the better-performing teams.

Figure 5.3 shows for each team, the distribution of the match and mismatch times, normalized by the duration of each team itself. While it is natural to think that instructions will be followed by a match for more successful teams in a structured and organized manner, we observe that teams that performed badly and teams that did not learn also have a certain period of matches. Therefore independent of their task success, **all teams have their IG's instructions matched to the IF's actions** to some extent. Match times have mean of the medians = 62.6% (combined $SD = 21.9\%$), and mismatches have mean of medians = 67.6% (combined $SD = 26.0\%$).

By inspecting the median of the match times with learning outcomes, we observe that they are not significantly different for teams with positive learning outcomes (Teams 7, 9, 10, 11, 17) vs. others (Teams 8, 18, 20, 28, 47) by Kruskal-Wallis H tests ($H = 0.54$, $p = .47$ for times in absolute values, and $H = 0.10$, $p = .75$ for normalized times).

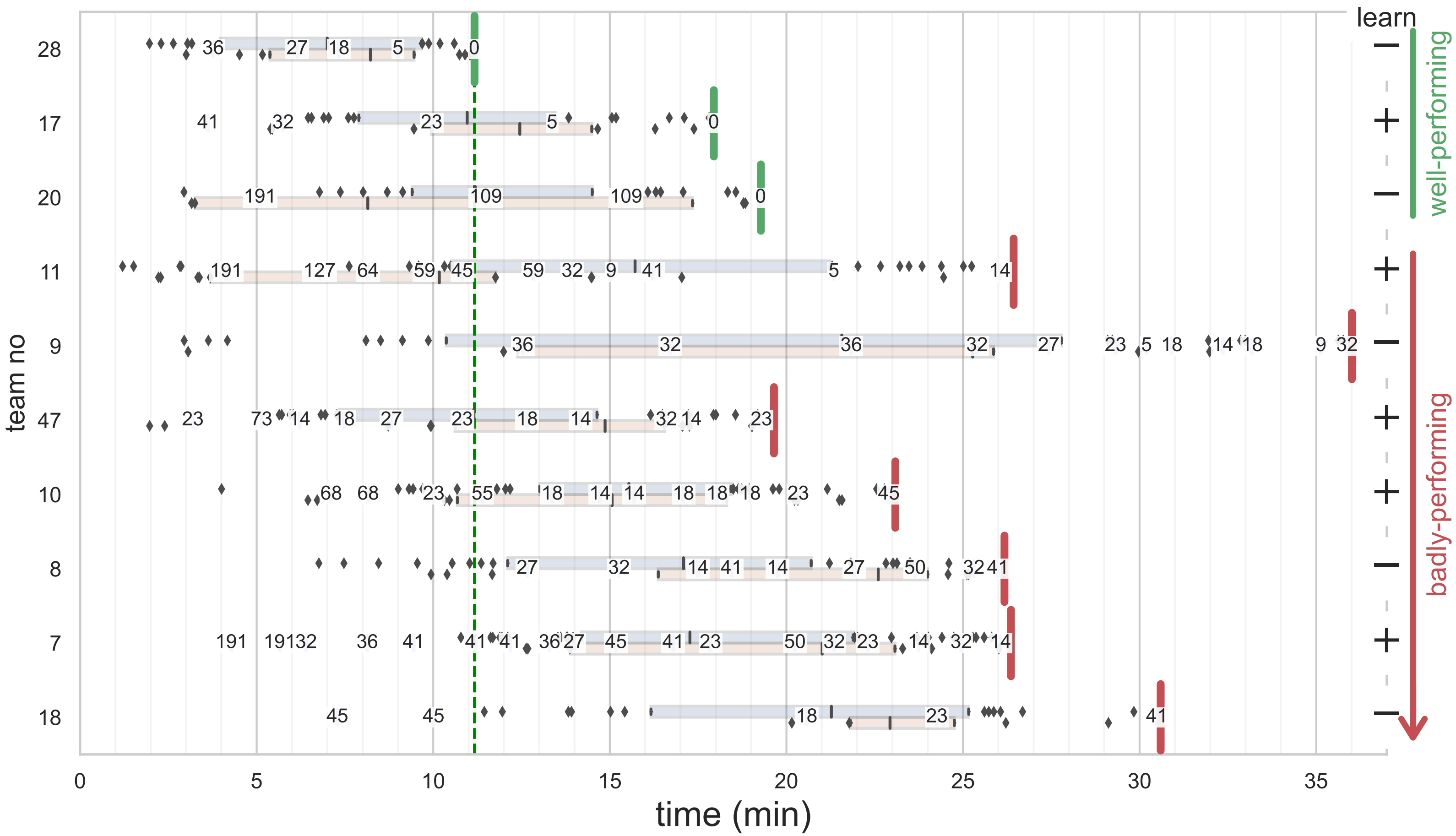


Figure 5.2 – Match and mismatch times in the JUSThink Dialogue and Actions Dataset, in blue and red respectively. The teams are sorted by decreasing task performance. The learning outcomes are indicated with a plus ('+') for learn > 0, or else a minus ('−'). Solid lines indicate the end of the interaction, by submitting a correct solution (in green) or timing out (in red). The thin blue lines indicate submission of a solution, with the number showing the error quantifying how far the submitted solution is from an optimal solution in terms of its cost.

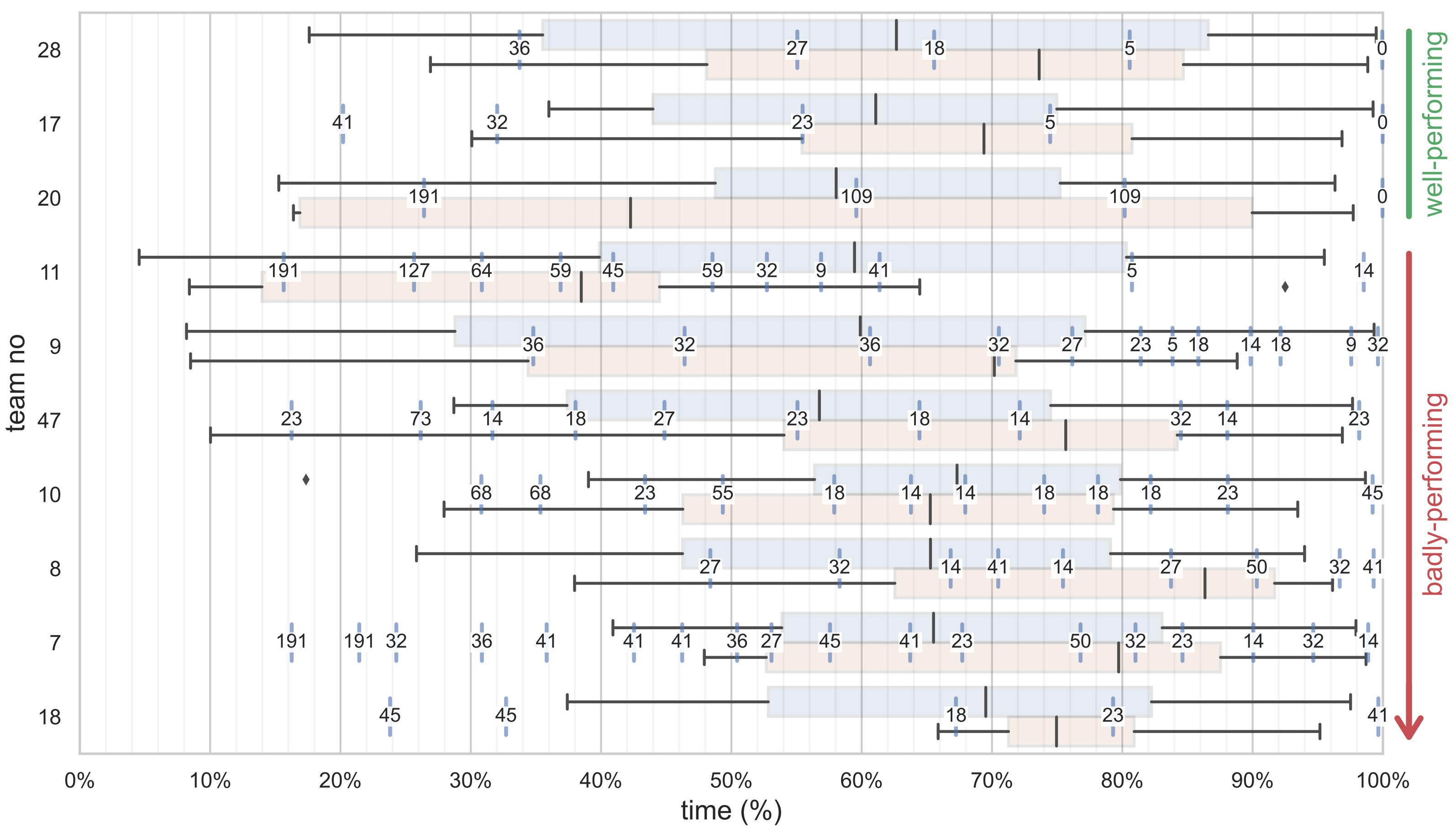


Figure 5.3 – Match and mismatch times in blue and red respectively, as normalized by the duration for each team separately. Sorting and annotations are as in Figure 5.2.

5.5 Discussion

By inspecting the outputs of the automatically inferred instructions in Table 5.2, occasionally, we see that the traditional roles of IG and IF are not maintained, even for the brief fixed time of views (i.e. within a turn). The IF, having just been switched from IG, may have their own ideas about the next best edit action, as seen in Utt. no. 199 – prior to Utt. no. 198, the views were swapped, with B being in the abstract view, meaning B was just previously the IG. Thus the interlocutors can also conduct a negotiation where they collaboratively decide which action to take. It is therefore not always the case that the IG is in the abstract view deciding and instructing what the IF in the visual view should do; and this is supported by the frequent swapping of views.

What is interesting, is that from the behavioral alignment algorithm, we see that all teams have more inferred matches than mismatches. Teams 18 and 20 that learned the least, have the highest match-to-mismatch ratio (3.6 and 4.0 matches of instructions-to-actions per mismatch respectively), as well as Team 7 (3.3). We observe that Teams 10, 11 and 17, who had the highest learning gain, comparatively have a lower ratio of matches-to-mismatches of 2.4, 1.8 and 2.3 respectively. While we could not conclude on the statistical significance, we interpret that it is important for learning that interlocutors have a good ratio of matches-to-mismatches, and not just be in total (blind) agreement with the other. This is consistent with previous research. By just inspecting performance, at the first glance, Team 20 performed well. However, they did not learn (in fact, "unlearned", as shown with a negative learning outcome). Positive learning teams seem to have conflict(s) that they resolved collaboratively while building and submitting solutions, whereas the others either had less or unresolved conflicts.

By inspecting Figure 5.3 for how the (mis)matches are distributed, and how they are positioned in relation to the costs of the submitted solutions, we observe nuances of conflicts and their resolutions. For instance, consider a high-learning Team 11. The team began with a high cost of 191%, i.e. they basically connected everything to each other, with many redundant connections. They initially misunderstood the goal of the task of connecting the graph minimally. Then, they did not agree on what to do, as indicated by the many mismatches they had (*"a mismatch period"*), which are followed by many matches ending with (at Q3 of matches) their best solution, getting very close (5%, i.e. they did not notice that they could replace a particular connection with a better/lower-cost one). Although the team could not find a correct solution, they learned a lot from this interaction flow. In comparison, Team 20 also started with that high cost of 191%. Yet, they could not resolve the conflict, as their (mis)matches did not result in them getting closer to an optimal solution, but rather still keeping to high cost solutions, and repeating these high cost submissions (109%, i.e. more than double the cost they could make with many redundant connections). Although they found a correct solution next, since they did not have the conflict resolution period that could be cued by (mis)matches, they did not end up learning.

5.6 Conclusion

In this chapter, to assess the collaboration and mutual understanding between children as they solve a problem together in the JUSThink activity, we studied their dialogue with the actions they took in the task. We presented algorithms to automatically measure the *behavioral alignment* (what was done in relation to what was said), i.e. when instructions provided by one interlocutor are either followed or not followed with physical actions by the other interlocutor. The results show that all teams are behaviorally aligned to some degree, irrespective of their task success. We observed a general trend that better performing teams tend to follow up their instructions with actions earlier in the task rather than later: this pattern was also observed for lexical alignment in Chapter 4. For learning, although the general trend is not statistically confirmed, we gained insight into the nuances of the dynamics of interaction, by inspecting the normalized (mis)match time plots alongside with the submission costs. There seems to be conflicts that may have been collaboratively resolved and resulted in learning, or may not have been resolved and had adverse effects on the learning outcomes.

While our measures discussed in this chapter for behavioral alignment (and in Chapter 4 on lexical alignment) do not show significant results for learning, we still think considering learning with performance is an important step in evaluating task success in these activities, as performance does not necessarily bring about learning (as observed in Chapter 3). Our measures still reflect some fine-grained aspects of learning in the dialogue (such as a collaborative period the interlocutors go through during the alignment process), even if we cannot conclude that overall they are linked to the final measure of learning. Our measures capture more aspects of performance than learning, as they focus on what was said about the environment, and the immediate apparent changes to the environment—the crux of the task. At a higher level, lack of understanding/learning etc. could be reflected by other phenomena (e.g. via multi-modal features, by Nasir et al. (2022a, 2022b, 2021)), other expressions ("I don't understand") etc. Since there are many ways to learn, and hence different behaviors that could result in learning, it is unsurprising that such patterns are difficult to capture.

Working with dialogues is complicated, due to (i) the nature of spontaneous speech (variable turn taking, disfluencies, etc.) and (ii) the lack of automatic evaluation criteria. For example, human annotators might instinctively be able to say that a particular team collaborated well after observing a dialogue, but it is hard to empirically pinpoint the exact reasons such a judgment was taken. Our results, albeit limited to a small dataset, highlight that in a situated educational activity, focusing simply on expressions related to the task and the actions taken within the context can still give good insights into the nuances of collaboration between the interlocutors, and its ultimate links to task success (with the awareness that there are several other aspects that remain to be observed in the dialogue such as studying the gaze patterns, prosodic features and so on). It would be very interesting in future work to look at the influence that levels of alignment have on each other, e.g. how the lexical can influence the behavioral and so on. We hope that our findings can inspire further research on the topic and contribute to the design of technologies for the support of learning.

Human-Robot Mutual Understanding Part II

6 Adaptation to a Human-Robot Activity to Probe Robot Mental Modeling

6.1 Introduction

In Part I, we developed methods a robot can utilize to assess how humans build a mutual understanding through dialogue and actions: as a use case, we designed and evaluated a robot-mediated human-human collaborative problem solving activity (the JUSThink activity) with the learning objective to improve the Computational Thinking (CT) skills of children, through reasoning on an optimization problem on networks. Motivated by the mechanisms of how humans collaborated to build a mutual understanding and converged to a shared solution, in this chapter, we revise the activity to a human-robot version.[I] In this activity, a humanoid robot collaborates with a human to construct a shared solution, via suggesting actions and (dis)agreeing with each other in turns. This adapted activity serves as an example domain in Part II, where we equip a robot with mutual modeling abilities to build a mutual understanding with a human, and investigate the potential benefits of mutual modeling on the interaction and the learning outcomes: we formalize a mutual modeling framework in terms of mental models maintained by the robot in Chapter 7, and we compare the effects of different mental models through an experiment (i.e. User Study 3) in Chapter 8.

In this chapter, as a starting point to examine what the robot needs to know to inform its mental model about the activity and the human, we design and evaluate a robot that has an incorrect mental model of the activity and no explicit model of the human as an agent. We are interested to see if such a robot can help trigger learning mechanisms by taking actions, and thereby support an effort by the human to build a mutual understanding about the task, even without the knowledge of a correct solution. Via a pre-experiment conducted remotely with 9 school children (User Study 2), we investigate (i) whether the interaction with the robot results in positive learning outcomes, (ii) how the collaboration evolves in a sequence of two collaborative activities, and (iii) whether the evolution of performance relates to the learning.

[I] The code that represents the activity and governs the interaction with the robot is publicly available online, from the GitHub Repositories https://github.com/utku-norman/justhink_world for the activity (that can also replay the data of User Study 2), and https://github.com/utku-norman/justhink-ros for the interaction. The analysis that generates the results in this chapter are available at https://github.com/utku-norman/justhink-preexp-analysis.

Results show positive learning outcomes in terms of finding better solutions in the post-test than the pre-test, suggesting that the collaboration with the robot might have helped trigger the learning mechanisms: since there was no control group in this pre-experimental study, we can not claim a causal relationship. Given that the robot was not aware of a correct solution, but rather the process of interaction, this result indicates a shift from a focus on the robot knowing and enacting a correct approach to the problem in order to facilitate learning, to rather better modeling the child. Then, the solutions of the participants tended to improve, with better performance in the second collaborative activity compared to the first: however, we did not observe an improvement trend for the specific problem-solving actions taken by the children. By inspecting the dynamics of the interaction, we see that children had no problems with (dis)agreeing with the robot: we observe a whole spectrum from episodes of "fight" where the children asserted their *beliefs* in one end, to systematic agreements with the robot including its incorrect choices in the other end. This suggests the children did "enter" into the "dialogic game" with the robot, as if it was a real dialogue. Finally, we see no evidence for a correlation between the performance change among the collaborative activities and the learning outcomes. Thus, actions and performance in the task to steadily improve and translate into acquired skills is too simplistic for how the collaboration evolved and the children learned. We believe a robot that maintains *beliefs about the human* would work at a more suitable level, with a better proxy to track the learning process; hence, we develop on this further in the next chapters.

This work corresponds to the following publications:

> **Norman**, **U.**, Chin, A., Bruno, B., & Dillenbourg, P. (2022). Efficacy of a 'misconceiving' robot to improve computational thinking in a collaborative problem solving activity: a pilot study. *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 1413–1420. https://doi.org/10.1109/RO-MAN53752.2022.9900775. (Norman, et al., 2022)

Within this work, the author of this thesis implemented the learning activity components and their interaction with the robot, designed the activity, robot behavior, and evaluation tests, directed the user study, and contributed to the data analysis.

6.1.1 Research Questions and Hypotheses

Table 6.1 presents our research questions and hypotheses for the pre-experiment. For RQ1, we postulate that the collaboration with the robot would have a positive impact in terms of the learning outcomes, as the effort to build a shared solution can help the participants realize their misconceptions. Hence, we hypothesize that they perform better in the post-test than in the pre-test; where the tests are designed as individual exercises, with the collaborative activities as counterparts that are worked on together with the robot.

For RQ2, we postulate that while collaborating with the robot, the participants would realize

Table 6.1 – Research questions and hypotheses for User Study 2. The difference between *valid* and *correct* is explained in Section 6.2.

No.	Research Question / Hypothesis
RQ1	*How are the learning outcomes after collaborating with the robot?*
H1.1	A participant provides a valid solution more in the post-test than the pre-test.
H1.2	A participant provides a correct solution more in the post-test than the pre-test.
H1.3	A participant provides a better solution (closer to a correct solution) more in the post-test than the pre-test.
RQ2	*How does performance in the task evolve during collaboration with the robot?*
H2.1	A participant submits better solutions (closer to a correct solution) later than earlier.
H2.2	A participant suggests correct actions more later than earlier.
H2.3	A participant (dis)agrees more with (in)correct robot suggestions later than earlier.
RQ3	*How does the evolution of performance in the task link to the learning outcomes?*
H3.1	The more a participant's submissions improve, the better are the learning outcomes.
H3.2	The more a participant's suggestions improve, the better are the learning outcomes.
H3.3	The more a participant's (dis)agreements improve, the better are the learning outcomes.

their misconceptions, and therefore improve their actions and solutions through time. Thus, we hypothesize that they perform better, by taking better actions and constructing better solutions, later in the collaborative activities compared to earlier. In particular, we hypothesize that they do better in the second collaborative activity, by learning from their mistakes in the first one.

For RQ3, we hypothesize that the more the participants *improve* their actions during the collaborative activities, the more they learn from them and *improve* in tests; in line with the idea "what children can do with others today, they can do alone tomorrow" (Vygotsky, 1934/1986). Our results in Chapter 3 and other studies e.g. by Kapur (2008) showed that the link between the performance in the task and the learning outcomes is not trivial: they do not necessarily have a linear or monotonic relationship. Success and failure can be productive or not in terms of the learning outcomes, as shown by Nasir et al. (2022b). Thus, in RQ3 we consider the performance *change* during collaboration, and investigate whether it is reflected in the learning outcomes as a similar kind of *change* from the pre-test to the post-test.

6.2 Design

In order to adapt the human-human JUSThink activity of Part I (described in Section 3.2) to a human-robot version, we revisit the design steps to review and revise the design choices. Thus, we: (i) identify the learning goals in Section 6.2.1, (ii) determine how to assess learning in Section 6.2.2, and (iii) design the pedagogical scenario in Section 6.2.3.

6.2.1 Learning Goals

As in the human-human JUSThink activity, we choose the minimum spanning tree (MST) problem as the underlying objective of the task, that children are not expected to be familiar with: we ask the children to reason on a network, and reasoning on a network is a general skill of CT. Thus, our desired result is that after completing this task, a participant will be able to *correctly choose a subset of connections on a given network*, so that (i) all nodes are connected to each other by some path, and (ii) the total cost on these connections is minimized.

The human-human design targeted learning goals on the underlying concepts that build to a correct understanding about the MST problem: it targeted the subskills to (i) *identify* whether a solution to a given instance of the MST problem exists, and to (ii) identify for *given* solutions whether they are valid[II] or (iii) correct[III]. As distinct from these goals, we now target more directly the ability to *construct* a valid and correct solution by oneself. This allows for an immediate comparison between how the participants solve the problem individually and how they do so collaboratively with the robot. In addition, it enables a finer grained assessment of the participants' understanding and misconceptions: we can analyze their action choices as they build a solution, that can point to any mistakes made by the participants in this process.

Concretely in the human-robot version, the learning goals are that given an instance of the MST problem, a participant will be able to:

> **LG1**: construct a valid solution to the instance (entails identifying if it is valid)
>
> **LG2**: construct a correct solution to the instance (hence identifying if it is correct)

These goals target the *create* level (along with the *understand* level) of learning objectives in the revised Bloom's taxonomy (L. W. Anderson & Krathwohl, 2001).

6.2.2 Assessment of Learning

The learning outcomes are measured by comparing the responses of the participants in the pre-test and the post-test, as in the human-human design of Part I. In contrast, instead of multiple-choice questions to evaluate the *identify* type of learning goals (as described above and in Section 3.2.2), for the current *construct* type of goals, the tests are designed as a sequence of instances of the MST problem to be solved individually by the participant. The learning goal of constructing a valid solution (LG1) is assessed by checking whether the responses (i.e. solutions) of the participant to the tests are *feasible* solutions to the corresponding instances of the MST problem or not. The learning goal of constructing a correct solution (LG2) is evaluated in two ways: by checking (i) whether the submitted solutions are *optimal* solutions or not, and (ii) how far the solutions are from the optimal solutions in terms of their cost.

[II] A *valid* or *feasible* solution to an instance of the MST problem connects all the nodes in the network of that instance to each other by some path.

[III] A *correct* or *optimal* solution is a feasible solution with the minimum possible cost.

The pre-test and the post-test contain five different instances of the MST problem: these items are in the same context with the collaborative activities, i.e. to connect rare metal mines as cheaply as possible on a fictional map of Switzerland. This is distinct from the human-human design, where the tests were in a context different than the collaborative activity (i.e. to connect houses for a postman to travel between them): it required a *far transfer* from the task to the test. Accordingly, results in Chapter 3 showed low learning gains and no evidence for a correlation between task performance and learning: these are difficult to achieve, as such a “transfer is the Achilles’ heel of formal education; students rarely spontaneously transfer what they learned in a specific situation to a new situation" (Dillenbourg, 2015). Thus, the redesigned tests assess for a more conservative goal and measure *near transfer* to different instances of the problem in the same setting: this can happen spontaneously, but might not indicate a high *abstraction* level as it does not require adaptation between contexts (Perkins & Salomon, 1992).

Regarding the format of the test items, in the human-human design, the children were asked multiple-choice questions with a single correct answer; e.g. to assess if they can *identify* the instance with the correct marked solution, by distinguishing it from other different instances with incorrect solutions—see Section 3.2.2. With the current design choice to *construct* correct solutions, the participants are no longer exposed to example solutions that might (i) help them learn while being evaluated e.g. by contrasting the options, or (ii) cause confusions, as it requires to think about *possible solutions* that can be different from the given solutions.[IV] We opt to minimize this potential confounding by providing no example solutions, and keeping the same objective throughout the tests and the collaborative activities[V]: the purpose is to make the goal as self-evident as possible, so that the choices of the participants reflect their understanding about the target concepts, without needing an intervention by an experimenter.

The tests are identical, and no feedback is given on the submitted solutions. The participant is asked to confirm before submitting. The networks of the instances in a test are created from the same underlying network structure: they are composed of 12 edges and 7 nodes, and are obtained by transforming (e.g. mirroring, rotating, scaling, and shifting) the layout and modifying the cost (e.g. by doubling or adding a constant), see Figure 6.1 for an example test item. Thus, the problem instances in the tests are of same complexity, and simpler than the instances in the collaborative activities (that have 20 edges and 10 nodes; see Section 6.2.3).

[IV] Indeed, in User Study 1, we observed that some of the participants occasionally misinterpreted the questions, and e.g. tried to find the best solution with the lowest cost among the answer options; as opposed to evaluating the options as problem instances within themselves to see whether the marked solutions satisfy validity or correctness as intended. In later studies conducted by colleagues, rephrasing the questions and an experimenter interrogating the participants to make sure they correctly understood the questions (and providing clarifications if needed) alleviated these issues: yet, through the help of an experimenter. The options to share the same network with different marked solutions could raise another issue: it can then become an exercise of contrasting cases, rather than inspecting the solutions themselves in terms of their validity or correctness which are the goal concepts. Instead of providing multiple options, asking about each option in a yes-no manner would still show positive and negative examples along the way; and can result in learning effects that are complicated to gauge.

[V] While also rephrasing the common goal: e.g. “Connect the mines while spending as little as possible, make sure miners can travel between all of them." in one question; and “What about these mines? Let’s connect them as cheaply as you can, to help the miners go from any mine to any other through the tracks." in another question.

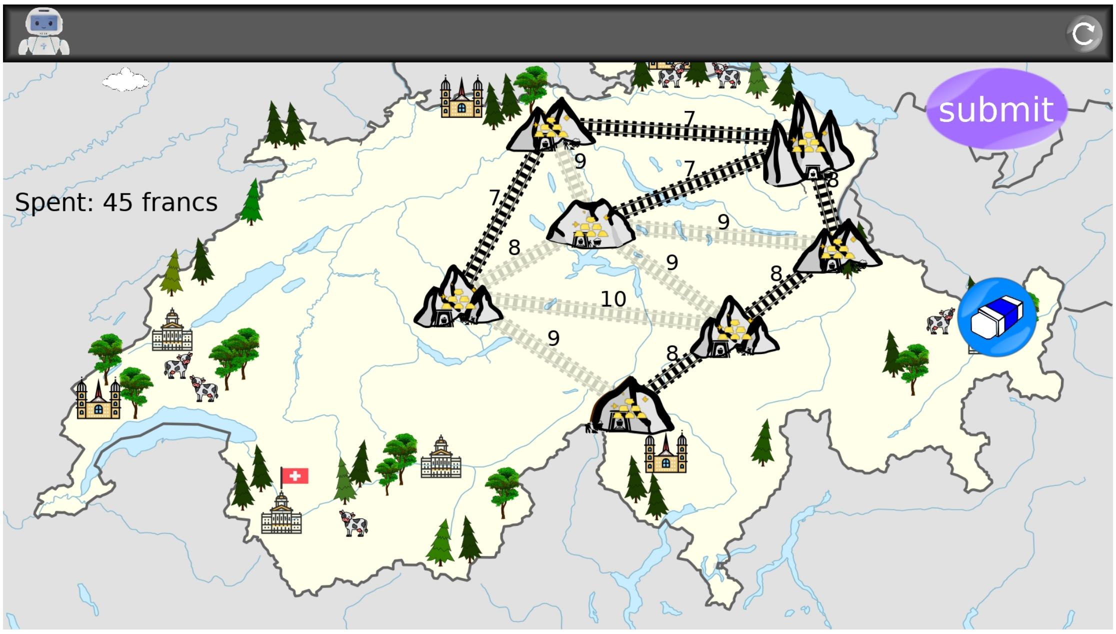


Figure 6.1 – An example test item in the JUSThink human-robot scenario. The selected connection forms an optimal solution.

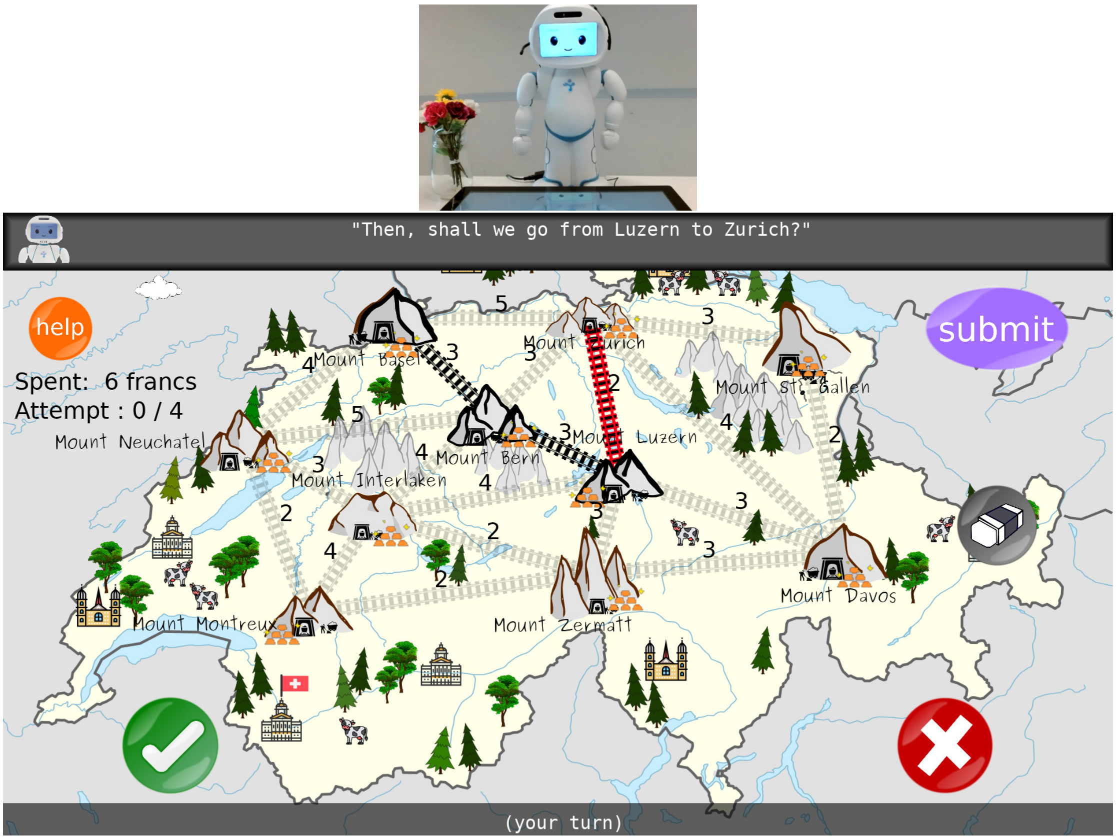


Figure 6.2 – An example game scene from the first collaborative activity in the JUSThink human-robot scenario and the robot as seen by the participant in Study 2

Table 6.2 – Sequence of activities in the JUSThink human-robot scenario. Instances indicate the number of problem instances in the group.

Activity Group	**What is the human supposed to do?**	**What does the robot do?**	**Time (min)**
Introduction	Listen to the robot, while looking at the screen that displays the context of the game.	Introduce the context and the goal of the game.	2
Tutorial	Learn how to interact with the activity on screen: (i) *pick* a connection (i.e. an edge), (ii) *clear* all the selections, and iii) *submit* a solution.	Guide the human through the tutorial. Explain the available actions.	3
Pre-test (5 instances)	In each instance, connect the mines by spending as little as possible for the network. Upon pressing the submit button, confirm the submission or continue modifying the selection. After confirmation, move on to the next instance.	Motivate the human to solve the problem instances in the test, state the (dual) goal of connecting all and spending as little as possible.	$\leq$ 5
Collaboration (2 instances)	Collaborate with the robot to connect the mines by spending as little as possible for a given network, by *suggesting* which connection to pick and *(dis)agreeing* in turns: selecting only if agreed. The human or the robot can submit $T \leq 4$ times per problem instance. Upon submission, receive feedback on whether the selection is correct.	When it is the robot's turn, suggest and (dis)agree according to a sub-optimal strategy based on greedily traversing the network; submit if there are no outgoing connections.	$\leq$ 30
Post-test (5 instances)	Solve the problem instances in the post-test, identical to the pre-test.	Motivate the human to solve the problem instances.	$\leq$ 5
Interview	Answer questions from the experimenter.	Robot is not present.	$\leq$ 15

6.2.3 Pedagogical Scenario

We design a human-robot interaction scenario where the robot orchestrates the sequence of activities described in Table 6.2. The robot introduces the scenario in the context of a game, which takes place on a fictional map of Switzerland with rare metal mines located in mountains. The goal of the game is to build a railway network to help the miners go from any mine to any other, and spend as little money as possible to build these railways. In each of the *collaborative activities*, the robot and the participant solve a problem instance together. The goal of the game is the same as the *tests*, where the participant individually solves a series of problem instances. At the end, an experimenter conducts a semi-structured *interview*.

Collaborative Activities

The human and the robot as same-status peers collaboratively construct a solution by deciding together which tracks to build; where each track, if built, connects one mine to another—see Figure 6.2. All the available connections and their costs are visible to both the human and the robot. This design choice makes the complete information about the problem instance (the possible connections and their costs) that is required to decide on the correct connections accessible to the human and (perceivably by the human) to the robot.

In contrast, in the human-human design as described in Section 3.2, there were two complementary views: only one of the views displayed the cost and availability of only the tracks that were built, and the other view showed only the tracks that are built or those available from a selected mine, without showing their cost. Furthermore, this availability and cost information was removed after each submission, and would only be accessible through an additional window that can be used to navigate through the history of previous submissions.[VI] With the current simplification for a shared view that displays all the available connections and their costs, the interaction is shifted more towards joint decision making on which connection to pick to build a correct solution: it is then no longer about discovering possible connections and their costs, and using interface features to consult this information. Thus, the robot can directly reason about the connections from its own perspective and the perspective of the human. It need not also model the information accessible to the human, or itself.

The human and the robot take turns in suggesting to build a specific connection, where the other agrees or disagrees with this suggestion, and then makes a new suggestion. A track will be built only if it is suggested by one and agreed by the other. The human as well as the

[VI] These features generated different behaviors in the collaborations, and thus allowed characterizing additional dimensions in the behaviors of the participants: for instance, Nasir et al. (2021) interpreted actions like opening the history as an indicator for the construct of *reflection* on previous solutions, and used this characterization to describe and discriminate between learner profiles that they identified. Yet, we observed in User Study 1 that the history window, although being described in the tutorial, was often ignored as a tool. The robot reminding the participants of the feature at a later submission was often either too late or unnoticed. Furthermore, hiding the connections that were not built occasionally resulted in the participants missing out on some of the crucial connections needed to build an optimal solution: therefore, it would add extra complexities to the interaction, that are not crucial for us to develop and evaluate mutual modeling skills for the robot.

Table 6.3 – The questions in the interview of Study 2

No	Question
1	How did you decide on which connections to pick?
2	How did you decide on your suggestions to the robot? Is it following some rule?
3	How does the robot decide on its own suggestions? Does it know the solution?
4	How did you decide whether to agree or disagree with the robot's suggestions?
5	How did the robot agree or disagree with your suggestions?
6	How good were the suggestions of the robot? Were they correct all the time? Incorrect?
7	How independent do you think was the robot? Was it controlled by someone?
8	What do you think was going on inside the robot's mind?
9	Did you select what you think the robot will do, when the robot asked you guess it?
10	Do you think the robot made guesses about what you would do?

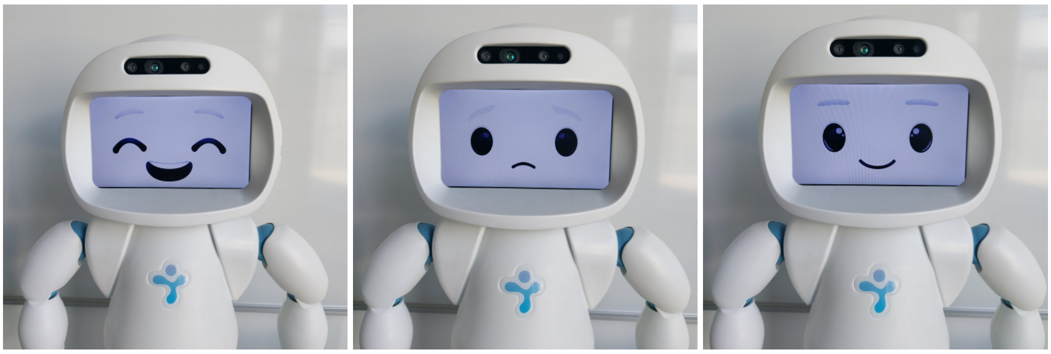
robot can submit the current solution and receive feedback on whether it is a correct solution or not: if it is a correct solution the activity ends, the solution is cleared otherwise. In each problem instance, a solution can be submitted up to $T = 4$ times. The human and the robot sequentially work on two problem instances of the same complexity, composed of 20 edges and 10 nodes. Figure 6.2 shows the first network. For instance, a *valid* solution for the problem is selecting all the possible connections: this is however not a *correct* solution as it contains many redundant connections. The second network is obtained by transforming the layout of the first network, renaming the mines, and modifying the costs on the edges.

Interview

In order to get a deeper insight on the impressions of the participants regarding the activity and their interaction with the robot, at the end of the scenario, an experimenter interviews the participants with 10 questions, listed in Table 6.3. The first half of the questions investigate how the participant perceived his/her own strategy and the strategy of the robot: how the participant and the robot made suggestions and how they agreed or disagreed with each other. Then, we ask about the participant's perception of the optimality and the autonomy of the robot, and finally if the participant and the robot made guesses about the other while taking actions. About guesses, the robot behavior is designed such that it occasionally asks the human to make a guess, see Section 6.2.4. Since we know by design of the pedagogical scenario and the robot behavior the ground truth for some of these questions, the responses give us a chance to reveal discrepancies in perceived and observed behavior of the participants.

6.2.4 Robot's Role and Behavior

We design the robot behavior such that (i) it is completely autonomous to orchestrate the interaction without a need for intervention by an experimenter, and (ii) it works together with the human to construct solutions, in such a way that elicits misconceptions and misunderstandings about the other and the activity.



(a) Happy (when human agrees) (b) Sad (when human disagrees) (c) Smile (upon suggestion)

Figure 6.3 – Example facial expressions of the robot, as given by Nasir (2022)

Throughout the Pedagogical Scenario

The robot fully automates the entire interaction by introducing the context and the goal of the game, (un)pausing the game, and moving the displayed activity to the next until the interaction concludes. Its behavior includes verbal explanations and support to motivate the human to try his/her best, facial expressions to convey emotions such as excitement (see Figure 6.3), and gestures like pointing to the participant or the self and looking at the screen e.g. to cue for whose turn it is.

During the Collaborative Activities

Unbeknownst to the participant, the robot does not know how to solve the problem correctly: it has a wrong conception of feasible and optimal solutions (i.e. no LG1 or LG2). The robot acts in a naive, simplistic and “convincible" manner, by making suggestions greedily from a node it assumes that they are at: it selects an outgoing connection with the lowest cost. This results in a sub-optimal strategy, that will traverse the map in a local-greedy manner and end up at a visited node; and hence selecting a sub-network that contains a cycle, which is a sub-optimal solution.

What the robot implicitly and functionally “believes" about the activity and its human counterpart can be summarized as follows: (i) “We are at a node and we move as we select edges.", (ii) “If we have nowhere new to select from the node we are at, we are done and we should submit.", (iii) “My strategy is correct and your strategy is incorrect.", and (iv) “If you are persistent, then you are correct.". These beliefs are manifested in the robot’s actions: when it is the robot’s turn, the robot suggests to pick one of the cheapest outgoing edges from a specific node, or submits if there is no edge to select from that node.

When it is the human’s turn, the robot either (i) asks the human to suggest what they should do (with 80% probability), or (ii) asks the human to make a guess on what it (the robot) would do (20%). The node from which the robot selects edges is not revealed to the human, and

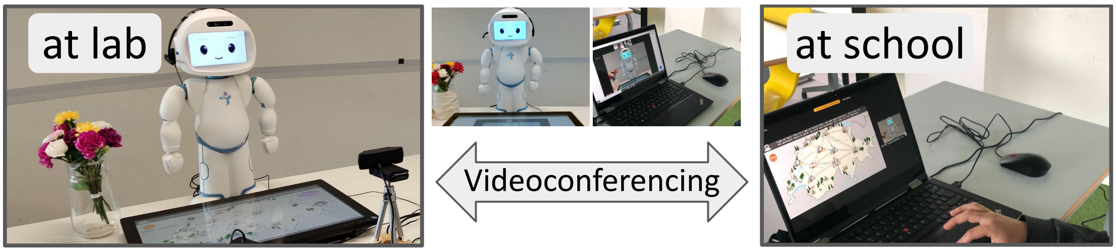


Figure 6.4 – Physical setup in Study 2: the participant interacts remotely from a school

moves as edges get suggested and selected. While responding to the human's suggestions, the robot agrees if (i) it is exactly the same edge it would pick or (ii) it is the second time the human suggests that edge; it disagrees otherwise: thus the robot is in a sense "convincible", a trait which also prevents the interaction from getting stuck if a human is to insist on selecting a particular connection.

6.3 Materials

6.3.1 User Study 2: Human-Robot Pre-experimental Pilot Study

Setup

To comply with Coronavirus Disease 2019 (COVID-19) school safety regulations, we had to resort to conducting the pilot study without bringing any external person or equipment on school premises. To make as meaningful reference to physical in-person interaction as possible, i.e. the way we intended, each child (physically located in a quiet room at school) interacted with the robot (physically located in our lab) remotely through a videoconferencing application, see Figure 6.4. We note that this was not the intended scenario, where the video-view of the robot is probably less engaging than having the robot in the room: however, this was the only possibility. Each session lasted about one hour, and followed the outline in Table 6.2. Ethical approval was granted by the EPFL Human Research Ethics Committee (HREC), No. 030-2021/06.04.2021, an amendment to the General Protocol HREC 051-2019/05.09.2019 (that authorized the human-human User Study 1 of Part I).

Participants

We collected a dataset of 9 children (3 females and 6 males), aged 10–11 years old, globally accounting for about 9 hours of interaction, of which around 5 hours spent in the collaborative problem-solving activities between the human and the robot ($\approx$ 3 hours in total in the first, and $\approx$ 2 hours in the second activity). The interviews took $M = 5.6$ min ($SD = 2.4$ min), with a maximum of 12.1 min.

6.3.2 The JUSThink Human-Robot Pre-experiment Dataset

The *JUSThink Pre-experiment Dataset* contains the interaction logs of $N = 9$ children, as a history of state transitions for each participant regarding how they constructed their solutions in each activity: individually in the tests and together with the robot in the collaborative activities. Table 6.4 presents descriptive statistics for this dataset.

The dataset is in the form of $< s_n, a_n, s_{n+1} >$ triples, where s_n and a_n denote the state of the activity and the action taken at the time step n, that results in the next state s_{n+1}.

An activity *state* encodes the current solution (i.e. the selected sub-network), the suggested connection if there is one, the agent that can take actions (for turn-taking), and the attempt number that counts the number of submissions made. It is also marked with a flag to indicate whether it is a submitting state or not (i.e. an editing state), to stage a confirmation box so that the participant can confirm submission or else continue editing the solution.

An *action* points to the type of the action, the actor that takes that action, and the object of the action such as the connection being suggested for a suggest action. We define two types of deterministic *transition functions* over any given network definition, to represent the actions and the next states available from a state: one for individual activities as in the evaluation tests, and another for the collaborative activities. This allows a uniform and standardized way of viewing and post-processing all problem solving steps in the scenario as state transition sequences. The type of the actions for an individual activity are: pick, clear and attempt-submit for editing states; and submit or cancel for submitting states. The action types in the collaborative activities are: suggest-pick, clear and attempt-submit if there is no pending suggestion, and agree or disagree to select or discard the current suggestion, respectively.

The JUSThink Human-Robot Pre-experiment Dataset is available from the following GitHub Repositories: https://github.com/utku-norman/justhink_world that implements the activities, and can visualize and replay the states and actions in the dataset, and https://github.com/utku-norman/justhink-preexp-analysis that also includes the analysis described in this chapter.

6.4 Methods

6.4.1 Measuring Learning Outcomes

We quantify the learning outcomes separately around each learning goal, by assessing the quality of the solutions of a participant in the pre-test and the post-test. For LG1, we compute the fraction of feasible solutions in the tests: from 0% (none of the solutions is feasible) to 100% (all solutions are feasible and hence valid). For LG2, we calculate the fraction of optimal solutions in the tests: from 0% (none of the solutions are correct) to 100% (all solutions are optimal and hence correct). As a finer assessment of performance in terms of how far a feasible solution is from an optimal solution, we define **error** as in Section 3.4.1, as the difference

Table 6.4 – Descriptive statistics for the JUSThink Pre-experiment Dataset ($N = 9$)

Variable	Activity	Mean	SD	Min	Max
total duration (min)	–	47.7	8.7	32.9	64.5
time in collaborative activity (min)	1	19.6	7.7	7.4	28.6
	2	13.0	5.3	5.3	19.7
number of submissions by the human	1	2.0	0.9	1	3
	2	2.0	0.9	1	3
number of suggestions by the human	1	22.8	10.0	8	34
	2	16.7	6.7	6	25
number of agreements by the human	1	16.0	8.8	2	29
	2	9.3	4.6	2	16
number of disagreements by the human	1	7.9	5.1	2	18
	2	7.9	2.8	3	13

between the cost of the solution and the cost of an optimal solution, normalized by the cost of an optimal solution.

We quantify the overall performance in a test by the *average error* per problem instance in that test. We quantify the change in the quality of responses in the post-test, compared to pre-test, on the basis of the **learning gain** of a participant, defined by the relative difference of the average error in pre-test and post-test.

6.4.2 Measuring Performance in the Collaborative Activities

Each collaborative activity allows submitting several solutions for a problem instance. Thus, we evaluate each solution in a collaborative activity separately, as the pair's best attempt for a correct solution in that activity: this may be any of the solutions, and necessarily the last solution if it is optimal. We use the *lowest error* of the solutions submitted by the human as the measure of overall performance of a participant for each collaborative activity: the lower the error, the higher the performance, with 0% meaning that an optimal solution was found. Note that the error is only computed for the feasible solutions submitted by the human participant. We measure the **performance gain** as the change of performance between the collaborative activities, by computing the relative difference of the lowest error in the first collaborative activity and in the second.

6.4.3 Characterizing Actions in Terms of Optimality

When the human and the robot are constructing a solution together, and while the human is individually constructing a solution in the tests, the actions taken can be qualified as *optimal* or not. To assign a quality label for every action, we consider the set of possible optimal solutions that can be constructed from the current solution state, by only adding more connections (or submitting). Thus, a submit action is optimal if and only if the submitted solution is optimal. A clear action is optimal, if there is no optimal solutions that can be constructed from the current state. We label a suggestion as optimal, if the suggested connection is part of at least one possible optimal solution that can be built from that state: by making more connections or submitting (and as sub-optimal otherwise). An agreement is optimal if the suggested edge is part of at least one of the possible optimal solutions, and a disagreement is optimal if the suggested edge is not part of any of the possible optimal solutions.

Optimality in this sense is towards finding an optimal solution in minimum number of actions: a sub-optimal agreement would result in selecting a connection that will need to be removed (for instance by the clear action) in order to obtain an optimal solution. Likewise, a sub-optimal disagreement is not “wrong", but rather not the quickest way to reach an optimal solution: in a later suggestion, the same or an equivalent connection could be suggested and agreed, and an optimal solution can be constructed, albeit with some more actions taken.

6.4.4 Quantifying the Trend of Change of Action Quality

For RQ3, we study the trend of improvement of actions of a specific type, such as suggestions (H3.2), agreements, and disagreements (H3.3). To quantify how and how much the quality of actions changes during a collaborative activity, we (i) filter for the specific type of actions taken by the human in that activity, then (ii) annotate each action with an action quality value by labeling optimal actions as 1 and sub-optimal actions as 0, (iii) annotate the action times as a fraction of the overall progress in that activity from 0 i.e. 0% at the beginning to 1.0 i.e. 100% at the end, and (iv) compute the slope of a fitted regression line between the action quality values and action times. Thus, a positive slope indicates an improvement, with e.g. the human suggesting optimal actions more later than earlier in the activity: the higher the slope, the stronger is the trend.

6.5 Results and Discussion

6.5.1 User Perception

Concerning their own strategy, all participants reported that they tried to select the cheapest connections that are available and that they suggested to the robot the same ones that they would pick. Participants were not sure how the robot decided on its suggestions: for two of the participants it “looks at the most efficient way", while for three it seemed “random" (“Sometimes it went for bigger numbers, sometimes for smaller").

About what the participants did when asked by the robot to make a guess on what it would do, the replies indicate they overall tended to rather follow their own choices. Five participants thought that their selection would essentially coincide with what the robot thinks, as the robot in that case would agree with it. Two followed their own choices, as they were unsure or

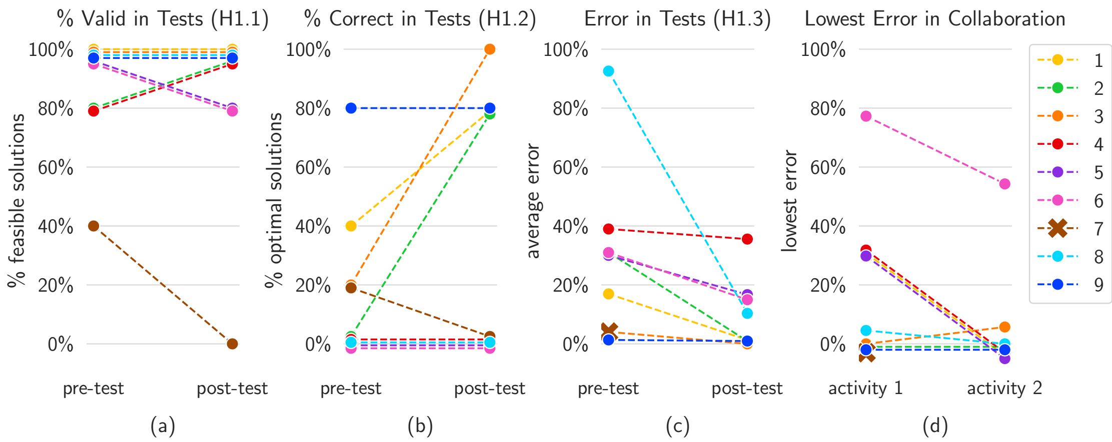


Figure 6.5 – (a-c) Learning outcomes for each participant, in terms of the quality difference of their responses from the pre-test to the post-test; (d) change in performance through the collaborative activities. The overlapping lines are demarcated by a slight offset. A cross indicates that the error for that participant can not be computed: as it is the case for Participant 7 in the post-test and the collaborative activity 2. The legend applies to all figures.

wanted to see the robot's reaction ("Because I didn't really understand the pattern it works in"). The rest occasionally ignored the request ("I picked what I thought the robot would do but sometimes I picked my own"). Therefore, in the following analyses we treat the actions taken when the robot asks to make a guess as suggestions being made to the robot.

6.5.2 RQ1 on the Learning Outcomes

Figure 6.5(a-c) illustrate how the quality of the solutions of the participants changed from the pre-test to the post-test. A Wilcoxon signed-rank test shows that the average error in the post-test is statistically significantly different than in the pre-test ($W(8) = 0.0,\ p = .008$).[VII] The error is lower in the post-test with a large effect size (Cliff's $\delta(8) = -.61$).[VIII] This supports our hypothesis H1.3: the participants performed better after collaborating with the robot, by submitting on average better solutions with lower error.

The data is inconclusive about a difference between the pre-test and the post-test in terms of submitting more feasible solutions (H1.1: $W(5) = 5.0$, $p = .63$) or more optimal solutions (H1.2: $W(4) = 1.0$, $p = .25$).[IX] On H1.1, we observe that all participants except Participant 7 submitted feasible solutions in 80-100% of the tests (see Figure 6.5(a)): this indicates that they already had a good conception for identifying valid solutions (LG1) prior to the study. For Participant 7, only 40% of the solutions were valid in the pre-test, and none in the post-test: we interpret that

[VII]Average errors in the post-test are not normally distributed (Shapiro-Wilk's $W(8) = .82$, $p = .044$).

[VIII]The magnitude of Cliff's Delta (δ) can be interpreted via thresholds $|\delta| < .147$ "negligible", $|\delta| < .33$ "small", $|\delta| < .474$ "medium", and otherwise "large" (Romano et al., 2006).

[IX]For the Wilcoxon tests for H1.1 and H1.2, there are 5 and 4 samples respectively, due to the ties (see Figure 6.5): we can not compute exact p-values with ties, and a p-value via normal approximation that allows ties is used for a sample size of typically > 50 (Conover, 1999).

Participant 7 was confused about the goals of the activity. On H1.2, we observe that one-third of the participants improved by 40 – 80%, whereas another third did not submit any optimal solutions in either of the tests: while the collaboration with the robot was beneficial for some, it did not seem to help everyone.

Overall, the results show positive learning outcomes for the participants in terms of finding better solutions, suggesting that the collaboration with the robot might have helped trigger the learning mechanisms: we can not claim a causal relationship, as there was no control group that e.g. did not collaborate with the robot. However, the interaction was not necessarily beneficial to all the participants, such as Participant 7, for whom it is likely that initial misconceptions remained unresolved.

6.5.3 RQ2 on the Evolution of Performance in the Task

Figure 6.5(d) shows how the performance changes from the first to the second collaborative activity, as measured by the best submission (i.e. the lowest error). Performance in the second activity is higher than in the first with a medium effect size (Cliff's $\delta(8) = -.38$), but the difference is not statistically significant ($W(6) = 2.0$, $p = .09$).[X] The data is consistent with H2.1: all participants except for Participant 3 performed better or the same (by finding an optimal solution in both of the activities); for Participant 3 performance only slightly decreased.

From the beginning (0% progress) to the end of the activity (100%), the median optimal suggestion times have the mean at 50.7% ($SD = 30.5\%$) and 49.4% ($SD = 31.1\%$) in the first and the second collaborative activity, respectively. In addition, median optimal (dis)agreement times have the mean 50.0% ($SD = 31.1\%$) and 59.2% ($SD = 31.4\%$). While we hypothesized in H2.2 and H2.3 that the optimal suggestions and (dis)agreements would come later in the interaction, we observe that they tend to occur throughout the interaction, spread around the middle of an activity.

All in all, participants tended to visibly improve across the collaborative activities, showing better performance in the second compared to the first, even though the robot did not know how to solve the problem. Surprisingly, within the collaborative activities, we do not observe such a trend of improvement in the quality of the actions over time: the notion that optimal actions tend to come later in the interaction is thus probably too simplistic a view for how the interaction evolves.

6.5.4 RQ3 on the Link Between Performance and Learning

Figure 6.6 shows how the participants are distributed in the space spanned by the change of performance across collaborative activities (i.e. performance gain) and learning outcomes (as measured by the learning gain). Performance gain has a weak correlation coefficient with the

[X] Lowest errors in the first and second collaborative activities are not normally distributed (Shapiro-Wilk's $W(8) = .80$, $p = .03$ and $W(8) = .48$, $p < .00001$, respectively).

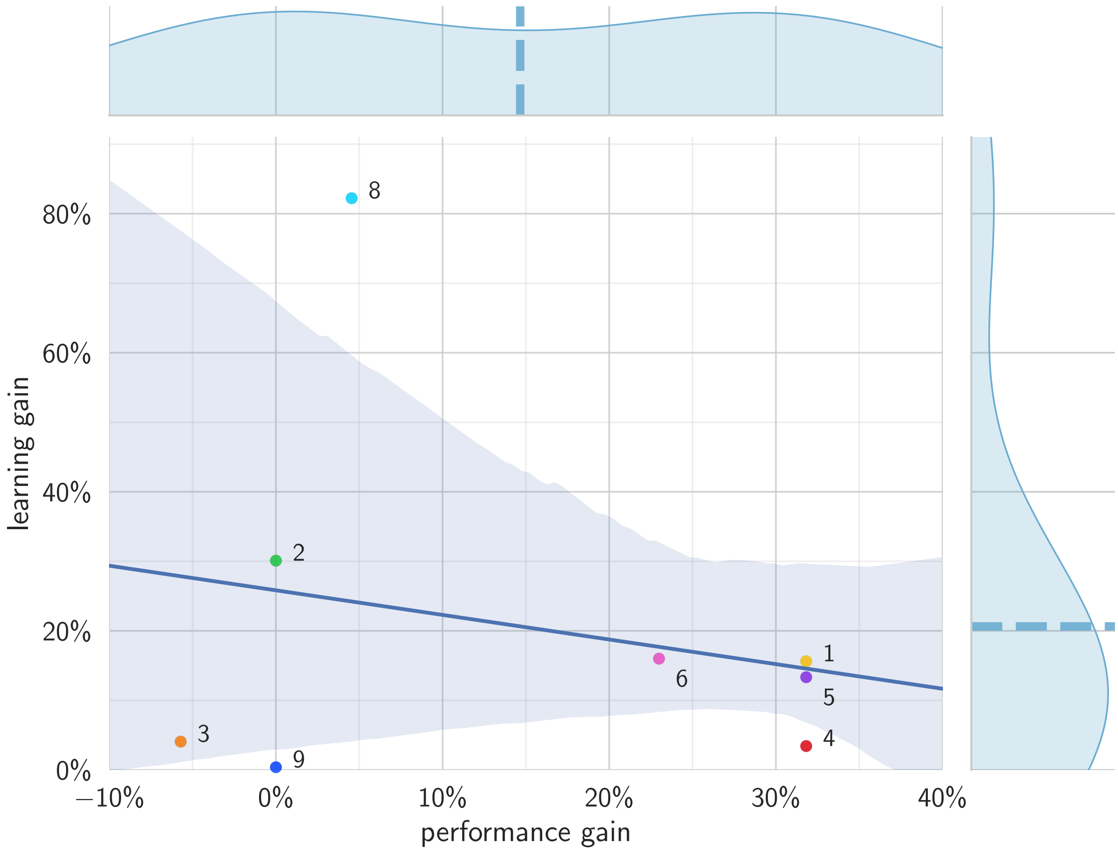


Figure 6.6 – Performance vs. learning plot of the JUSThink Pre-Experiment Dataset ($N = 9$). The gains for Participant 7 could not be computed. The mean values are shown as dashed lines, with the fit of a univariate kernel density estimate. Fit of a linear regression model is given with a solid line and a 95% confidence interval. Numbers denote the index of the participants.

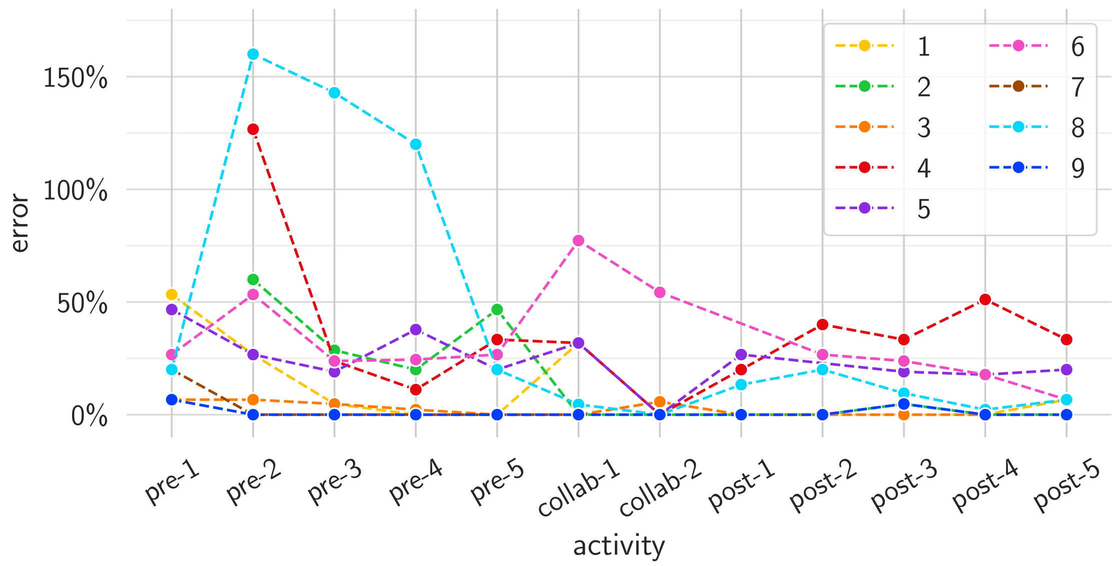


Figure 6.7 – Evolution of performance through the JUSThink scenario in Study 2

learning gain (Pearson's $r(8) = -.21$, $p = .60$).[XI] Thus, the data does not support H3.1.

When we look into the specific action types, the trend for suggestions only has a moderate correlation coefficient with the learning outcomes (first activity: Spearman's $\rho(8) = -.31$, $p = .46$; second: $\rho(8) = .64$, $p = .09$): hence, the data is inconclusive on H3.2. The trend for optimal (dis)agreements has a strong negative correlation with the learning outcomes for the first activity ($\rho(8) = .83$, $p = .01$) and a very weak correlation coefficient for the second ($\rho(8) = .07$, $p = .87$). To investigate this result in the first activity further, we separately check optimal agreements and disagreements: for these, the trends have very weak to moderate correlation coefficients ($\rho(8) = -.14$, $p = .74$ and $\rho(8) = -.41$, $p = .32$, respectively). Thus, there seems to be a non-trivial combined effect of treating them together, without a clear trend for either optimal agreements or disagreements occurring later (the median slope of the fitted regression lines for both are zero). Therefore, we conclude that H3.3 is not supported.

Figure 6.7 shows how the performance of the participants evolve through the scenario: from the pre-test instances, to collaborative activities and to post-test instances. We observe in general better performance in the collaborative activities compared to the tests, even though the robot does not know how to correctly solve the problem by itself. Furthermore, we see that high performance in the collaborative activities does not always result in a high performance in the test afterwards.

Overall, the results indicate that the trend of change in the optimality of actions is not the right level of assessing and predicting the learning outcomes: how the participants think about the activity is not necessarily synchronized with their behavior in the interaction as objectively seen by the actions they take. It is likely that their improvement reflects in the mental representations of the participants about the activity and their collaborator—the robot.

6.5.5 Observations on the Dynamics of Interaction

In the collaborative activities, the robot essentially implements a dialogue model. Thus, the goal of this study was, beyond testing the learning activity, to see how much the children "enter" into the "dialogic game" more or less as if it was a real dialogue, not just an interaction based on a few rules. For this purpose, we inspect the dynamics of interaction to see patterns of good and bad agreements and disagreements between a child and the robot over time.

Figure 6.8 shows the distribution of the optimal and sub-optimal actions taken by each participant. We observe that the children have no problems with disagreeing with the robot. Furthermore, there are some episodes of "fight", in which the robot's proposed connections are repeatedly rejected. For example, Participant 4 systematically disagreed with the robot's suggestion five times consecutively near the end of the first collaborative activity, and all the disagreements were optimal. Since the robot has an incorrect strategy, optimal disagreements

[XI] The magnitude of Spearman's ρ can be interpreted by: .00 – .19 "very weak", .20 – .39 "weak", .40 – .59 "moderate", .60 – .79 "strong", and ≥ .80 "very strong" (Evans, 1996).

are necessary to achieve the task's goal: i.e. the participant needs to say "no" to an unnecessary connection, or a costlier connection which if selected would lead to a sub-optimal solution. Thus, it seems that the child asserted his/her belief that the robot was making incorrect choices.

We also observe intervals of possible dis-engagement, in which the child agrees systematically. For example, Participant 8 simply agreed with every suggestion of the robot within the last attempt (seven times, during $t > 20$ min). The second agreement was sub-optimal, and thus the latter agreements and suggestions lead to a sub-optimal solution.

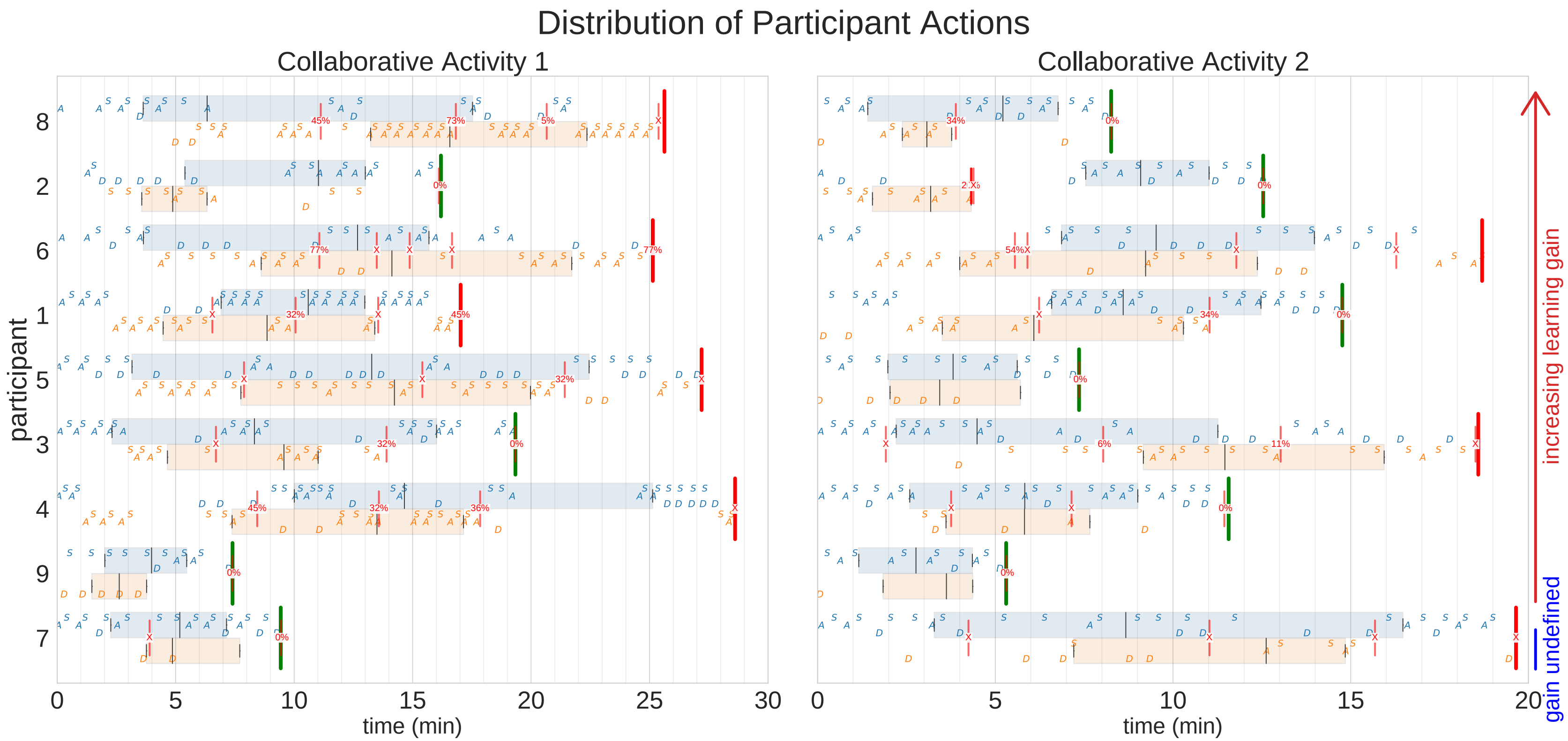


Figure 6.8 – Distribution of optimal and sub-optimal action times in the collaborative activities in Study 2, shown in blue and orange, respectively, and sorted by decreasing learning gain. Suggestions, agreements and disagreements are designated by ‘S’, ‘A’ and ‘D’, respectively. The box plots show the distribution of optimal and sub-optimal actions through the total duration of interaction. Thick lines indicate the end of the interaction, by submitting a correct solution (in green) or timing out (in red). The thin red lines indicate submission of a solution, with the number showing the error quantifying how far the submitted solution is from an optimal solution in terms of its cost, while ‘×’ indicates the submission was not feasible and therefore the error can not be computed.

6.6 Conclusion

In this chapter, we adapted the human-human collaborative problem solving activity from Part I to a human-robot version: this revised activity serves as an example domain in Part II, where our goal is to equip a robot with mutual modeling abilities to build a mutual understanding with a human, and investigate the potential benefits of mutual modeling on the interaction and the learning outcomes.

As a starting point to examine what the robot needs to know to inform its mental model about the activity and the human, we designed and evaluated a robot that has an incorrect mental model of the activity and no explicit model of the human. We conducted a pre-experiment with 9 school children. Results show positive learning outcomes for participants in terms of finding better solutions after collaboration with the robot: this suggests that the collaboration with the robot might have helped trigger learning mechanisms, though we can not claim a causal relationship. Furthermore, we observed better performance in the collaborative activities compared to the tests, even though the robot did not know how to solve the problem at hand; however, this high *collaborative* performance was not always carried over to the *individual* post-test. We did not observe a correlation between learning outcomes and the evolution of the quality of actions, which indicates a need to delve deeper into the participant's representation of the activity and the robot.

As the robot in our study is not aware of a correct solution, but rather of the process of the interaction, this indicates a shift from the focus on a robot knowing and enacting a correct approach to a problem, to rather better modeling the child. We hypothesize that a robot that maintains *beliefs* about the human would be able to work at a more suitable level, with a better proxy to track the learning process: it could guide the interaction using these beliefs and thus bring about learning more effectively. This is exactly what we investigate further in the next chapters. Also, we believe that this type of behavior is easy to port to other activities; as a correct approach need not be included in the behavioral design. Lastly, although the study was conducted with a remote-telepresent robot and was limited to a small number of participants, we believe the results allow further investigation.

7 Robot Mental Modeling Framework to Build Mutual Understanding

7.1 Introduction

In this chapter, we present a mental modeling framework to equip a robot with mutual understanding abilities, to facilitate its interaction with a human. We focus on the robot's *agency*, where an *agent* can be seen as anything that interacts with its *environment* by perceiving it and acting on it. The environment can contain other agents, that can be human agents or artificial agents: we refer to the remaining objects that are not agents as constituting the *world*. The world presents a particular context of interaction, within an *activity* that is shaped by a *domain*, where the domain can offer a *goal* for the agent to try to achieve in its interaction.

Our framework enables the robot to build and maintain a *mental model* of its environment: with its own *beliefs* about the world, the human, as well as the human's beliefs about the robot. We define an example domain, the *JUSThink domain*, that generalizes the human-robot collaborative problem solving activity in Chapter 6, as well as its initial robot-mediated human-human version in Part I. On this domain, we illustrate (i) what and how the robot can *think* about its environment, i.e. the world in the activity and the human, (ii) how it can decide on an action based on what it thinks, and (iii) how it can generate linguistic expressions and explanations about it.[I] Using this framework, in the next chapter (Chapter 8), we explore the potential benefits of these robot abilities on the interaction and the learning outcomes, by comparing the effects of robot behaviors that are guided by different mental models through an experiment.

In our framework, the robot is specifically designed to behave as an *intentional system*, i.e. a rational agent with its own beliefs, desires, and other mental states about its world and the other agents, that does what it ought to do given those beliefs and desires.[II] As discussed in

[I] A reference implementation for the JUSThink domain and the mental modeling framework presented here is publicly available online, from the GitHub Repository https://github.com/utku-norman/justhink_world. We use the `pomdp_py` Python package by Zheng and Tellex (2020) for representation.

[II] We refer to an agent as *rational*, even if it is not able to plan optimally due to e.g. incorrect premises or information: "independent of the accuracy of the model, an agent is rational if it behaves in accordance with its model of the world. A rational agent with a correct model is sound." (Goodwin, 1995).

Section 1.2.4, Dennett (1971, 1987) distinguishes between three basic stances or strategies, that humans adopt to understand an object's behavior: by adopting (i) the *physical stance*, we predict that a stone will fall to the ground if we release it by its physical constitution and laws of physics; by the (ii) *design stance* we predict that an alarm clock will sound at set time assuming it is designed that way; and by (ii) the *intentional stance* we make predictions by treating the object as an intentional system with mental states. An alarm clock is too simple that this anthropomorphism is unnecessary: no further *predictive power* is gained by adopting the intentional stance. Meanwhile, towards people, the intentional stance often is the only practical strategy. Likewise, it is also intuitive and likely very practical for us to consider a robot that we would play a game with as a rational agent who *wants* to win, who *knows* the rules, who *thinks* its choice is correct etc.: "an agent is a system that is most conveniently described by the intentional stance; the simplest consistent description requires the intentional stance." (Wooldridge & Jennings, 1995). We design the robot to *explicitly* maintain its mental states, and manifest them in what it says and does: we believe this can induce the child playing with the robot to adopt the intentional stance towards the robot.[III]

In our framework, the robot as an intentional system does *intention modeling*, i.e. it attributes mental states to the other agents in its environment about the world, the same way its has its own beliefs about the world. As discussed in Section 2.4, in Artificial Intelligence (AI) as the "study of constructing machine intelligence from an agentive perspective" (Zambak, 2009), a central area of research focuses on designing agents that can have an effective interaction with other agents, where the agent would be capable of reasoning about the actions, goals and beliefs of other agents (Albrecht & Stone, 2018). An agent can construct a *model* of the other agents, where in general a model is a *mapping* from the observed history of interaction between the modeling and the modeled agent to predicted properties of interest regarding the modeled agent. Dennett's intentional stance allows for a useful characterization of these modeling methods; modeling is either *intentional* or it is not, where intention modeling attributes meaningful mental states to the modeled agents like beliefs and desires in order to explain their behavior (Doshi et al., 2020). Within the model, intention modeling can entail the modeled agent to also model others, that results in *recursive modeling*: the agent can explicitly describe nested beliefs in the form "I think that you think that I think (...)". We believe the robot as an intentional system with its own beliefs, that also performs intention modeling by ascribing beliefs to the other, can build a mutual understanding with the other through their interaction: this can improve the quality of the interactions, and hopefully the learning outcomes in an educational setting.

We formalize the robot's interaction with its environment and model the sequential decision

[III] From the design stance, conspicuously for its designers, the robot of course acts according to the way its program prescribes, that could be used to describe and predict its behavior. However, as the activity goal and the robot's reasoning becomes more complex, the predictive power gained from the intentional stance can surpass by that from the design stance: it can become more practical and intuitive to reason with the robot's actions in terms of beliefs than its algorithm. We believe this could tend to happen for non-trivial collaborative activities with the purposes to achieve the activity goals: especially if the robot is designed that way, i.e. promoting the intentional stance.

making in this multi-agent environment in a finitely-nested Interactive-POMDP (I-POMDP)-like manner, by focusing on the state of the world and the mental state representations the robot can have about the world and the agent, i.e. the human learner. We leverage the fact that the interaction occurs in a well-defined activity and in a designed environment, to readily represent the possible, relevant beliefs about the world: from the robot's and the human's point of view, as inferred by the robot. Actions *leak* information, on what we think, the other thinks, the other thinks that we think, etc. Thus, we consider the actions taken by the human and the robot as apparent indicators for the beliefs of the human and the robot, with less uncertainty in terms of perception and interpretation: e.g. human's speech as an interaction modality could serve as an alternative realization of this framework, with a different uncertainty attached to the inferences, that we do not explore in this thesis.[IV] As a result, the robot can maintain a mental model of the human by observing the actions. Then, it can use its mental model to *steer* the human towards conceptual change. Using the mental model, it can detect miscommunication phenomena: misunderstandings, non-understandings, disagreements and misconceptions. We believe the robot can address these with the help of its mental model that represents the human, in such a way that can improve learning.

The chapter is organized as follows. Firstly, we describe what a world is in the example domain: how we represent it, how a single agent or multiple agents can interact with it, and what they can think about their environment in an activity, in Section 7.2. Then, we formalize how the agent can represent and maintain its mental model about its environment in Section 7.3. Next, we look into how the agent can use its mental model to reason about its environment, and generate linguistic descriptions and explanations in Section 7.4. We conclude in Section 7.6.

7.2 An Example Domain: JUSThink

7.2.1 A JUSThink World

In a *world* within the JUSThink domain, there exist rare metal mines located in the mountains on a fictional map of Switzerland. The mountains can have proper names, e.g. "Mount Luzern". A subset of mine pairs can be directly connected to each other with railway tracks, and each track has a cost. The goal in this domain is to build a railway network to help the miners go from any mine to any other through some path (i.e. a route along the tracks), and spend as little money as possible to build these railways: this is in human terms. Meanwhile, the more workable form for an artificial agent is to minimize a known cost function, i.e. the total cost on a subset of edges such that there is a path between every node pair. This goal corresponds to the minimum spanning tree (MST) problem, which is the learning context we target in this thesis—see Section 3.2.1 for details on the learning goals.

[IV] Machines are currently not so good at understanding natural language online and in real-time: especially of children as we discussed for Automatic Speech Recognition (ASR) results in Section 4.2.3.

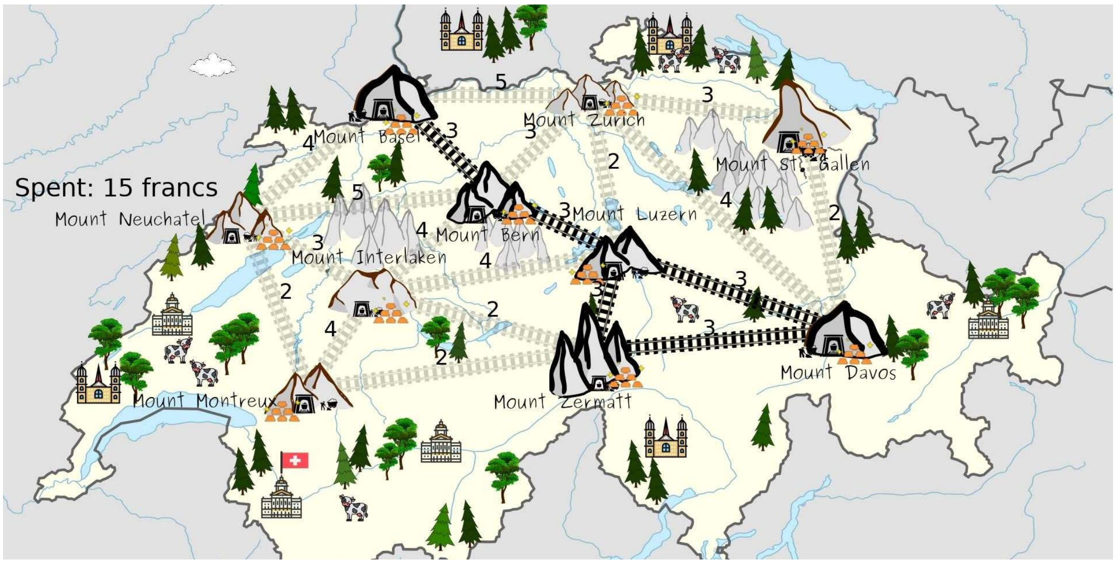


Figure 7.1 – An example situation for a world in the JUSThink domain

Figure 7.1 shows an example *situation* for a world in the JUSThink domain: the mines at "Mount Davos" and "Mount Zermatt", in brief Davos and Zermatt, are connected to each other with a track that costs three francs, there is a possible connection between Zermatt and Montreux that costs two francs, and there is no possible connection to directly link Montreux and Luzern. Different worlds can have different maps on the background, different names and positions for the mines, and different possible connections and costs between the mines. For instance, in the human-robot collaborative activity in Chapter 6, there is a sequence of two collaborative activities: the world in the second activity has different costs and topology.

A situation in the world is represented as a *world state*, that encodes the information on the current *solution*: i.e. the selected sub-network as a set of edges, and the background network with the node names, their positions, the possible connections between them and their costs. A *transition function* describes which actions are available at a given state, and determines how they change by state transitions. For example in the situation in Figure 7.1, a pick(Zermatt, Montreux) action will result in a new state with Zermatt and Montreux connected to each other, in addition to the present connections. A minimal set for the type of actions that allows solving this task correctly is {pick, submit}: it consists of pick actions to connect any possible mine pair, and submit action to finalize the current selection. To allow for changes in the selection, a more comprehensive set of actions can include a clear action to remove all the selections, or an unpick action to remove a selected connection for a subset of selections.

An agent can solve the task optimally by following a deterministic, greedy algorithm such as Jarník's algorithm or Kruskal's algorithm[V]. If it is restricted in the world by design, modeled

[V] Jarník's algorithm is also known as Prim's algorithm or Dijkstra-Jarník-Prim (DJP) algorithm: developed by Jarník (1930), then rediscovered by Prim (1957) and Dijkstra (1959). Kruskal's algorithm is described by Kruskal (1956), and rediscovered by Loberman and Weinberger (1957). Algorithms for the MST problem go back to Borůvka through its transition function, such that one can only pick connections from the already selected nodes, Kruskal's algorithm might not be applicable as it does not respect this rule and can try to select edges from nodes that are not connected yet.[VI] Alternatively, the agent, likely an artificial agent, can use a reinforcement learning (RL) algorithm such as Q-learning (Watkins & Dayan, 1992): the agent can receive the edge costs as negative rewards for when it selects edges, and a large positive reward when an optimal selection is made. Then, it can learn an optimal *policy* for the given world. However, this method needs that there is no limit on the number of times an action can be tried out in any state: a simulation of the world as it is intrinsically based on trial and error. Furthermore, the learned policy does not translate directly into a general set of rules to solve the task in other worlds, with different connections and costs: thus, it would not be clear to reason with and communicate why a choice is correct.

7.2.2 Single-Agent Activities

As a starting point, characterizing how the robot can interact with the activity by itself would help establish its relationship with the world, as well as its reasoning about the other agents. In a single-agent activity in the JUSThink domain, there are no other agents in the environment: one agent (an artificial agent or a human) is asked to solve the task on a given world by itself. It could be a human learner presented with the task, as an exercise or a game, to be done alone. Or else, it could be an artificial agent that is presented with the task, and we want the agent to *learn* a policy by some machine learning method, or test if a particular algorithm is correct.

What the agent can think about its environment is only about the world, as there are no other agents. Firstly, the agent can form *explicit beliefs* about the state of the world based on its perception: e.g. "I think Davos-Zermatt (connection) exists.", "I think Davos-Zermatt costs three.", "I think Davos-Zermatt is selected.". These can be graded to indicate the uncertainty of the agent in perceiving the world, e.g. via quantifying by a probability distribution.[VII] We do not investigate this type of beliefs in this thesis, and assume the world can be fully observed by the robot. Then, the agent can form beliefs based on its own strategy, about the *correctness of the choices* it can make: "I think Basel-Bern is a correct choice.", "I think Basel-Zurich is *not* a correct choice.". These beliefs can describe and predict the agent's behavior: the agent can then select the connection Basel-Bern, and not select Basel-Zurich. The agent can also form beliefs about its own overall strategy: "I think my strategy is correct.". Beliefs on the strategy can influence the beliefs on the choices, which in turn influence the beliefs on the strategy.

(1926a, 1926b)—see Nešetřil et al. (2001) for the translations of the Borůvka's works and a historical survey.

[VI] The transition function in the human-human collaborative activity in Chapter 3 (User Study 1) and the final design of its human-robot version for the experiment (User Study 3) have this restriction, to allow for inspecting the patterns of *growing* a solution. Meanwhile, in the first design of the human-robot version for the pre-experiment (User Study 2 in Chapter 6), the transition function allowed the agents to pick any connection: this lead to complications in the robot strategy and we reinstated the rule in the transition function of final design.

[VII] from a subjective Bayesian interpretation of probability theory, that "treats the degree of probability as a measure of the feelings of certainty or uncertainty, of belief or doubt, which may be aroused in us by certain assertions or conjectures" (Popper, 1935/2002). See Gillies (2000) for a survey on the interpretations of probability.

As a search problem A single-agent activity can be formulated as a *search problem* as follows:

$$\text{search problem } = \langle \mathcal{S}, s_{start}, Actions, Succ, Cost, IsEnd \rangle, \tag{7.1}$$

Equation 7.1 – Search problem formulation of a single-agent JUSThink activity

where $\mathcal{S}$ is the set of states, i.e. the selectable sub-networks; $s_{start} \in \mathcal{S}$ is the starting state i.e. the empty solution; $Actions(s)$ is the set of actions available at s with $Actions: \mathcal{S} \rightarrow \mathcal{A}$, such as pick actions for the remaining connections; $Succ: \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{S}$ is the successor function, e.g. transitioning into a new solution for a pick action; $Cost: \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$ is the cost function, i.e. the cost of the action; and $IsEnd: \mathcal{S} \rightarrow \{\text{True}, \text{False}\}$ is the goal test, i.e. True for $s \in \{s \text{ is an optimal solution}\}$ and False otherwise. The *planning* problem for the agent is about how it should pick its next action: to choose an optimal action, that steps towards an optimal solution. The solution to this planning problem is a sequence of actions.

As an MDP MDPs generalize search problems for uncertainty in state transitions (Bellman, 1957). A search problem has an equivalent deterministic MDP. By formulating it as an MDP, an artificial agent can use methods like Q-learning to solve the planning problem. The equivalent MDP formulation for the search problem is a 5-tuple:

$$\text{MDP} = \langle \mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \gamma \rangle, \tag{7.2}$$

Equation 7.2 – MDP formulation of a single-agent JUSThink activity

where $\mathcal{S}$ is the finite set of states of the world (same as in the search problem); $\mathcal{A}$ is a finite set of actions, e.g. $\{\mathsf{pick}, \mathsf{submit}, \ldots\}$; $\mathcal{T}: \mathcal{S} \times \mathcal{A} \rightarrow \Pi(\mathcal{S})$ is a transition function, for equivalence $\mathcal{T}(s', s, a) = P(s'|s, a) = 1$ if $s' = Succ(s, a)$ and 0 otherwise; $\mathcal{R}: \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$ is an immediate reward function such that $\mathcal{R}(s, a, s') = -Cost(s, a)$; and $\gamma \in [0, 1]$ is the discount factor for reward on later actions: $\gamma = 1$ for no discounting. A solution to an MDP is a policy $\pi: \mathcal{S} \rightarrow \mathcal{A}$.

As a POMDP POMDPs generalize MDPs to partially observable contexts (Kaelbling et al., 1998). An agent no longer knows the state, but maintains a *belief* over states, i.e. a probability distribution over the states. In a single-agent activity, and also for multi-agent activities, these can define the agent's beliefs on what is an optimal solution. This mathematical representation of beliefs entails *linguistic expressions*; at least so that the agent can verbalize its beliefs. Then, given a belief as a probability distribution, the solution with the highest probability can be decomposed into a corresponding probability assignment over the edges of the solution as those that should be picked, about the correctness of choices: e.g. if high, "I think Basel-Bern is a correct choice." or, if low "I think Basel-Zurich is not a correct choice.".

A single-agent activity can be formulated as a POMDP, as a 7-tuple:

$$\text{POMDP} = \langle \mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \Omega, \mathcal{O}, \gamma \rangle, \tag{7.3}$$

Equation 7.3 – POMDP formulation of a single-agent JUSThink activity

where $\mathcal{A}, \mathcal{T}, \mathcal{R}$ are as in the MDP formulation; $\mathcal{S} = \langle \mathcal{X}, \mathcal{Y} \rangle$ is decomposed into a fully-observable state space $\mathcal{X}$ (selectable sub-networks, given the world state is fully-observable), and a partially-observable state space $\mathcal{Y}$ (optimal sub-networks); Ω is the finite set of observations the agent can experience in the world, e.g. a "prize" for finding an optimal solution, or an observation on the world state: perfectly if the world is fully-observable or with some uncertainty if it is not. $\mathcal{O} = \mathcal{S} \times \mathcal{A} \to \Pi(\Omega)$ is the observation function $O(s', a, o) = P(o|s')$. $o \in \Omega$ describes the world states with $y \in \mathcal{Y}$ that are consistent with $x \in \mathcal{X}$, and $\gamma \in [0, 1]$ is the discount factor.

7.2.3 Multi-Agent Activities

Observer's world (with an observing robot) In a multi-agent activity, it could be that the robot observes how human(s) solve the task, or vice versa. For instance, in the robot-mediated human-human JUSThink activity in Part I, the robot observes two humans solve the task together. In the tests for the assessment of learning in the JUSThink human-robot scenario of this part (Part II), the robot observes the human solve the task individually, multiple times.

What the robot as an agent can think about its environment is no longer just about the world, as there are also other agents. The robot can attribute to the human(s) the same kind of beliefs it can have about the world, recursively: e.g. perceptually about the state of the world, "(I think that) the human thinks Davos-Zermatt is selected.". Then, it can think about the human's strategy, with attributions about the correctness of choices: "(I think that) the human thinks Basel-Bern is a correct choice.". The robot can infer these beliefs from what the human does, and ascribe them to the human. For instance, if the human connects Luzern to Zurich, the robot can form: "I think that the human thinks that Luzern-Zurich is a correct choice.". This inference is based on treating actions as intended by the human, assuming the human wanted to do it (Cohen et al., 1981). These beliefs can be graded, as they come with some uncertainty: to represent this quantitatively, the robot can attach a probability value on the beliefs.

There are also other kinds of beliefs, or mental states in general, the robot can have about the other agent; such as when humans project intentions on computers ("computers don't like me", as many people say). In our framework, these would count as zeroth-order beliefs about the other; but not as first-order beliefs, i.e. attributions to the other as beliefs held by the other. However, they can be updated similarly by observation: if the computer does not turn on, then the agent can form the belief: "I believe that the computer does not like me.". In our framework allows for extending to beliefs of these kinds. See a more extensive discussion about the shortcomings of our modeling framework in Section 9.2.2.

Collaborative world (with a collaborating robot) In a multi-agent activity, it could be that the robot works together with human(s) to solve the task. For example, in the JUSThink human-robot scenario, the human and the robot as same-status peers collaboratively construct a solution by deciding together which tracks to build, via suggesting connections and (dis)agreeing with each other in turns, as described in Section 6.2.3.

What the robot as an agent can think about its environment, in addition to its own beliefs about the world and the human's beliefs about the world, now also includes the human's beliefs about the robot's beliefs about the world. For instance, if the robot connects Luzern to Zermatt, this gives a cue to the human regarding the robot's beliefs about the world. The robot can form a belief: "I think that the human thinks that I think that Luzern-Zermatt is a correct choice.". Overall, in the mental model built and maintained by the robot, the beliefs we described can be grouped in terms of the following levels: *zeroth-order beliefs* are about the world itself, *first-order beliefs* are beliefs ascribed to the human about the world, and *second-order beliefs* are beliefs ascribed to the human about the robot's beliefs about the world.[VIII]

As an I-POMDP POMDPs simply ignore other agents, treating their influence on the transitions and observations as noise: I-POMDPs generalize POMDPs to handle their presence (P. Gmytrasiewicz & Doshi, 2005): they consider the subjective view of an individual agent in the multi-agent system, rather than the objective view presented by decentralized POMDPs (Bernstein et al., 2002). I-POMDPs extend the RMM method by P. J. Gmytrasiewicz and Durfee (2000), and relate to the decision-theoretic or epistemic approaches to game theory (P. Gmytrasiewicz & Doshi, 2005). The normative aspect of decision theory is about identifying how a perfectly rational agent is to make the optimal decisions; while the prescriptive aspect is for the agent to select its own action in practice, given constraints and assumptions. In this thesis, we focus more on the descriptive aspect of decision theory to describe observed behaviors, that analyzes how the agents make the decisions they do. An I-POMDP of agent i is given by:

$$\text{I-POMDP}_i = \langle IS_i, \mathcal{A}, \mathcal{T}_i, \Omega_i, \mathcal{O}_i, \mathcal{R}_i \rangle, \tag{7.4}$$

Equation 7.4 – I-POMDP formulation of a multi-agent JUSThink activity

where $IS_i = \mathcal{S} \times M_j$ is the set of *interactive* states, such that $\mathcal{S}$ is the state of the world and M_j is the set of possible models of agent j. Each model $m_j \in M_j$ is defined as a triple $m_j = \langle h_j, f_j, O_j \rangle$, where $f_j : H_j \to \Delta(\mathcal{A}_j)$ is the agent j's function that maps possible histories H_j to $\Delta(\mathcal{S})$, i.e. the set of probability distributions over a state space $\mathcal{S}$. Finally, O_j is a function for

[VIII] In the three agents approach for second-order modeling by A. D. Jacq (2020) and A. D. Jacq et al. (2016), these levels are considered as independent agents that do not perform mutual modelling. The model becomes non-recursive, breaking an infinite regress if the robot were to use its own architecture recursively to model the other agent, e.g. as in Breazeal et al. (2006). Here, we use a mental model similar to the finitely-nested I-POMDP that rather keeps the recursive nature of the beliefs and breaks regress via constructing finitely nested states bottom-up.

the way the agent receives its observations from its environment. Solution to an I-POMDP is a policy π on a history of observations to $P(a)$. A compressed representation is a probability distribution over states: i.e. the agent's belief $b(y)$, which is the degree of belief in the world state $y \in \mathcal{Y}$, and is described as a probability distribution over $\mathcal{Y}$. b in this domain gives agent's beliefs on what is an optimal solution. Agent's desires are encoded in its reward function.

Agent's belief $b_t(y)$ in environment y at time t is updated via a deterministic function of prior belief b_{t-1}, observation o_t, agent state x_t, and world state y by satisfying $b_t(y) \propto P(o_t|x_t, y)P(x_t|x_{t-1}, y, a_{t-1})b_{t-1}(y)$: i.e. by a Bayesian update.[IX] This approach is e.g. used by C. L. Baker et al. (2017) that refined C. L. Baker et al. (2009, 2011) to model mentalizing.

As a finitely-nested I-POMDP In an I-POMDP, the other agent's inferred model is *sub-intentional*: it does not represent the beliefs of other agents, but is only about the world. Finitely-nested I-POMDPs overcome this limitation via recursion, as described in P. Gmytrasiewicz and Doshi (2005). A finitely-nested I-POMDP of agent i is as given by the 6-tuple:

$$\text{I-POMDP}_{i,l} = \langle IS_{i,l}, \mathcal{A}, \mathcal{T}_i, \Omega_i, \mathcal{O}_i, \mathcal{R}_i \rangle, \tag{7.5}$$

Equation 7.5 – Finitely-nested I-POMDP formulation of a multi-agent JUSThink activity

where parameter l is the number of levels that the finitely-nested I-POMDP admits. Interactive state $IS_{i,k}$ of agent i is formally defined as:

$$IS_{i,0} = S, \qquad \Theta_{j,0} = \{\langle b_{j,0}, \hat{\theta}_j \rangle : b_{j,0} \in \Delta(IS_{j,0})\}, \qquad M_{i,0} = \Theta_{j,0} \cup SM_j \tag{7.6}$$
$$IS_{i,1} = S \times M_{j,0}, \qquad \Theta_{j,1} = \{\langle b_{j,1}, \hat{\theta}_j \rangle : b_{j,1} \in \Delta(IS_{j,1})\}, \qquad M_{i,1} = \Theta_{j,1} \cup M_{j,0} \tag{7.7}$$
$$\cdot \qquad\qquad \cdot \qquad\qquad \cdot$$
$$IS_{i,l} = S \times M_{j,l-1}, \qquad \Theta_{j,l} = \{\langle b_{j,l}, \hat{\theta}_j \rangle : b_{j,l} \in \Delta(IS_{j,l})\}, \qquad M_{i,l} = \Theta_{j,l} \cup M_{j,l-1} \tag{7.8}$$

$\Theta_{j,l}\langle b_{j,l}, \hat{\theta}_j \rangle$ is a set of *types* for agent j, gathering the parameters to compute the agent's behavior: it contains j's beliefs $b_{j,l}$, and j's frame $\hat{\theta}_j = \langle \mathcal{A}_j, \Omega_j, \mathcal{T}_j, \Omega_j, \mathcal{O}_j, \mathcal{R}_j, OC_j \rangle$ with optimality criterion OC_j on how the agent uses the rewards its receives. An agent's first-order beliefs are probability distributions over world states and zeroth-order models $M_{j,0}$ of the other agent. Zeroth-order belief $b_{j,0}$ for agent j is a belief over world states S. $\mathtt{k}^{\text{th}}$-order belief $b_{j,k}$ is defined over world states and models consisting of types that admit beliefs of up to level `k-1`. We use this recursive structure that formalizes "I think that you think p." to construct the mental model of the robot, and a simplified belief update rule to maintain the model. In this thesis, we restricted ourselves to $l = 2$, i.e. up to and including the second-order beliefs.[X]

POMDPs quickly become impractical and suffer from the curse of dimensionality, as the

[IX] "Bayesian inference stipulates how rational learners should update their beliefs in the light of evidence." (Griffiths et al., 2008).

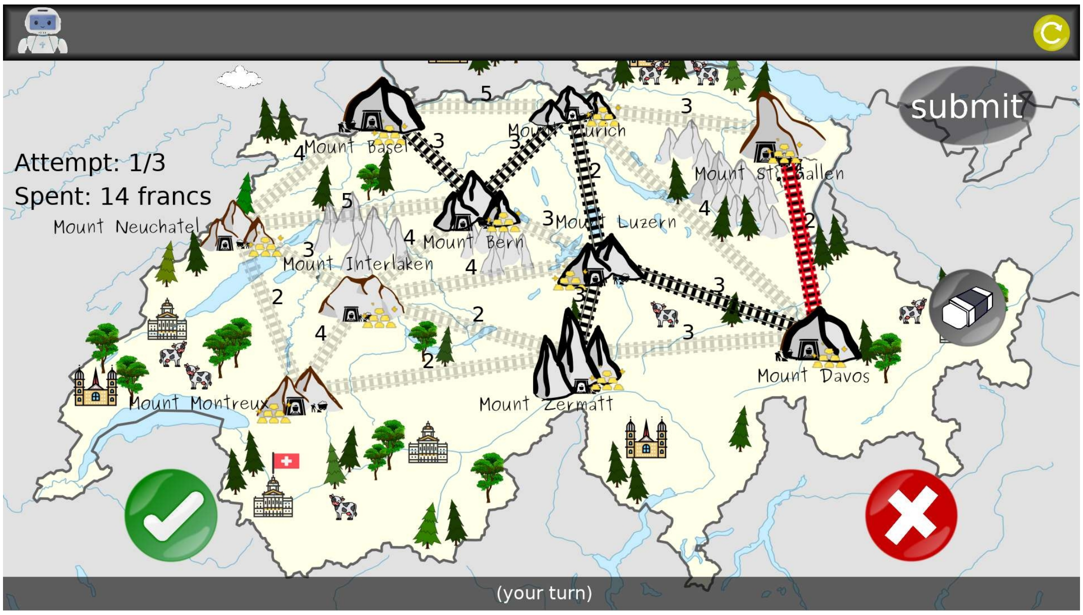
[X] See the opening of Clark and Marshall (1981) that cites *Knots* by R.D. Laing for the higher-order adventures of Jack and Jill, that goes like: "Jack thinks he does not know what he thinks Jill thinks he does not know. . . ".

size of the belief scales exponentially with the number of states. I-POMDPs$_{i,l}$ also have the added depth of recursive levels: it becomes equivalent to $O(M^l)$ POMDPs, which is undecidable as also for POMDPs (P. Gmytrasiewicz & Doshi, 2005). To alleviate this and reduce the dimensionality, the hypothesis space of beliefs is in general discretized. In addition, the state space has been factorized, e.g. into independent objects by Wandzel et al. (2019) or sampling methods have been used to solve the planning problem, e.g. via Monte-Carlo tree search for large POMDPs by Silver and Veness (2010). Yet, it still remains a challenge with practical considerations for complex games as in the JUSThink domain with more than a few nodes and edges and the human involved: the model overall can still fail to accurately estimate beliefs of a human due to assumptions regarding the human such as rationality, e.g. as observed by Rueben et al. (2022).

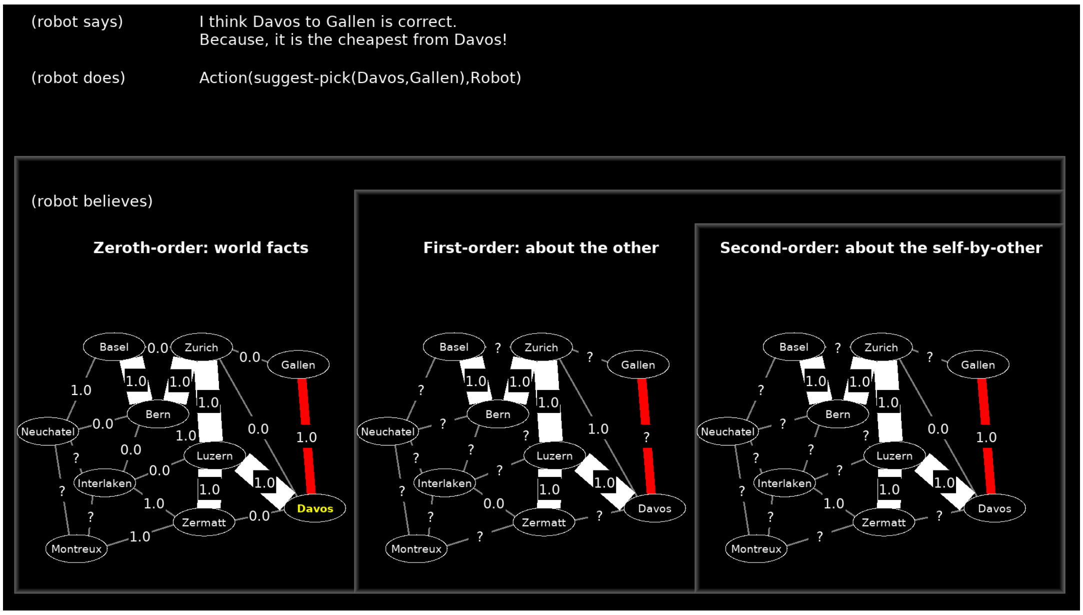
7.3 Mental Modeling for the Robot: An Example

We focus on a collaborative world variant of the JUSThink human-robot collaborative activity in Chapter 6. Inspired by finitely-nested I-POMDPs, we present a way for the robot to represent and maintain a mental model about its environment: the human and the world. For our research objective to explore and compare the effects of different mental models by the robot (in Chapter 8), we want to make the differences more evident: based on distinct, “black and white" beliefs. Thus, we use a special case of Bayesian update to maintain the beliefs, with the likelihood P(belief | observation history) being either 0 or 1. The reason that we use a continuous probability approach (Bayesian) but develop a binary model (true of false) is that it could evolve towards more continuous values of P. See Section 9.2.2 where we discuss about this and other shortcomings of this modeling framework. The zeroth-order beliefs are updated by the robot's reasoning on the world, e.g. a correct strategy—see Section 7.4. The first-order beliefs are updated as follows: whenever the robot witnesses the human picking a connection, by making a suggestion or agreeing with a suggestion, the probability that the robot thinks that the human thinks it is a correct choice is set to 1. Conversely, whenever the robot witnesses the human disagreeing with one of its suggestions, the probability that the robot thinks that the human thinks it is a correct choice is set to 0. The second-order beliefs are updated with the robot taking an action: its suggestions, and agreements set its own beliefs about the human's beliefs about itself to 1, and disagreements set it to 0. These update functions can be modified to accommodate any uncertainty the robot can have about its beliefs.

Figure 7.2 shows an example situation of the world, and a visualization of the mental model of the robot with level $l = 2$, that is maintained by an interaction with a human. The robot has zeroth-order beliefs, i.e. its own mental model about the world. Then, it has first-order beliefs, that constitute (its own) mental model of the human about the world. Finally, it has second-order beliefs, i.e. its own mental model of the human's model of the robot about the world. From the mental model visualized in this figure, Table 7.1 presents examples for the robot's beliefs in linguistic form, as well as logical forms to highlight the order of a belief.



(a) An example game scene in a human-robot JUSThink activity. It is the human's turn.



(b) The corresponding mental scene that visualizes the mental model of the robot with order $l = 2$, i.e. it has zeroth-, first-, and second-order beliefs. A probability assignment $P = 1.0$ on an edge at a belief order indicates the robot thinks that the edge is a correct choice for that order; $P = 0.0$ means the robot thinks it is not a correct choice, and a question mark (?) indicates the robot has not formed a belief about the choice. See Table 7.1 for example linguistic expressions that describe the beliefs.

Figure 7.2 – An example game and mental scene for a human-robot JUSThink activity

Table 7.1 – Example beliefs in the mental model of the robot in a human-robot JUSThink activity that illustrate the different orders of beliefs. They are from the mental state of the robot as visualized in Figure 7.2. B_r and B_h point to beliefs by robot and human, respectively. B denotes the belief operator, where $B_a\phi$ would read as “a thinks that ϕ.". Predicate $P(u,v)$ denotes “$u-v$ is selected.", and $Q(u,v)$ stands for “(Connecting) $u-v$ is a correct choice".

Belief (Linguistic/Verbal Form)	**Logical Form**	**About**	**Order**
“I think Basel-Bern is selected."	$B_r P(\text{Basel}, \text{Bern})$	world state	zeroth
“I think Basel-Neuchâtel is not selected."	$B_r \neg P(\text{Basel}, \text{Neuchâtel})$		
“I think Basel-Bern is a correct choice."	$B_r Q(\text{Basel}, \text{Bern})$	my strategy	
“I think Basel-Zurich is not a correct choice."	$B_r \neg Q(\text{Basel}, \text{Zurich})$		
“I think that the human thinks that Zermatt-Luzern is a correct choice."	$B_r B_h Q(\text{Zermatt}, \text{Luzern})$	your strategy	first
“I think that the human thinks that Zermatt-Interlaken is not a correct choice."	$B_r B_h \neg Q(\text{Zermatt}, \text{Interlaken})$		
“I think that the human thinks that I think that Davos-Luzern is a correct choice."	$B_r B_h B_r Q(\text{Davos}, \text{Luzern})$	my strategy by the other	second
“I think that the human thinks that I think that Davos-Zurich is not a correct choice."	$B_r B_h B_r \neg Q(\text{Davos}, \text{Zurich})$		

To illustrate how beliefs about choosing a connection are formed and can differ among the different orders of belief, consider Figure 7.2 and the connection Zermatt-Interlaken: the robot thinks it is a correct choice (as a zeroth-order belief with $P = 1$, left column), but at the same time, the robot thinks that the human does not think it is a correct choice (a first-order belief with $P = 0$, middle column). Then, there is a second-order belief: the robot thinks that the human thinks that the robot thinks it is a correct choice ($P = 1$, right column). These beliefs were maintained as follows: the zeroth-order belief on Zermatt-Interlaken being a correct choice was formed by the robot's strategy. Then, the robot suggested choosing this connection (according to this zeroth-order belief). By the robot's understanding, this gives the human a cue that the robot thinks it is a correct choice. Thus, the robot now also forms a second-order belief to reflect this: “I think that the human thinks that I think Zermatt-Interlaken is correct.". Finally, the human disagreed with this suggestion: this gives a cue for the robot to form the first-order belief “I think that the human does not think that Zermatt-Interlaken is correct."

7.4 Reasoning with the Mental Model

We consider three strategies the robot can follow as it works together with a human in a collaborative world in the JUSThink domain. We use these strategies as the action selection procedures for robot behaviors in Chapter 8, where we analyze the effects of using a mental model of the human maintained by the robot to drive the robot's behavior with first-order beliefs about the human; compared to having only zeroth-order beliefs about the activity only and using a correct vs. incorrect reasoning strategy, via User Study 3. Other strategies can be designed to use the first-order beliefs differently, or also use second-order beliefs; and the rules in a strategy can be learned on data.

7.4.1 A Correct Strategy (Jarník's Algorithm)

One of the strategies the robot can follow is a *correct strategy*, i.e. it leads to an optimal solution. We design one such strategy, where the robot decides on an optimal action by forming correct, explicit zeroth-order beliefs about choices it can make in the world, without explicitly representing the human, as described in POLICY-COLLABORATIVE-ZEROTH in Algorithm B.2. It correctly reasons with the situation of the world to find equivalence classes for intended and remaining choices by the Jarník's algorithm, and updates its beliefs to reflect these preferences, as described in INTENTION-CORRECT in Algorithm B.4 and BELIEF-UPDATE-ZEROTH in Algorithm B.3. The intentions are guaranteed to be part of an optimal solution by the *cut property* of the MST problem[XI], if the current solution is part of an optimal solution (hence *growing* an optimal solution): the robot does not consider taking a clear action to remove incorrect choices if any, but rather manifests a correct way to to reason about new choices.

We control the assertiveness of the robot with a *persistence criterion*: disagree threshold $T_{dis} = 2$ for the maximum number of times it disagrees for a specific connection. If infinity, the robot essentially demonstrates a correct solution to the human, by agreeing with the human only for the optimal choices it inferred, and disagreeing with any other suggestion. If finite, a possible deadlock is avoided, if the human were to suggest an incorrect connection repeatedly, where the robot would not accept because it does not think the choice is correct: this can happen only up to T_{dis} times. For the (sub-)optimal robot behavior in Chapter 8, we set $T_{dis} = 2$ to allow the robot disagree up to two times, thereby asserting its beliefs, and only agreeing afterwards. For the aligning robot, we set $T_{dis} = 1$: the design of the robot behaviors is described in Section 8.3.

A verbal explanation for this strategy when the robot suggests or agrees with a connection due to its own beliefs is "I think it is a correct choice, because it is the best from the ones that are connected.": this refers to an intuitive understanding about the cut property. When the robot agrees not because of its own beliefs but due to the persistence criterion, it says "Fine, since you insist so much! I still do not think it is a correct choice.". For examples of the explanations the robot can give to the human, see POLICY-COLLABORATIVE-ZEROTH in Algorithm B.2.

7.4.2 An Incorrect Strategy (A Locally Greedy Algorithm)

Another strategy the robot can use is an *incorrect strategy*, i.e. it will not lead to an optimal solution. We design such a strategy, where the robot decides on a possibly sub-optimal action by forming likely incorrect, explicit zeroth-beliefs about choices it can make in the world; similarly without explicitly representing the human. It incorrectly reasons with the situation to find equivalence classes for intended and remaining choices by a locally greedy algorithm, updates its beliefs to reflect these preferences, and selects an action according to its

[XI] A *cut* of a graph is a set of edges that if we remove them, the graph will be split into two components without any connection between them. The *cut property* of the MST problem shows that an edge in the cut is part of an optimal solution, if it has the lowest cost among the edges in the cut.

beliefs; as described in INTENTION-INCORRECT in Algorithm B.5, BELIEF-UPDATE-ZEROTH in Algorithm B.3, and POLICY-COLLABORATIVE-ZEROTH in Algorithm B.2, respectively. The belief update and action selection policy are the same as in the correct strategy; only the reasoning changes from a correct to an incorrect one. The intentions are no longer guaranteed to be part of an optimal solution: in this reasoning, the robot adopts a locally greedy approach, thinking the selection is to be made from a node, and tries to traverse greedily from that node. The node is moved as selections are made. If there is no selection available, it moves to another node that is selected, and tries to select the best (cheapest), local i.e. outgoing connection from that node. If the human would agree with all the robot's choices, this strategy will result in a sub-optimal solution with all the connections selected.

Verbal explanations for this strategy, when the robot suggests or agrees with a connection based on its own beliefs, come in two forms: *from*-explanations, and *to*-explanations. A from-explanation is of the form: "I think it is a correct choice, because it is the best from ⟨current node⟩.". A to-explanation has the form: "I think it is a correct choice, because it is better than going to ⟨destination nodes in the remaining choices⟩.", where the remaining choices are enumerated by the reasoning algorithm: by INTENTION-INCORRECT. Both type of explanations given an intuitive understanding about the robot's incorrect reasoning; with the from-explanations explicitly stating the node the robot thinks the selection needs to be made, while this is implicit in a to-explanation. When the robot agrees not because of its own beliefs but due to the persistence criterion, it will still say "Fine, since you insist so much! I still do not think it is a correct choice.", similar to the optimal strategy.

7.4.3 A Mutual Modeling Strategy (An Aligning Algorithm)

The third strategy we consider uses a representation of the human to build a mutual understanding on what a correct solution is, and converge to a shared solution that they agreed upon in order to solve the problem. For this purpose, the robot forms likely incorrect explicit zeroth-order beliefs about choices it can make and explicitly represents the human with first-order beliefs attributed to the human, about the choices the human makes in the world. We design the strategy such that the robot prefers choices in their common ground that consists of their aligned choices, or else if not available chooses a sub-optimal action, described in POLICY-ALIGNING in Algorithm B.6. It still incorrectly reasons with the situation to find equivalence classes for intended and remaining choices by a locally greedy algorithm as in the incorrect strategy, by INTENTION-INCORRECT in Algorithm B.5. However, in this case, the robot update uses first-order beliefs about the other as well as its zeroth-order beliefs in its reasoning, to refer to its first-order beliefs while revising its zeroth order beliefs; by using BELIEF-UPDATE-ALIGNING in Algorithm B.7, instead of BELIEF-UPDATE-ZEROTH in Algorithm B.3. This results in an adaptive-greedy strategy, where the responsibility to reach an optimal solution is still on the human as in the incorrect strategy.

Verbal explanations for this strategy depend on how the choice is made. The aligning strategy accommodates the shared choices into its belief system, by forming beliefs like “I think that ⟨choice⟩ is correct." for its own beliefs, and “I think that you think ⟨choice⟩ is correct." for the beliefs ascribed to the human. Thus, a verbal explanation for an aligned choice is in the form: “I think that *we both* think it is a correct choice.", as intentional explanations, due to a belief attribution. The first-order attributions and their maintenance brings in another aspect to the interaction: the robot can state its beliefs and belief updates with utterances such as: “Great! You agree! Then, we both think it is a correct choice.". When there is a contradiction, such as the human not agreeing with a previously-agreed joint decision the next time the same choice is considered, the robot can say “Oh really? I thought you did not think it was a correct choice!" to reveal its surprise, and revise its beliefs accordingly again. When the robot agrees not because of its own beliefs but due to the persistence criterion, it is now changing its own belief to align on the choice and add it to their common ground, e.g. to say: “I agree, I think that it is correct, although previously I did not think it was correct.". For examples of the explanations the robot can give to the human, see POLICY-ALIGNING in Algorithm B.6.

7.5 A Demonstration

Table 7.2 presents a demonstration of two artificial agents playing together in terms of what they think and say (that implies what they do), where one is following the optimal strategy and the other the mutual modelling strategy. We see that the modeling agent adapts to the optimal agent, and maintains first- and second-order beliefs about the optimal agent. On the other hand, the optimal agent does not form beliefs in these orders.

7.6 Conclusion

In this chapter, we presented a mental modeling framework to equip a robot with mutual understanding abilities, enabling the robot to build and maintain its own explicit mental model of its environment: with its beliefs about the world, the human, as well as the human’s beliefs about the robot. As an example domain, we extended and generalized the human-robot collaborative problem solving activity in Chapter 6, and designed three strategies the robot can adopt to select its actions, depending on how it reasons with its environment, and whether it represents the human explicitly or not. Based on these strategies, in the next chapter (Chapter 8), we will explore the effects of these mental modelling abilities on the interaction and the learning outcomes, by comparing the effects of robot behaviors that are guided by different mental models through an experiment.

Table 7.2 – A demonstration of the models' execution for two agents A and B, that presents the model updates and utterances of the agents. A follows the optimal strategy and B follows the mutual modeling strategy. The turn-taking starts with A, in the world shown in Figure 7.2a, with an initial state where no connection is selected.

Model Update of A	**Utterance of** A	**Model Update of** B	**Utterance of** B
(initial beliefs by strategy) "I think Bern-Basel or Bern-Zurich or Bern-Luzern is correct.", "I think Bern-Interlaken and Bern-Neuchâtel are not correct."	Let's start at Bern. Shall we go to Basel? (suggests Bern-Basel)	"I think that you think Bern-Basel is correct."	–
"I think Bern-Zurich or Bern-Luzern is correct.", "I think (remaining action choices, including Basel-Neuchâtel) are not correct."	–	"I think that you think that I think Bern-Basel is correct.", "I think we are at Basel.", "I think Basel-Neuchâtel is correct.", "I think Basel-Zurich is not correct."	I agree, we both believe it is a correct one. (agrees with the suggestion, hence Bern-Basel is connected)
–	–	"I think that you think that I think Basel-Neuchâtel is correct."	I believe Basel to Neuchâtel is a good choice. Because it is the cheapest from Basel. What do you think? (suggests Basel-Neuchâtel)
–	I don't think it is correct! I disagree. (disagrees with the suggestion)	"I think that you think Basel-Neuchâtel is not correct."	–
–	Rather than that, I think Bern to Zurich is a good one. Because, it is the cheapest from the ones we have connected. (suggests Bern-Zurich)	"I think that you think Bern-Zurich is correct.", "I think we are at Bern.", "I think Bern-Zurich or Bern-Luzern is correct.", "I think Bern-Interlaken and Bern-Neuchâtel are not correct."	–
"I think that Zurich-Luzern is correct.", "I think that (remaining action choices) are not correct."	–	"I think that you think that I think Bern-Zurich is correct.", "I think we are at Zurich", "I think that Zurich-Luzern is correct.", "I think that Zurich-(remaining) are not correct."	We both think it is a good connection: I agree. (agrees with the suggestion, hence Bern-Zurich is connected)
…	…	…	…

8 Effects of Mental Models of a Robot on Mutual Understanding

8.1 Introduction

In this chapter, we investigate the potential benefits of the mental modeling abilities we developed for the robot in Chapter 7 on the interaction and the learning outcomes. For this purpose, we use the JUSThink human-robot collaborative problem solving activity in Chapter 6. The learning objective is to improve the CT skills of children, through reasoning on a network optimization problem; the humanoid robot collaborates with a human, via suggesting problem-solving actions and (dis)agreeing with each other in turns. The human and the robot need to build a mutual understanding on what a correct solution is, and converge to a shared solution that they agree upon in order to solve the problem. For this activity, we design three robot behaviors that are guided by different mental models maintained by the robot: (i) the *optimal robot* with a correct model of the activity, and no model of its human counterpart, (ii) the *sub-optimal robot* with an incorrect model of the activity, and (iii) the *aligning robot* with a model of the human, together with an incorrect model of the activity.[I] The aligning robot adapts to the human, by preferring their agreed-upon choices, as they build a shared solution; while the other robots do not explicitly model nor adapt to the human. We compare the effects of the robot's different behaviors via a between-subjects experiment with 61 children in schools, in User Study 3.

Results show positive learning outcomes irrespective of the robot's behavior, in terms of finding more correct solutions in the post-test than the pre-test. The data provides not enough evidence for a difference in the learning outcomes between the robot behaviors at the first glance, despite the fact that the robots differ in their conception of the problem from incorrect to correct, and from no model of the human with persistent/non-adaptive choices to a model that adapts to the human. By considering the participants that have prior knowledge separately from those without any prior knowledge, we see that the sub-optimal robot is less effective in helping the participants without prior to learn, compared to the optimal or the

[I]The code that represents the activity and governs the interaction with the robot is publicly available online, from the GitHub Repositories https://github.com/utku-norman/justhink_world for the activity (that can also replay the data of User Study 3), and https://github.com/utku-norman/justhink-ros for the interaction.

Table 8.1 – Research questions and hypotheses for User Study 3

No.	Research Question / Hypothesis
RQ1	*How are the learning outcomes after collaborating with the robot? Do they differ by condition?*
H1.1	Participants learn for all robot behaviors, i.e. provide more correct solutions in the post-test than the pre-test.
H1.2	Participants that collaborate with the aligning robot learn more than those that collaborate with (a) the sub-optimal robot or (b) the optimal robot.
H1.3	Participants that collaborate with the optimal robot learn more than those that collaborate with the sub-optimal robot.
RQ2	*How does the evolution of performance in the task link to the learning outcomes? Does it differ between conditions?*
H2.1	The actions of the participants that have some prior knowledge are better during collaboration than those that have no prior knowledge.
H2.2	The actions of the participants that learn improve during collaboration, while the actions do not improve for those that do not learn.
H2.3	The actions of the participants that collaborate with the aligning or sub-optimal robot exhibit these distinctions in H2.1 and H2.2 more clearly, than those that collaborate with the optimal robot: i.e. namely, (a) high performance for those with some prior knowledge than no prior knowledge, (b) actions improve for those that learn than those that do not.

aligning robots. By looking into the actions during the collaboration and how the quality of the actions evolves, we see that the sub-optimal and the aligning robots can differentiate between those who learned and those who did not. The suggestions to and the agreements with the sub-optimal robot as well as the agreements with the aligning robot became worse later in the collaboration for those who did not learn, whereas we do not observe this pattern for those who learned. The sub-optimal and the aligning robots can make this distinction more clearly than the optimal robot. Although mutual modeling by the robot helps it form a belief about the human alongside its own beliefs about the world, it does not directly result in a single best way to use the model to improve learning. A future robot that leverages a correct conception of the problem together with its mental model about the human can act adaptively between the optimal and the sub-optimal robots, give right and wrong suggestions to gauge the human's understanding, and at the same time help realize misconceptions and steer the human towards learning.

8.1.1 Research Questions and Hypotheses

Table 8.1 presents our research questions and hypotheses for the experiment. In RQ1, we ask about the learning outcomes, and whether they differ by the robot's behavior. As in the pre-experiment in Chapter 6, we postulate that collaboration with the robot has a positive impact on the learning: the fact that the participants work together with the robot, and need to make an effort to converge on a shared solution, can help them realize their misconceptions if any. Thus, we hypothesize in H1.1 that this is the case, irrespective of the robot's specific strategy that drives its behavior: i.e. whether optimal, sub-optimal, or aligning. We already observed evidence for this in the pre-experiment, where the children interacted with the first version of the sub-optimal robot that does not model the human; there were positive learning outcomes, in terms of finding better solutions in the post-test than the pre-test, suggesting that the collaboration with the robot might have helped trigger the learning mechanisms.

In the pre-experiment in User Study 2, there was no control group so we could not claim a causal relationship. In this case for our experiment in User Study 3, we can now compare the effects of different robot behaviors, where each behavior serves as an experimental condition. In a sense, we have two control groups, as baselines: those who interact with the optimal or the sub-optimal robot. These robots do not model the human, but differ in their own conception of the problem, i.e. following a correct or an incorrect strategy. Then, the experimental group interacts with the aligning robot. Although the aligning robot still has an incorrect strategy, it builds and uses an explicit model of the human and assists their convergence to what can still be an incorrect solution: this can induce the human to model the robot as well, to reach a mutual understanding of each other about the task. The participant can then engagingly reason with why the robot acts in the particular way, as they still end up with an incorrect solution. In contrast, the sub-optimal robot only follows its own incorrect approach, and asserts its perspective in a less amenable way than the aligning robot. The sub-optimal robot may simply leave the participants confused: it does not know how to solve the problem, and does not make any particular effort to converge on a solution like the aligning robot does. Hence, we hypothesize in H1.2 that the participants who collaborate with the aligning robot learn more than those who work with the sub-optimal robot, i.e. the non-adaptive version of it, as well as the optimal robot, which essentially leads the participant to a correct solution without leaving much room to think about the decisions and make wrong ones. Moreover, we hypothesize in H1.3 that those interacting with the optimal robot learn more than those with the sub-optimal robot. An optimal robot is like a tutor in an exercise session, suggesting optimal choices, and interactively agreeing with the optimal choices and disagreeing with the sub-optimal ones by the human. Since the optimal robot essentially demonstrates a correct solution, it could be easy or even straightforward that this is picked up by the human, and is directly translated into positive learning outcomes.

In RQ2, we consider the relationship between how the participants collaborate, whether they learn or not in the end, and if this differs by robot behavior. We postulate that the participants' actions and performance in their collaboration with the robot initially reflect how much prior knowledge they begin with; and it can point to whether they learn or not, as measured by any change in their knowledge after the collaboration. Thus, in H2.1 we hypothesize that those with some prior knowledge perform better in the collaboration, than those without. As they collaborate with the robot, the participants can realize their misconceptions, and therefore improve their actions and solutions through time. This can happen more easily if they have some prior knowledge, and it could be possible to detect and predict the learning from whether or how much the participants improve their actions. In this manner, we hypothesize in H2.2 that those who learn improve their actions through the collaboration, meanwhile those who do not learn do not improve as such. These distinctions, of improving or not, lie within the actions taken by the human. The aligning or the sub-optimal robot do not know how to solve the problem, and hence can not give away a correct solution, whereas the optimal robot guides the participant to a correct solution without many possibilities to make mistakes without being immediately corrected by the robot. Therefore, we hypothesize in H2.3 that the distinctions in H2.1 for prior knowledge and H2.2 for learning are much more evident for the participants that interact with the aligning or the sub-optimal robot, than the optimal robot.

8.2 Activity Design

In Chapter 6, we described the design of the JUSThink human-robot activity and scenario, and its evaluation via a pre-experiment. For the experiment in this chapter, we adopt the same human-robot interaction scenario, where the robot orchestrates the sequence of activities described in Table 6.2. We target the same learning objective, i.e. after completing the task, a participant will be able to *correctly choose a subset of connections on a given network*, so that (i) all nodes are connected to each other by some path, and (ii) the total cost on these connections is minimized: we specifically focus on the learning goal on constructing a correct solution (LG2), see Section 6.2.1. We assess the learning the same say, by a pre-test and a post-test where the participant individually solves several problem instances, see Section 6.2.2.

We slightly revise the collaborative activities as follows. The transition function is modified so that the collaborators can only select and suggest connections from the already connected mines, or any connection if no mines are connected yet. This simplifies the action selection strategies for the robot, so that it can decide on an action at any solution state and generate a consistent explanation for its choice, see Section 7.4. As in the pre-experiment, if time permits, each participant works together with the robot on two problem instances of the same complexity, see Section 6.2. Differently, we now allow $N \leq 3$ submissions as *attempts*, instead of $N \leq 4$ in the pre-experiment. In principle, only the human would submit with the sub-optimal and the aligning robots, whereas the optimal robot submits when there is a valid solution that connects the nodes: optimal if the persistence criterion is not reached. The sub-optimal and the aligning robots would submit only if all the connections are selected, whereas the first version of the sub-optimal robot would submit if there were no connections available from the node it thought they were at: this often resulted in submissions that are due to connecting a cycle with no further available connections. The activity still provides minimal feedback: i.e. terminate if the solution is correct, otherwise try again up to N times.

For the scenario, we adopt a new *termination criterion* to have a better control on the total duration of the interaction, so that it will likely last less than an hour: if the total duration in collaboration $T > 20$ min, the experimenter stops the collaborative part at the next submission, without any signaling. This means the second collaborative activity is introduced only if the termination criterion is not met. Note that there is no minimum duration: the participants can take actions as quickly as they want, only bounded by the response times of the robot.

We define the experimenter's role in terms of a research protocol, as we target sessions with fully autonomous human-robot interaction. The only interventions given by an experimenter are: (i) at the beginning of a session, introduce the experimenter's name and role as a researcher, ask about prior experience with robots, and say "You will play a game with QTrobot, it will tell you everything you need to know and do."; (ii) upon meeting termination criterion during collaboration, say "We need to move on.", and press the keyboard shortcut to move to the next activity; and (ii) at the end of the interaction, say thanks, and conduct an interview to get an insight on how the participant perceived the activity and the interaction with the robot. In this chapter, we did not evaluate the interviews; we will analyze them in our future work.

8.3 Design of Robot Behaviors

We design three robot behaviors, based on the action selection strategies the robot can follow to work together with the human that are described in Section 7.4: (i) *optimal robot*, that follows a correct strategy by the Jarník's Algorithm, (ii) *sub-optimal robot*, that follows an incorrect strategy by a locally greedy algorithm, and (iii) *aligning robot*, that performs mutual modeling to align with the human. For all robot behaviors, we set the probability of the robot verbalizing a strategy explanation to $P_e = 0.5$; sampled every time the robot makes a suggestion or agrees with the human. For the sub-optimal and the aligning robots, we set probabilities for the types of the explanation, so that there is 50% chance to give a from-explanation (i.e. 25% chance overall), or else a to-explanation (i.e. with 25% chance).

8.3.1 Optimal Robot

The optimal robot has only its own zeroth-order beliefs about the world: to select its actions, it uses a correct strategy described in Section 7.4.1. We set the persistence criterion $T_{dis} = 2$, i.e. the robot disagrees with a suggestion for a connection up to two times.

8.3.2 Sub-Optimal Robot

The sub-optimal robot also has only zeroth-order beliefs, but it uses an incorrect strategy described in Section 7.4.2. We set the persistence criterion $T_{dis} = 2$, as in the optimal robot.

8.3.3 Aligning Robot

The aligning robot uses an aligning strategy as described in Section 7.4.3, based on its first-order beliefs about the human's beliefs about the world and its own zeroth-order beliefs about the world. It prefers choices in their common ground that consists of the choices that are aligned between the human and the robot; or else if not available it chooses an action that can be sub-optimal the same way the sub-optimal robot does. We set the persistence criterion $T_{dis} = 1$, so that it disagrees with a specific suggestion at most once: then, the robot would revise its mental model, so that the choice becomes part of their common ground, "I agree, then *we both think* it is correct, although previously I did not think it was correct.". The

Figure 8.1 – An example scene of JUSThink human-robot scenario from Study 3

probability of giving a strategy explanation is $P_e = 0.5$, as in the optimal and the sub-optimal robots. In addition, the probability of verbalizing an intentional attribution is set to $P_I = 1.0$: then, it always verbalizes e.g. "(I think that) *we both think* it is a correct choice.", and "What a surprise! I thought *you did not think* it was a correct choice!".

8.4 Materials

8.4.1 User Study 3: Human-Robot Experiment

Setup

The human and the robot sit across each other, and a touch screen is placed horizontally in front of them. They can see each other and the screen, as shown in Figure 8.1. Each session took about one hour, and followed the outline given in Table 6.2. Ethical approval was granted by the EPFL HREC, No. 030-2021/06.04.2021 (that also authorized User Study 2 in Chapter 6).

Participants

We collected a dataset of 61 children (30 females and 31 males), aged 9–16 years old ($Mdn = 11$), from six schools in Switzerland, in a between-subjects study with three conditions: the children interacted with a humanoid robot (QTrobot), where the robot's behavior was one of the three behaviors described in Section 8.3. The study globally accounts for about 50 hours of interaction, of which around 20 hours spent in the collaborative problem-solving activities between the human and the robot. Assignment to a condition was by rotation with allocation

Table 8.2 – Descriptive statistics for the JUSThink Experiment Dataset ($N = 61$)

Variable	Condition	Mean	SD	Min	Max
total duration (min)	optimal	31.4	6.2	19.6	39.9
	sub-optimal	41.7	9.7	20.8	60.7
	aligning	44.6	10.6	20.7	58.9
time in collaborative activities (min)	optimal	6.2	2.5	4.0	12.1
	sub-optimal	17.6	9.0	3.8	36.6
	aligning	18.6	8.6	4.3	28.7
total number of submissions	optimal	2.1	0.2	2	2
	sub-optimal	2.5	1.1	1	3
	aligning	2.5	0.9	1	3
number of suggestions by the human	optimal	11.9	2.7	8	20
	sub-optimal	22.7	9.0	8	38
	aligning	20.0	7.1	8	32
number of agreements by the human	optimal	10.6	1.9	8	16
	sub-optimal	19.1	6.8	8	30
	aligning	15.5	5.4	9	31
number of disagreements by the human	optimal	1.8	2.1	0	8
	sub-optimal	5.1	4.5	0	15
	aligning	6.0	3.6	0	13

concealment. This allocation resulted in 20, 21 and 20 children that interacted with the optimal, the sub-optimal and the aligning robot, respectively.

8.4.2 The JUSThink Human-Robot Experiment Dataset

The *JUSThink Experiment Dataset* contains the interaction logs of $N = 61$ children. It consists of the history of state transitions for each participant regarding how they constructed their solutions in each activity: individually in the tests and together with the robot in the collaborative activities. The dataset is in the same form as JUSThink Pre-experiment Dataset described in Section 6.3.2: it primarily encodes $< s_n, a_n, s_{n+1} >$ triples, where s_n and a_n denote the state of the activity and the action taken at the time step n, that results in the next state s_{n+1}.

Table 8.2 presents descriptive statistics for this dataset. It reveals large differences of the learning time with the optimal robot compared to the other robots. This is a variable we ideally target to keep constant. Yet, in our activity, once an optimal solution is found and submitted, the task's goal is achieved; there is no further steps that can be taken about that problem. In an effort to balance the difference, the second collaborative activity is introduced if the termination criterion is not satisfied; thus, the median number of total submissions over the two collaborative activities is two, i.e. every participant builds about two solutions.

8.5 Methods

8.5.1 Measuring Prior Knowledge and Learning Outcomes

We quantify the overall performance of a participant in a test with a *test score*, computed as the fraction of optimal solutions in the test: from 0% (none of the solutions of the participant are optimal and hence correct) to 100% (all solutions are correct). This measure was also adopted in the pre-experiment, see Section 6.4.1. Regarding their $\mathsf{prior\ knowledge}$ about the task, the participants are labeled as having *no prior knowledge* (L) if their pre-test score = 0%, and *some prior knowledge* (H) otherwise: with a test score > 0%, the participant found at least one optimal solution in a test item. This discretizes the prior knowledge as being zero or above. If the participant has a wrong conception of the problem, we expect the participant to select redundant connections or more expensive connections in all items of the test. All pre-test items they are composed of 12 edges and seven nodes with exactly one optimal solution: the participant need to select six specific edges out of the 12 edges, see Section 6.2.2 for the design of the tests.

We measure the learning by the change in the quality of responses in the post-test, compared to pre-test: on the basis of learn, i.e. the relative learning gain of a participant, defined as the difference between the post-test and the pre-test scores normalized by the margin of improvement or decline, see Section 3.4.2 for the formula. The participants are labeled as having *learned* (G) if $\mathsf{learn} > 0$, as *did not learn* (NG) if $\mathsf{learn} = 0$ and pre-test score $\neq$ 100%, and as already an *expert* (E) in the task otherwise, with the pre-test and post-test scores = 100%.

8.5.2 Characterizing Actions in Terms of Optimality

Each action in this task, i.e. selecting a connection or submitting in a test, and suggesting a connection or (dis)agreeing with a suggestion in a collaborative activity, is qualified as *optimal* or not (*sub-optimal*), in terms of whether it is leading to an optimal solution or not. We assign these quality labels to every action the same way as in the pre-experiment, see Section 6.4.3.

8.5.3 Quantifying the Trend of Change of Action Quality

To inspect how the quality of the actions changes through the collaboration, we label the first half of the actions by count as the *earlier* actions, and the second half as the *later* actions. Then, we look into how the fraction of optimal actions for each type of actions (i.e. suggestion, agreement or disagreement) changes from the earlier to the later actions of that type.

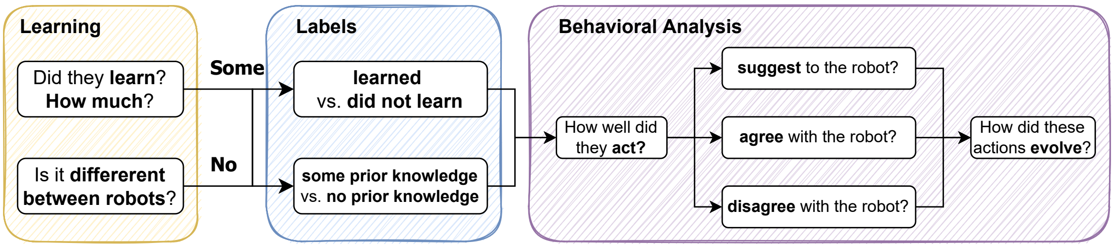


Figure 8.2 – Analysis flowchart for the JUSThink Experiment Dataset

8.6 Results and Discussion

We analyze the JUSThink Experiment Dataset and revisit our research questions and hypotheses, following the steps illustrated by the flowchart in Figure 8.2.

8.6.1 Preliminary Checks

Between conditions, there is no statistically significant difference in the pre-test scores of the participants (by a Kruskal-Wallis H test, $H(20,21,20) = 1.42, p = .49$) or the age of the participants ($H(20,21,20) = 0.01, p = .99$). The gender ratio (female to male) was 60%, 62%, and 25% for those who interacted with the optimal, the sub-optimal and the aligning robot, respectively: note that we did not specifically control for the gender, but assigned the participants regardless of their gender to conditions by rotation, to primarily keep the number balance overall and across schools that had differing age groups. Thus, across schools, there is no statistically significant difference between number of participants from a school per condition ($H(20,21,20) = 0.39, p = .82$).

8.6.2 RQ1 on the Learning Outcomes

Figure 8.3 shows how the participants' performance in the tests changed after collaborating with the robot, from the pre-test to the post-test. We observe positive learning outcomes, in terms of a higher median for the number of correct solutions after collaboration: a Wilcoxon signed-rank test shows that the post-test scores are statistically significantly different than the pre-test ($W(61) = 103.5$, $p < .00001$[II]), and higher with a small effect size (Cliff's $\delta = 0.31$). This is the case for the participants in each condition, who interacted with a different robot as characterized by the robot's behavior: for the optimal robot with a medium effect size ($W(20) = 12.5, p = .002, \delta = 0.44$), the sub-optimal robot with a small effect size ($W(21) = 14.0, p = .02, \delta = 0.17$), and the aligning robot with a medium effect size ($W(20) = 10.0, p = .004, \delta = 0.35$). These outcomes support our hypothesis H1.1: the participants performed better after collaborating with the robot, by submitting more correct solutions, irrespective of

[II] Estimated by normal approximation for the p-value with the presence of ties ($N_{ties} = 27$); via dropping their ranks, that is typically applicable for sample sizes > 50 (Conover, 1999). The pre- and post-test score distributions are not normal (Shapiro-Wilk's $W = 0.72$ and $W = 0.82$ with $p < .00001$ for pre-test and post-test, respectively).

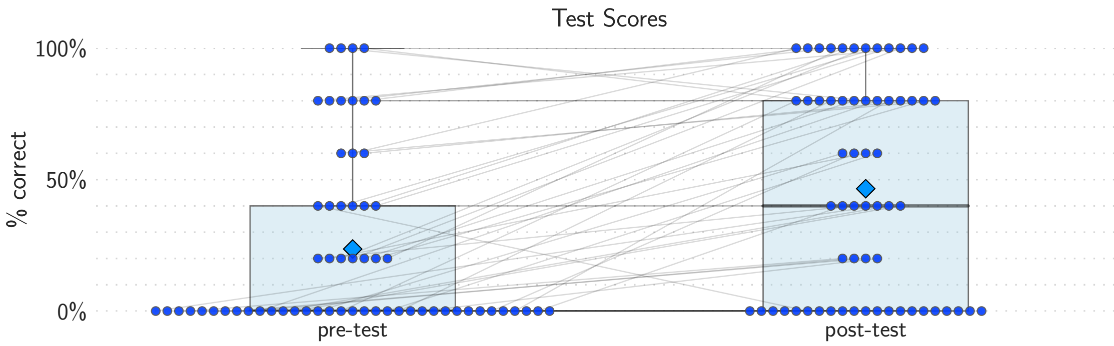


Figure 8.3 – Distribution of pre- and post-test scores in JUSThink Experiment Dataset ($N = 61$). A participant is indicated with two circles linked by a line, the mean is marked by a diamond.

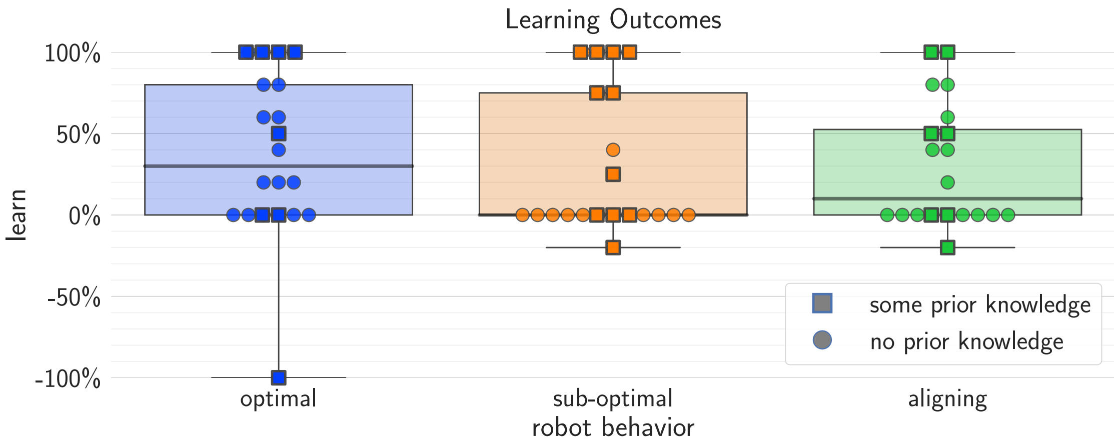


Figure 8.4 – Learning outcomes as measured by learn per condition in the JUSThink Experiment Dataset ($N = 61$). learn $= -100\%$ means the post-test score $= 0\%$, i.e. the participant's score from the pre-test declined as much as it could: e.g. the single participant with learn $= -100\%$ had 40% in the pre-test (correctly answering 2/5 questions) and 0% in the posttest (i.e. 0/5); see Figure 8.5 for the participants in the prior knowledge vs. learn space.

the condition.

Figure 8.4 shows the distribution of the relative learning gain (learn) of the participants in each condition. Between robot behaviors, a Kruskal-Wallis H test shows no evidence for a difference in the learning ($H(20,21,20) = 1.45$, $p = .48$). Thus, the data directly supports neither H1.2 on the aligning robot being more effective than (a) the optimal or (b) the sub-optimal robot, nor H1.3 on the optimal robot being more effective than the sub-optimal one. To delve deeper, we consider the participants that demonstrated having some prior knowledge, and those who exhibited no prior knowledge separately as two groups. Figure 8.5 shows the participants in the knowledge space as spanned by their prior and learning. There are 35 children without prior, 15 of which have learned and 20 did not; and 26 children with some prior, 16 of which learned, eight did not, and two are experts with perfect scores in the tests that we treat separately.

Figure 8.6 shows for each prior knowledge group the learning outcomes per condition. For

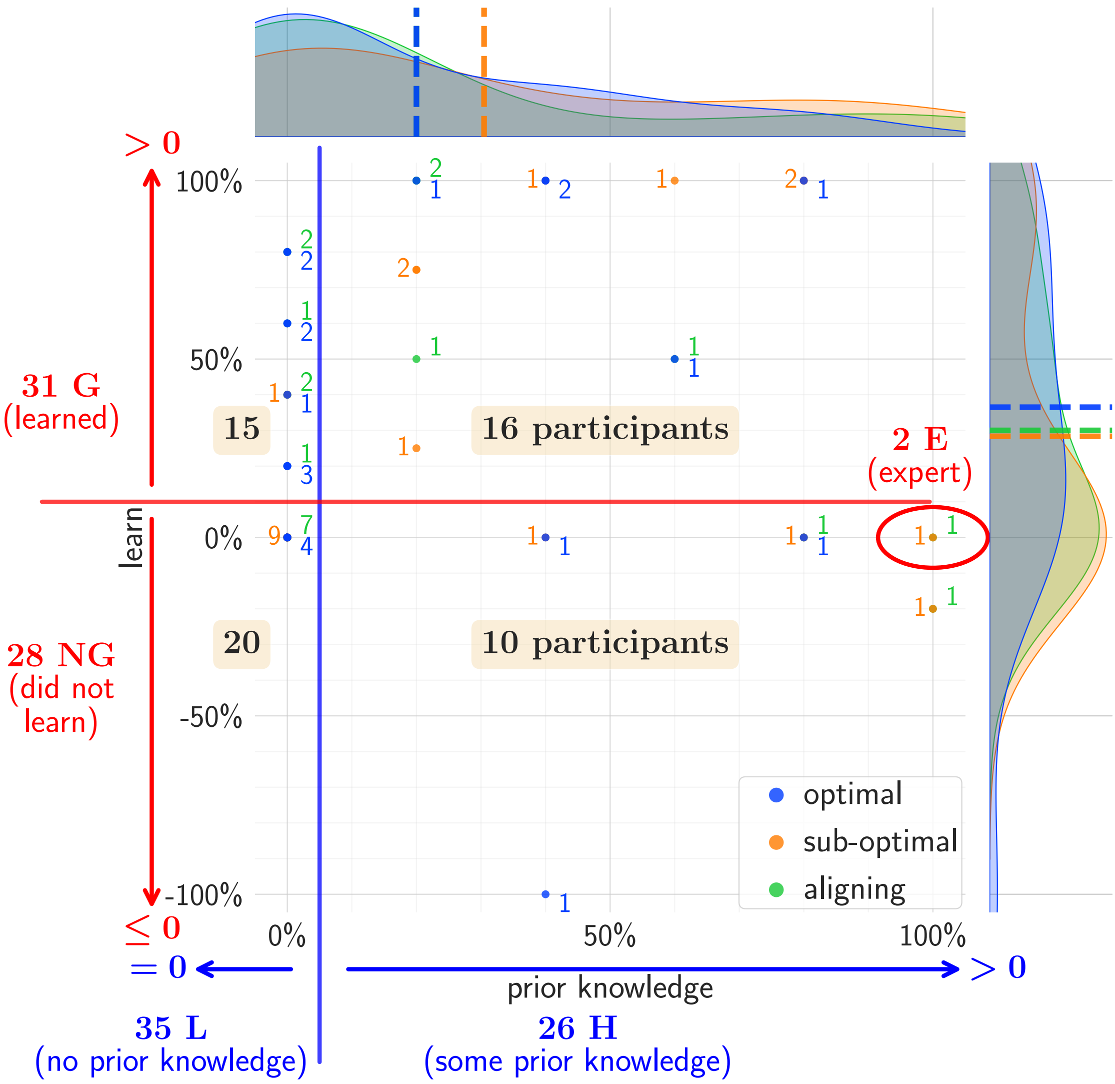
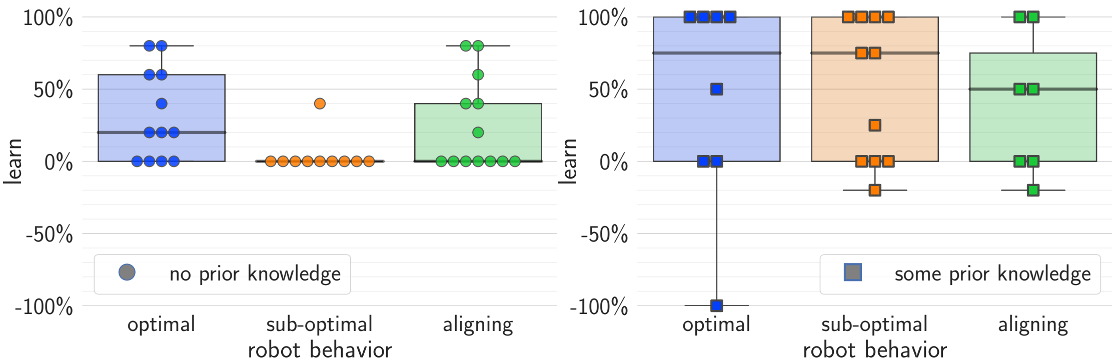


Figure 8.5 – Prior knowledge vs. learning plot for the JUSThink Experiment Dataset ($N = 61$) for each condition (i.e. robot behavior). The colored numbers in triangular formation indicate the number of participants in a condition for the indicated score tuple, zeroes are not shown; see the legend for the color coding for the conditions. The mean values for each condition are shown as dashed lines, with the fit of a univariate kernel density estimate. We annotate in blue our grouping by prior knowledge (i.e. participants with prior knowledge ≥ 0 are labeled as having some prior knowledge; and prior $= 0$ as not having any prior knowledge: pre-test score is 0%), and in red our grouping by learning (e.g. learn ≥ 0 as for those that learned, i.e. had positive learning gain by scoring better in the post-test than the pre-test). The total number of participants for each quadrant after this grouping is given with sand color background.



(a) participants without prior knowledge ($N_L = 35$) (b) with some prior knowledge ($N_H = 26$)

Figure 8.6 – Learning outcomes as measured by learn per condition for each prior knowledge group

participants without prior, the learning gains differ between conditions ($H(12,10,13) = 6.40$, $p = .04$), while there is no evidence for a difference for participants with prior ($H(8,11,7) = 0.35$, $p = .84$). For participants without prior, learning is higher with the optimal robot than the sub-optimal robot, with a large effect size ($H(12,10) = 6.5$ with $p = .01$, $\delta = -0.58$): this supports H1.3, partially for participants without prior. Indeed, inspecting Figure 8.6a on participants without prior, we see that only one participant learned with the sub-optimal robot (out of 10, i.e. 10%), whereas with the optimal robot most such participants learned (8 of 12, i.e. 67%).

Participants without prior tended to learn better with the aligning robot than the sub-optimal robot with a medium effect size, which is not statistically significant ($H(10,13) = 6.5$, $p = .06$, $\delta = 0.38$). Meanwhile, there is not enough evidence for a difference between aligning and optimal for those with prior ($H(12,13) = 0.51$, $p = .48$, $\delta = -0.16$) and without ($H(8,7) = 0.30$, $p = .58$, $\delta = -0.16$). Hence, there is only inconclusive support for H1.2(a), for only participants with prior learning more with aligning than sub-optimal. H1.2(b) that foresees the aligning robot to be more effective than the optimal is not supported. The robot behavior by itself is not sufficient to determine if the learning happens or not. Therefore, we look into the actions during the collaboration in RQ2 to see what they can tell us about the learning processes.

8.6.3 RQ2 on the Link Between Performance and Learning

Overall performance in the task The duration of the *task*, i.e. the collaborative activities, is bounded by the maximum number of attempts per activity ($N \leq 3$), and by the termination criterion (i.e. ending the task with the next submission if $T > 20$ min). Figure 8.7 shows how the participants performed in terms of success by finding an optimal solution, for each robot behavior and prior knowledge group. In the first activity, the participants with prior knowledge are more successful than those without any prior knowledge with a small effect size ($H(25,35) = 6.69$, $p = .01$, $\delta = -0.33$). Meanwhile, there is not enough evidence for a difference between these prior groups in the second activity ($H(22,22) = 1.08$, $p = .30$, $\delta = -0.09$). The

(a) first activity per condition over attempts

(b) second activity per condition over attempts

(c) first activity per condition by prior

(d) second activity per condition by prior

(e) first activity per prior knowledge group

(f) second activity per prior knowledge group

Figure 8.7 – Performance in the collaborative activities in the JUSThink Experiment Dataset

reason could be that they reached a similar level of competence regardless of their prior; as most of the participants who arrived to the second activity were successful in the first (39/44).

The success rate in the first activity is 61% (39/61): 100% (20/20) with the optimal, 50% (10/20) with the sub-optimal, and 43% (9/21) with the aligning robot. Due to the termination criterion, only 65% (44/61) of the participants could work on the second activity. Nonetheless, these participants performed well: the success rate in the second activity is 91% (40/44) overall; 100% (20/20) with the optimal, and 83% (10/12) each with the sub-optimal and the aligning robots. Success with the optimal robot follows by the design of its behavior, as it has a correct conception of the problem and leads the child to a correct solution; this makes reaching an optimal solution in the first attempt very likely, unless e.g. the human enforces selecting a sub-optimal connection and the robot agrees by the persistence criterion only after several tries, or the human submits prematurely before they build an optimal solution. Only one participant (out of 20) that interacted with the optimal robot went beyond the first attempt, and it is for the second activity, as seen in Figure 8.7b. In contrast, with the sub-optimal and the aligning robots, it is on the human to figure out an optimal solution, or at least to recognize and submit the solution if they construct one. For our comparisons in the following analyses, we focus on the actions in the first activity, as it has been experienced by all participants.

Overall quality of actions: H2.1 and H2.3(a) Figure 8.8 shows the distribution of the quality of the participants' actions, as the fraction of optimal actions taken by the participants for each action type (suggest, agree, and disagree) in the first collaborative activity: it compares the participants who learned with those who did not learn. Figure 8.9 shows the same type of comparison, in this case between the participants with prior and those without prior.

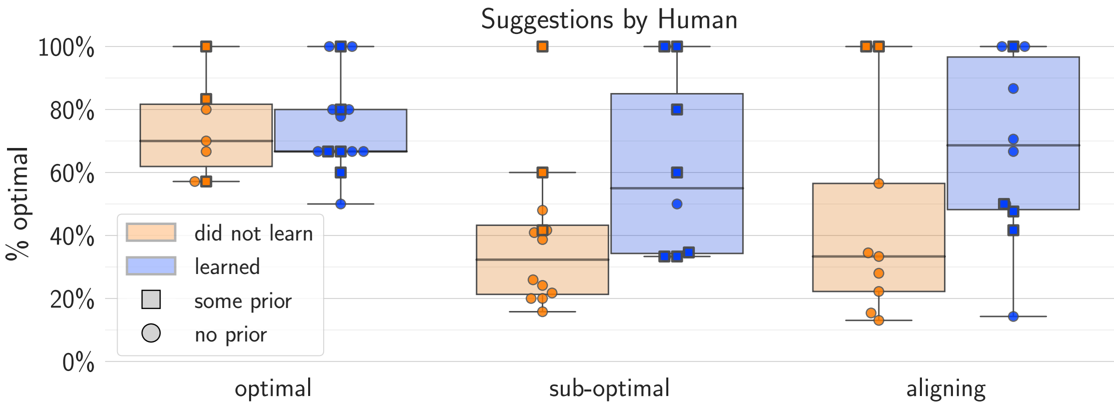
About suggestions, for the sub-optimal and the aligning robots, the participants who learned had a tendency to make better suggestions than those who did not learn, with large and medium effect sizes; yet these are not statistically significant ($H(8,12) = 3.59, p = .06, \delta = 0.51$ and $H(8,12) = 2.45, p = .12, \delta = 0.42$, respectively), see Figure 8.8a. There is a difference for the sub-optimal robot when we compare the participants without any prior and those with some prior: those with prior made better suggestions than those without prior ($H(10,10) = 7.45, p = .006$), with a large effect size ($\delta = 0.72$). For the aligning robot, those with prior also had a tendency to make better suggestions; though insignificant ($H(13,6) = 3.56, p = .11$, medium $\delta = 0.46$), see Figure 8.9a. Meanwhile, there is no evidence for a difference with optimal robot.

For agreements, we see a pattern similar to the suggestions: those who learned tended to agree better ($H(8,12) = 3.18, p = .08$ with large $\delta = 0.48$ for the sub-optimal robot, and $H(10,9) = 2.01, p = .15$ with medium $\delta = 0.39$ for the aligning robot). Furthermore, those with prior do also agree better as shown in Figure 8.9b ($H(10,10) = 6.67, p = .01$ with large $\delta = 0.68$ for the sub-optimal robot, and $H(13,6) = 4.37, p = .07$ with large $\delta = 0.53$ for the aligning robot). We see no difference for the optimal robot in terms of the quality of agreements: by the robot's behavior, as it only makes optimal suggestions, it is not possible to agree with it sub-optimally.

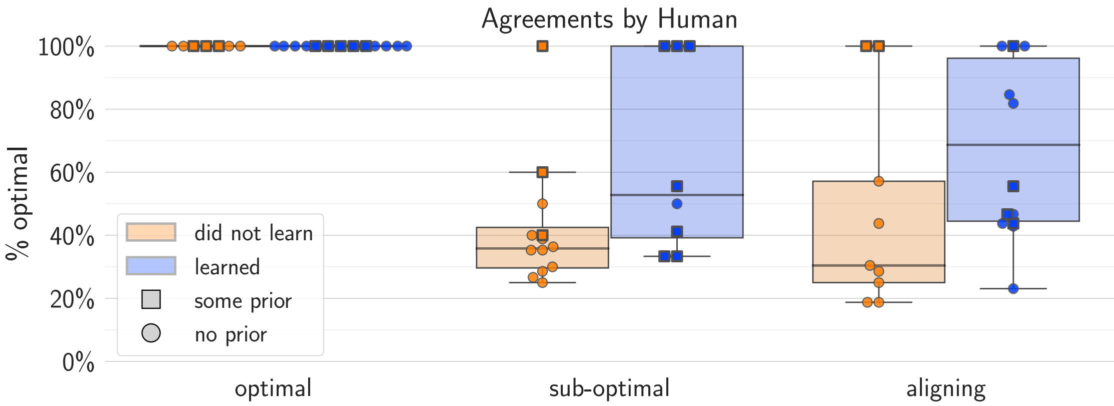
The pre-experiment showed that children had no problem disagreeing with a humanoid robot: then, the robot's strategy can have a genuine effect in surfacing the beliefs of children. We see that this is the case here too; the median of the number of disagreements is one even with the optimal robot, and higher with the other robots; see Figure 8.10 for the action counts per attempt. Regarding disagreements, we see that those who did not learn slightly tended to disagree better (i.e. correctly disagreed with the sub-optimal suggestions of the robot) than those who learned, with small effect sizes ($H(9,6) = 1.03, p = .31, \delta = -0.32$ for the sub-optimal, and $H(7,9) = 0.36, p = .55, \delta = -0.18$ for the aligning). There is no evidence for a difference if we split by the prior, as shown in Figure 8.9c. For the optimal robot, as it only suggests optimal choices that are part a correct solution, it is not possible to disagree with it optimally: any disagreement would be sub-optimal. Hence, we do not see any difference for the participants that worked with the optimal robot, for either splits about disagreements. Therefore, H2.1 is partially supported for suggestions and agreements; the participants with prior tend to act better than those without prior. This becomes especially evident for the sub-optimal and the aligning robot, compared to the optimal; which supports H2.3(a). With the optimal robot, it is not possible to differentiate between those who learned vs. who did not learn, or those with prior vs. without. The sub-optimal suggestions of the human are too quickly discarded before the human grasps why they are wrong, and the optimal suggestions are agreed with too easily without any challenge for the human to reason about why they are actually correct: the robot does not present the necessary opportunities for reflection.

Evolution of the quality of actions: H2.2 and H2.3(b) Figure 8.11 shows how the quality of actions evolved for the participants that worked with the sub-optimal and the aligning robots, by comparing the quality of earlier actions (in the first half of the actions) with the later actions (in the second half), for those who learned vs. did not learn. With the sub-optimal robot, the participants who did not learn made worse suggestions later than earlier, with a large effect size (by Wilcoxon signed-rank test $W(12) = 12.0$, $p = .04$, Cliff's $\delta = -0.49$). In contrast, those who learned had a tendency to improve their suggestions, with a small effect size; which is not statistically significant ($W(8) = 8.0$, $p = .23$, $\delta = 0.31$). We do not observe this distinction with the aligning robot, where both those who learned and did not learn tended to slightly improve with small effect sizes, and these are not statistically significant ($W(10) = 17.0$, $p = .44$, $\delta = 0.24$, and $W(9) = 13.0$, $p = .34$, $\delta = 0.15$, respectively).

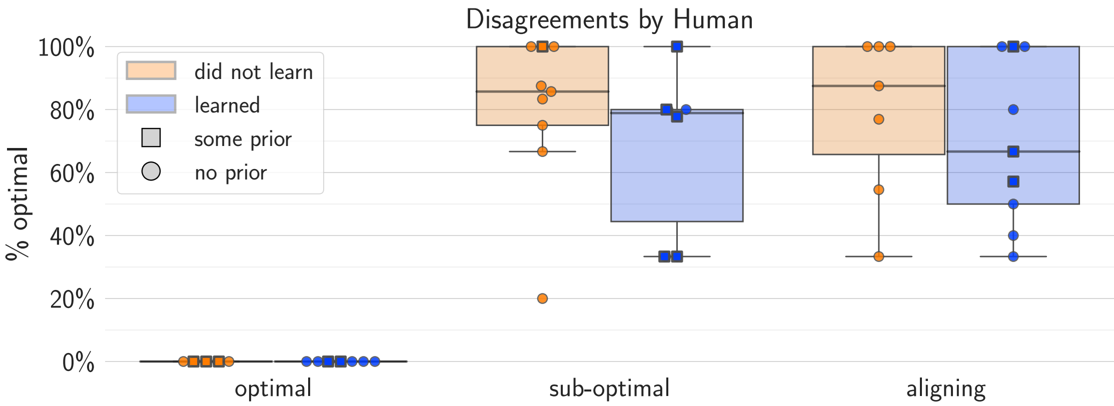
For the agreements by the participants, we see a pattern similar to the suggestions: those who did not learn agreed worse later than earlier with a large effect size ($W(12) = 12.0$, $p = .04$, $\delta = -0.58$), whereas those who learned rather tended to slightly improve in their agreements ($W(8) = 10.0$, $p = .47$, $\delta = 0.20$). We only see negligible-small effect sizes for the aligning robot, in the same directions ($W(9) = 18.0$, $p = .72$, $\delta = -0.11$ for did not learn, and $W(10) = 23.0$, $p = .88$, $\delta = 0.19$ for those who learned). Thus, the data supports H2.2 especially for suggestions to and agreements with the sub-optimal robot: those who learned improve during collaboration, while the actions do not improve as such for those who did not learn. It seems that when the aligning robot builds a common ground that consists of sub-optimal choices, they would still



(a) suggestions that participants made to the robot for those who did not learn vs. learned

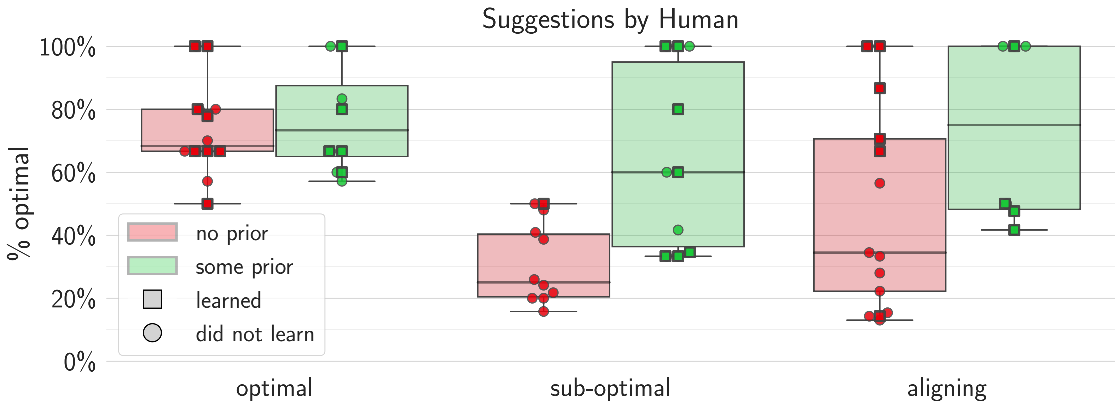


(b) participants' agreements with the robot for those who did not learn vs. learned

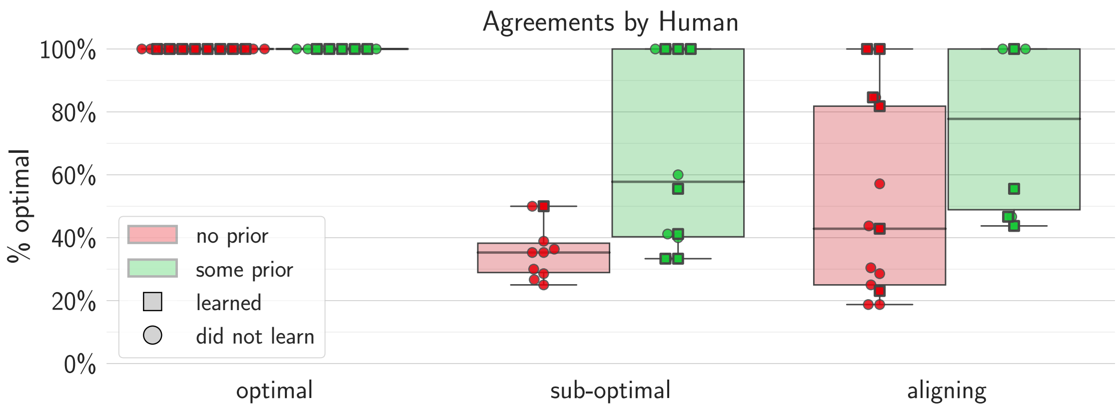


(c) participants' disagreements with the robot for those who did not learn vs. learned

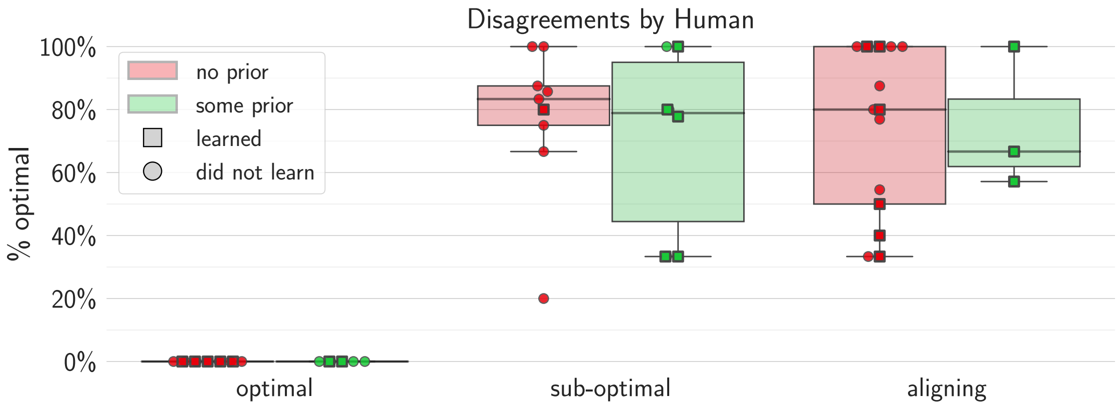
Figure 8.8 – Quality of actions by learning group in the JUSThink Experiment Dataset



(a) suggestions that participants made to the robot for those without prior vs. with some prior

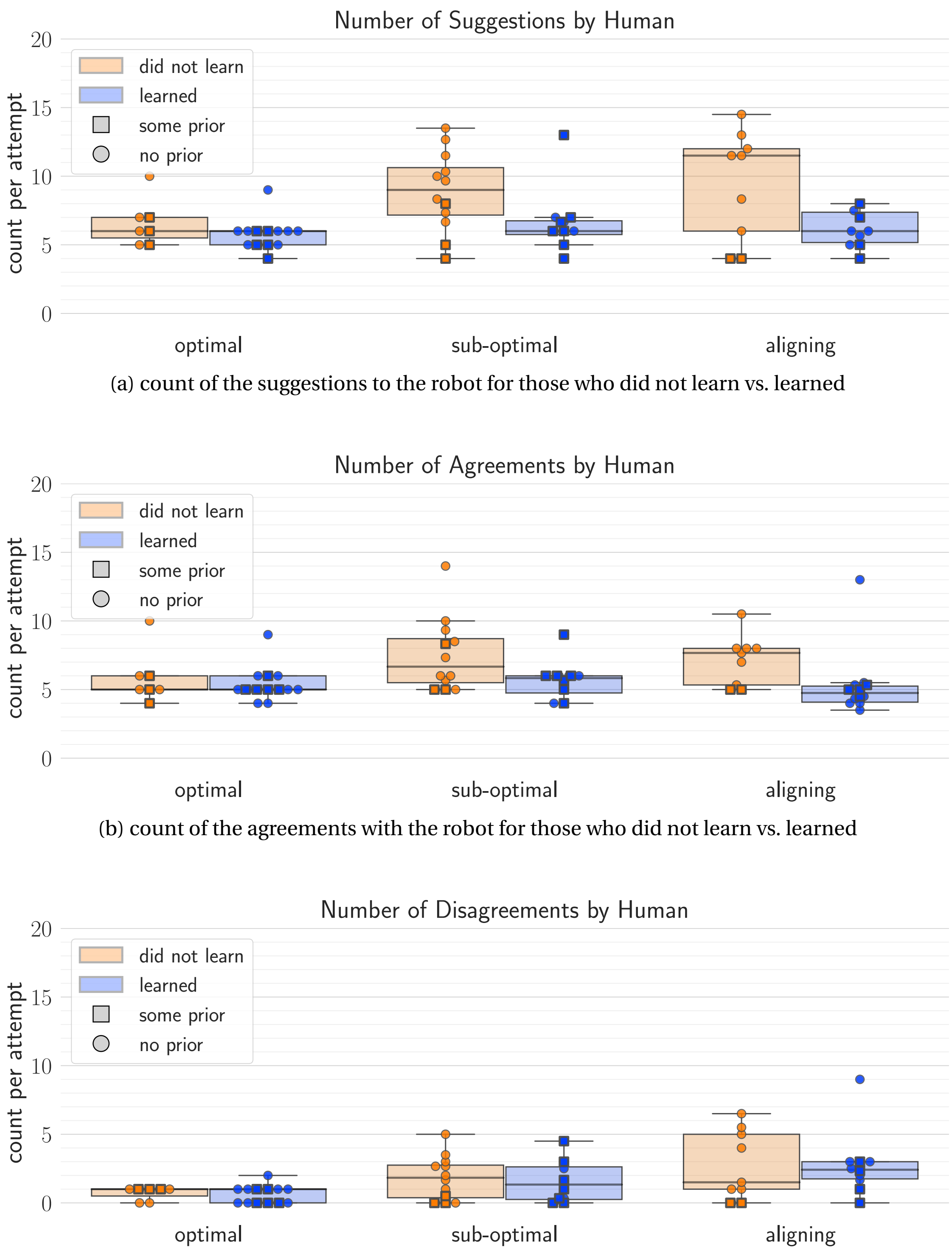


(b) participants' agreements with the robot for those without prior vs. with some prior



(c) the participants' disagreements with the robot for those without prior vs. with some prior

Figure 8.9 – Quality of actions by prior knowledge in the JUSThink Experiment Dataset



(a) count of the suggestions to the robot for those who did not learn vs. learned

(b) count of the agreements with the robot for those who did not learn vs. learned

(c) count of the disagreements with the robot for those who did not learn vs. learned

Figure 8.10 – Number of actions per attempt in the JUSThink Experiment Dataset

be agreed e.g. again in a later attempt; yet, these choices are still sub-optimal. Then, as these choices are chosen without much challenge (after being established as agreed-upon choices), objectively classifying actions in terms of their optimality is able to discriminate less about the prior or the learning. This does not happen for the sub-optimal robot, as any choice of the robot or the human is likely challenged, and how the participants respond to these challenges can reflect their prior, and bring about the conceptual change that leads to the learning.

Figure 8.12 shows the evolution of the quality of actions evolved for the participants that interacted with the optimal robot. Contrary to the sub-optimal and the aligning robots, the suggestions of those who learned actually got worse later compared to earlier with a medium effect size ($W(13) = 9.0$, $p = .88$, $\delta = -0.43$). Meanwhile, those who did not learn had a tendency to do worse later, which is not significant ($W(7) = 7.0$, $p = .27$, $\delta = -0.33$). These support H2.3(b), that the collaboration with the optimal robot presents less distinction between those who learned and those who did not, than the other robots. For those who learned with the optimal robot, there could be some sub-optimal choices that the participants tried to select, that got rejected by the robot and lead them to recognize the misconceptions. For instance, if the participant tries to make a cycle by connecting two nodes that are already connected by some other path, the optimal robot would disagree as this is a sub-optimal choice. Such a disagreement can be helpful, only if the human actually tries to make such a connection. In contrast, with the sub-optimal or the aligning robot, the robot can actually agree with these kind of suggestions: this would happen rather frequently, whenever it is compatible with the locally greedy strategy.

8.7 Conclusion

In this chapter, we explored the effects of the mental modeling abilities that we developed for the robot in Chapter 7, by comparing three robot behaviors that are guided by different mental models maintained by the robot. We utilized the JUSThink human-robot collaborative problem solving activity from Chapter 6 as the test bench, in which the humanoid robot collaborates with a human to construct a shared solution. The *optimal robot* and the *sub-optimal robot* do not model the human, but have a correct and incorrect model of only the activity; whereas the *aligning robot* explicitly models and adapts to the human, by preferring their agreed-upon choices as they build a shared solution. By conducting an experiment, we investigated the effects of the robot's behavior on the interaction and the learning outcomes.

The results highlight that the children tended to learn, irrespective of the robot's behavior. At first glance, there is not enough evidence for a difference in learning outcomes between the robot behaviors, although the robots are rather different: on one axis, their conception of the problem ranges from incorrect to correct with the sub-optimal to the optimal robot. On another axis, their adaptivity and cognitive capability ranges from no-adaptation and no model of the human, to adaptive with a model of the human, with the (sub-)optimal to the aligning robot. By grouping the participants by their prior knowledge, we see that the

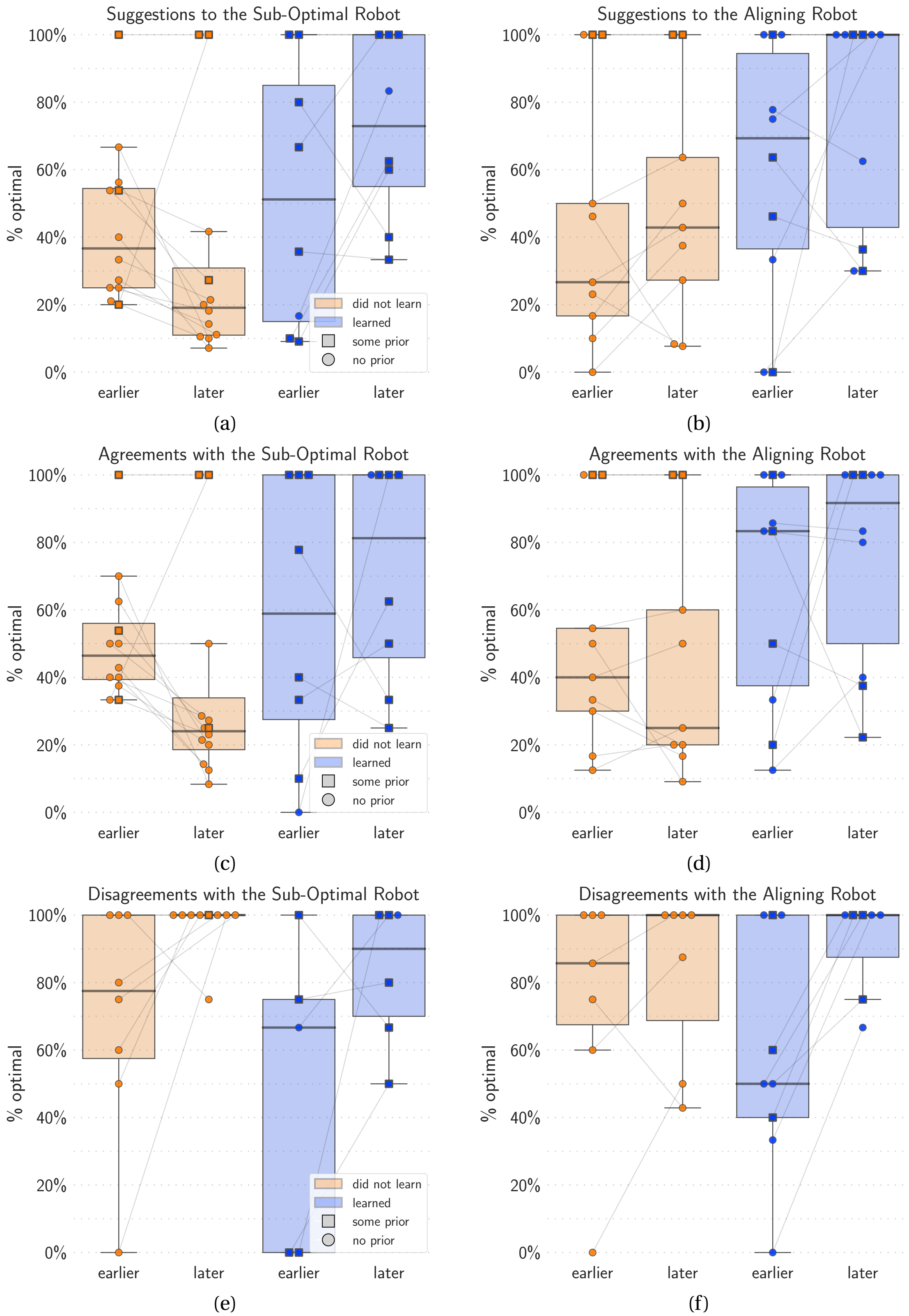
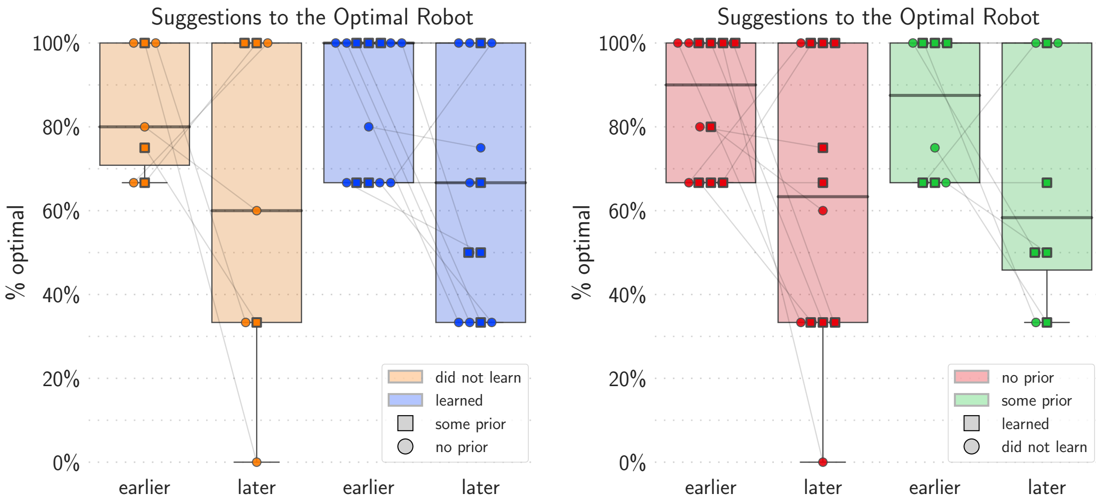


Figure 8.11 – Change in the quality of participant actions for sub-optimal and aligning robots



(a) suggestions for learning groups

(b) suggestions for prior groups

Figure 8.12 – Change in the quality of participant actions for the optimal robot

sub-optimal robot is less effective in helping the participants without prior to learn, compared to the optimal or the aligning robots. Yet, overall, the robot behavior by itself is not sufficient to determine if the learning would happen or not. In this thesis, we did not evaluate the interviews; we note that they could give interesting insights about the children's perception of the robots and their behavior, that we plan to analyze in our future work.

By inspecting how the quality of the actions evolves, we saw that the sub-optimal and the aligning robots can tell, later on in the activity, from the way the human interacts with them, whether they are learning or not, while the optimal robot can not. This is very important because it enables those two robots to take actions on the basis of whether the human is learning or not, e.g. changing strategy. Conversely, it is very interesting that the optimal robot is the one with the best learning for students without prior, as it seems it is effective in overcoming their initial “ignorance" and giving them opportunities to reflect on its own actions and understand the underlying principles. The mutual modeling abilities for the robot helps it form a belief about the human together with its own beliefs about the world: yet, it does not directly point to a single best way to use the model, in order to improve learning.

9 Conclusion

9.1 Retrospective and Scientific Contributions

This thesis began with the observation that humans, unlike robots, are highly competent in detecting and addressing misunderstandings, building a *mutual understanding* about the task at hand, and converging to a shared solution in *collaborative learning* situations. As *social robots* are becoming a part of educational environments, we hypothesized that they can be equipped with these abilities to participate in the interaction by monitoring and contributing to it, and thereby support building a mutual understanding which can result in learning. To explore this hypothesis, in this thesis, we presented the design and development of *mutual modeling abilities* for humanoid social robots along two directions, so that they can: (i) *assess how humans build a mutual understanding* through dialogue and actions, and (ii) *build a mutual understanding with a human*, by building and maintaining a *mental model* with their own beliefs about the activity and the human, in order to support human learning.

The relevant literature pointed to diverse roles and behaviors for the robot, various ways to analyze dialogue that can give an account of the mutual understanding between humans, and modeling approaches to represent the human in order to build a mutual understanding. Since learning is a deeper goal than immediate, performance goals that machines tend to target, it is not clear which of these approaches are effective in an educational setting, and how we should design robot skills so that they can effectively observe and interact with humans to help improve learning. In this thesis, we addressed these challenges by bringing together different perspectives, and developing on them with concrete learning goals in naturalistic situations with real students, as complementary activities in schools; by this, we could rethink earlier findings and intuitions from the literature, and bring fresh insights as well as perspectives.

To investigate how humans build a mutual understanding and develop methods for its assessment, in Part I, we designed a robot-mediated human-human collaborative problem solving activity that aims to improve the CT skills of children. By creating interdependence between the learners, it is designed to elicit situated and task-oriented dialogue, in the context of their common environment that they refer to, with the aim to achieve the task goals. What humans say and do in this activity is not only tied to how they perform together in the task, but also what they ultimately learn from it. On data we collected in a user study with 78 children doing this activity at schools, we saw the activity was successful in inducing genuine dialogues with high levels of *speech activity*, containing pauses, gaps and overlaps. Yet, the link between learning and performance was not trivial, and the speech activity levels were not sufficient to differentiate between those who performed well and who did badly, or those who learned and those who did not. Subjective measures like self-assessment of competence and mutual understanding reflected their performance in the task, but might not indicate the learning.

By delving into the *speech content* (i.e. what is said), and gathering perspectives on dialogue and collaborative learning, we explored the complex and tangled relationship between what humans say and what humans do, and the outcomes of this in our learning context. For this purpose, we focused on how humans *align* in their dialogue and actions (i.e. develop shared representations at different linguistic levels (Pickering & Garrod, 2004, 2006)); to get an insight on how they build a mutual understanding, as shallower behavioral cues the robot can detect more easily without trying to understand their minds. We developed novel rule-based algorithms the robot can use, to automatically and empirically assess the collaboration between children: in terms of how they align within their dialogue, from their dialogue to their actions, and how these lead to their performance in the task and learning outcomes. We treated alignment as a procedure that occurs in stages and over time, that the robot can keep track of; compared to a holistic approach used in other works to quantify the overall alignment. Thus our measures study more in depth how alignment forms (e.g. via priming and establishment), and then builds into a function of the discourse. To study alignment, we focused on the formation of expressions related to the activity that the robot can specifically aim to detect, as these expressions target information necessary to progress in the task.

We transcribed a portion of the collected data and saw that our measures were capable of capturing alignment in our context. Our measures reflected some fine-grained aspects of learning in the dialogue; such as an exploratory and collaborative period the children go through during the alignment process that can be automatically inferred with our measures, even if we can not conclude that overall they are linked to the final measure of learning. Our measures captured more aspects of performance than learning, as the measures focus on what was said specifically about the environment, and the immediate apparent changes to the environment. Our results, albeit limited to a small dataset, highlight that in a situated educational activity, focusing simply on expressions related to the task can still give good insights into the nuances of collaboration as an effort to build a mutual understanding through dialogue, and its ultimate links to task success; with the awareness that there are several other aspects that remain to be observed in the dialogue such as gaze patterns, prosodic features, and so on. We made this dataset and tools to study alignment in dialogues publicly available.

To explore how a robot can build a mutual understanding with a learner, in Part II, we equipped the robot with mental modeling abilities; to enable it to process a dialogue on-line, track and perform *grounding*, and thus build a mutual understanding. Rather than the shallower behavioral cues the robot can detect and replicate without trying to understand the human's mind, we gave the robot its own mind, with its own beliefs about the activity, the human, as well as the human's beliefs about the robot. To this end, we adapted the human-human collaborative problem solving activity from Part I to a human-robot version: the robot collaborates with a human to construct a shared solution, via suggesting actions and (dis)agreeing with each other in turns; a joint decision is made only if it is suggested by one and agreed by the other. Therefore, the human and the robot need to build a mutual understanding on what a correct solution is, and converge to a solution that they agree upon in order to solve the problem.

The adapted activity served as a “dialogic game", with the content to be grounded as the joint decisions that build up in the collaboration. In this thesis, we chose to realize this through actions on the activity rather than spoken dialogue, as speech is currently difficult with robots; which was also shown by the ASR results on the dataset in Part I. In our pre-experiment, we saw that children had no problems with disagreeing with the robot; we observed a whole spectrum from episodes of “fight" where the children asserted their beliefs in one end, to systematic agreements with the robot, including its incorrect choices, in the other end. This suggested the children did “enter" into the dialogic game with the robot, as if a real dialogue.

We concentrated on how the robot can represent the activity and the human learner in its mental model, and how it can use its mental model to guide its behavior. About the activity, the robot builds and maintains *zeroth-order beliefs*, i.e. its own mental model about the activity. The beliefs are verbalized in the form of “I think c is a correct choice.", on the choices the robot can make, in terms of the degree that the robot thinks that they are correct by the robot's own understanding. The zeroth-order beliefs are updated by the robot's reasoning on the activity. We designed two ways the robot can reason about the activity in our context: with (i) a *correct* conception of the problem, leading to a correct solution, or (ii) an *incorrect* conception of the problem, with which the robot can not solve the problem by itself. These illustrate, and reflect in our context, different amounts of knowledge a robot can have about a given activity, in terms of how to achieve the goals of the activity; to solve the given problem in our context. In both ways of reasoning, given a belief “I think c is a correct choice.", the robot generates verbal explanations “Because it is better than choosing ... ", that reflects its algorithmic reasoning.

To represent the human, the robot builds and maintains *first-order beliefs* about the human, i.e. beliefs ascribed to the human about the activity. The beliefs are verbalized in the form “I think that the human thinks c is a correct choice.", on the degree that the human thinks the choice is correct by the robot's understanding about the human's understanding. These beliefs are maintained by the interaction with the human via Bayesian update; a special case of it in our work. As a way the robot can utilize its mental model in the interaction, we designed a robot behavior where the robot tries to reconcile and align its first- and zeroth-order beliefs to build a *common ground* with the human, which consists of their agreed-upon choices. When a choice is grounded, i.e. added to their common ground, the robot verbalizes this explicitly, e.g. “Great! Then, we both think it is a correct choice.". Then, the robot adapts to the human, by preferring choices in their common ground the next time, e.g. if they fail and try again. The robot raises discrepancies in its beliefs about the human and the observed human behavior; e.g. when the human makes a choice that contradicts with an ascribed belief, it says "Really? I thought you did not think it was a correct choice!", and revises its belief "Then you think it is a correct choice.". Thus the robot instigates a dialogue, where it makes an explicit effort to build a mutual understanding with the human, and assist their convergence to a shared solution.

With the zeroth-order beliefs, the robot has its own beliefs about the activity that reflect the way it reasons and takes actions; the robot is de facto an *intentional system*, that can be understood by the human through adopting the *intentional stance* towards the robot. With the addition of the first-order beliefs, the robot now recognizes the human as an intentional system as well. The robot essentially adopts the intentional stance towards the human to understand and predict the human's actions, and hence establishing a reciprocity; but terminating a recursion that can go on at level 1. We extended the belief system we described to *second-order beliefs*, that the robot builds and maintains about the human's beliefs about the robot. These beliefs are verbalized in the form "I think that the human thinks that I think *c* is a correct choice.". In this thesis, we illustrated how second-order beliefs can be maintained in our context; but concentrated on designing robot behaviors that use zeroth-order beliefs only, or zeroth-order beliefs together with first-order beliefs: our goal was to see what the robot needs to know about the activity and/or the human to have an effective dialogue in a collaborative learning scenario. For this purpose, we investigated the potential benefits and effects of the simplest, predictable mental models the robot can have to drive its behavior in our educational context.

To explore the effects of the mental modeling abilities that we developed for the robot, we conducted a between-subjects experiment with 61 children in schools, where we compared three robot behaviors: (i) the *optimal robot* with a correct model of the activity, and no model of its human counterpart, (ii) the *sub-optimal robot* with an incorrect model of the activity, and (iii) the *aligning robot* with a model of the human, together with an incorrect model of the activity. The aligning robot explicitly models the human, contributes to the dialogue accordingly by acknowledging and reflecting on the beliefs the human can have about the activity as discussed above. It adapts to the human, by preferring their agreed-upon choices as they build a shared solution; while the other robots do not explicitly model nor adapt to the human. These robot behaviors allowed comparing effects of the mental models along two axes: on one axis, their conception of the problem ranges from incorrect to correct with the sub-optimal to the optimal robot. On the other axis, their adaptivity and cognitive capability ranges from no-adaptation and no model of the human, to adaptive with a model of the human, with the (sub-)optimal to the aligning robot. We made the code that represents the activity and governs the interaction with the robot for the three behaviors publicly available.

We inspected the dialogues these robots generated, what kind of collaboration dynamics they induced, and the effects on the learning outcomes to validate the modeling process that we developed. We witnessed the modeling actually produces credible dialogues, and saw that the collaborations resulted in positive learning outcomes irrespective of the robot's behavior; while we did not observe evidence for a difference in the learning outcomes between the robot behaviors overall. The results showed as the activity proceeds, the sub-optimal and the aligning robots can tell from the way the human interacts with them whether they are learning or not, while the optimal robot can not. Therefore, those robots can act on the basis of whether the human is learning or not, e.g. by changing strategy. In contrast, the optimal robot resulted in the best learning for children without prior knowledge, as it seems to be effective in helping them overcome their initial incomprehension about the problem, by providing them with the opportunities to reflect on the robot's own actions and understand the underlying principles.

In the case of the optimal robot, showing the human a correct way to solve the problem while giving explanations, correcting the human's mistakes, and hoping the human will grasp it did not work for everyone. Or else, not knowing how to solve it, but building a common ground of choices with the human, explicitly acknowledging their contrary and shared beliefs, and hoping the human will recognize where they are wrong was also not sufficient: neither robots are challenging the choices of the human as much as they should for the learner to reflect and recognize where they were wrong. Thus, the differences we observed in the dynamics of collaboration and the effects on learning outcomes paved the way for a future robot behavior that can do this: by intelligently using its knowledge about both the problem and the human.

The future robot can leverage correct and incorrect conceptions of the problem, together with its mental model about the human. Based on its initial interaction with the human, as well as the prior knowledge of the human if available to the robot, the robot can change its strategy about the problem during the interaction with respect to its model about the human. It can start like the sub-optimal robot, giving incorrect suggestions to gauge the human's understanding and represent this in its mental model; "making a mental note" of the human's misconceptions. Then, it can gradually act more like the optimal robot, helping the human recognize the misconceptions, that can lead to a conceptual change and hence learning. Here, the second-order beliefs can come into play: they are essentially the image of the robot in the human's mind. The robot is effectively by changing its behavior manipulating this image to the ends for the human to encounter a socio-cognitive conflict with the robot. Thus, second-order beliefs can be used to create, keep track of, and deliberately induce differences between zeroth-order beliefs and second-order beliefs: e.g. the robot makes a choice, showing the human (as what the robot thinks that human thinks) that the robot thinks it is a correct choice, encoded as a second-order belief, while the robot actually thinks it is not, as captured by a zeroth-order belief. This kind of clash, that can be anticipated and induced, can help learning.

9.2 Shortcomings and Future Directions

In this section, we discuss the shortcomings that are relevant to the mutual modeling skills we developed in this thesis, as well as future directions towards a more natural and intuitive interaction with the robot; where the robot can more actively use its embodiment and multi-modal capabilities to enhance the interaction, within our context as an example. For these directions, we expand on our initial explorations and takeaways that can inform future investigations.

9.2.1 Towards Natural Language Understanding Skills for the Robot

In Part I, our focus was on developing measures to assess the mutual understanding between humans, in terms of how they aligned in their dialogue and actions: the proposed alignment measures are *automatic*, yet we evaluated them on the manual transcripts of the corpus, seen here as gold-standard Automatic Speech Recognition (ASR) outputs. The state-of-the-art ASR performance is evolving quickly and we did not want to adapt our metrics to specific ASR errors. Other works that we discussed in Chapter 2 have used similar terminology (e.g. automatic methods), but are subject to the performance of the ASR system, or even dependent on specific forms of segmentation. Our investigation for the feasibility with ASR showed only poor results, as described in Section 4.2.3. A future design of the activity can contain referents that are simpler to detect by ASR e.g. “House", “Market", instead of “Mount Neuchâtel", “Mount Luzern" that are subject to too many variations in pronunciation. Then a robot using ASR can be capable of tracking the alignment between the humans by adapting our measures.

In Part II, we focused on cognitive, mental modeling capabilities for the robot, how to represent the activity and the human, and use its model to guide the behavior; aiming robust full autonomy for the robot. Thus, in the JUSThink human-robot activity we developed, as the robot and the human construct a solution together, we chose the human to use a touch screen that displayed the game scene to take actions, while the robot takes actions via direct commands to the activity as well as verbalizing its intentions and actions. Although the human can understand the utterances of the robot, the robot currently relies on the human's use of the touch screen, and can not comprehend what is said if the human were to speak; although it is very natural to respond verbally, especially to a humanoid robot. Furthermore, speaking out loud, externalizing one's thoughts, can help the human think and reflect about the choice.

We explored how we can make the robot understand the intentions of the human from speech in our context. In a collaborative effort with Yiwen Ma, whose semester project was co-supervised by the author of this thesis, an intention recognition system was iteratively designed and evaluated. The system recognized the verbal content of speech via ASR using Google Speech Recognition API (v2), with a feedback system indicating if the robot is listening to the human or not, and then inferred the human's intention through keyword analysis. The robot responded with follow-up questions to confirm or repair its understanding of the utterances, and finally executed the action. Two ways of giving feedback were compared: visual with colored light (green for listening, red for processing) and audio by making sound effects. Case studies with adults suggested that it is easy to recognize intentions like (dis)agreeing and submitting, by detecting keywords such as “yes", “no", “submit". However, more complex intentions like suggesting a particular connection benefited from a feedback system, where the audio feedback performed better than the visual feedback: the participants found the audio feedback less distracting and more engaging: they reported feeling less stressed, and wanted to interact with that robot more.

These early findings point to a need for further research to explore ways a robot can incorporate the comprehension of speech, as a joint activity between the robot and the human: this calls for signalign mechanisms to acknowledge understanding and repair mechanisms to address misunderstandings, that can be enacted through multiple modalities. These need to be intuitive to the human, by the robot reflecting to the human the awareness of and attention to the human, that becomes crucial for the functional goal to convey one's message to the robot. Crucially, the acceptance of the robot is more important than the functional purposes being achieved: a big red blinking light is too alarming for a human unless something is really wrong, but whatever the message is, it would be well received. If anything, the interaction with that robot would likely be kept as short as possible, and this, would be missing the whole point.

9.2.2 Building and Maintaining the Mental Model

Contents of the mental model In the mental model of the robot, we chose to represent beliefs over the particular choices available in an activity, in terms of whether they are believed to be correct by the robot's understanding. This assumes the robot perceives the activity, has access to the list of available actions, and has a way to quantify and compare the choices in terms of what would happen, e.g. how much it would cost, by some understanding of its own.

There are other kinds of mental states that can be considered than beliefs held by the robot as its own or as ascribed to the human. For instance, in the traditional belief-desire-intention models that implement Bratman (1987)'s theory of human practical reasoning, the intentions or desires (goals) are also characteristic parts for the reasoning about the agent's actions. The goals and motivations of the human can be different than achieving the goals of the given task in the activity; the human might just want to have fun, go against the robot to see its limits, or anything else than work together with the robot to find a correct solution to the given problem in our context. These differing goals can be also represented by the robot as beliefs over the goals, about the activity or the other purposes the human can have: "I think the human wants p.", where p="achieve the task's goals", "just have fun (and not achieve the task's goals)", etc.

Enumerating which goals or mental states are relevant to consider, or trying to determine the mental state might not be possible; as they would be under-determined by evidence. In any case, it might not matter as much to be right in the inference. Did you ever think that your alarm clock is very nasty with you, forcing you to wake up; humans also project intentions on computers ("computers don't like me" say many), cars ("please be nice with me and turn on please?"). Similarly, any difficulty in interaction, or any inconvenience for that matter can be interpreted in terms of favoring relationships. Whatever our robot says, does, or thinks, the learner can simply see it as being nasty: "It doesn't like me", and with a good reason: "because it disagrees with me". Luckily for us, a social robot has the opportunities and modalities to try to convince the learner otherwise; while in the end "it wants you to learn and have fun".

The beliefs of the robot need not be only about the activity, but can be about the human, in the form "I think that the human is p.": p can be any attitude of the human regarding the overall interaction, the activity, the robot, or none of these (such as affective states or personality traits): for instance, p =“engaged in the interaction", “competent in this activity", “looking at the robot", “excited", or “creative". These can be inferred with some uncertainty by processing the social signals, e.g. inference of what the child is looking at from the gaze direction. These are still zeroth-order beliefs: they are about the human, not attributed to the human as beliefs of the human. Likewise, the first-order beliefs that are attributed to the human can be about the robot (and not about the activity): “I think that the human thinks that the robot is p.", with p='competent", etc. We did not consider such beliefs in this work; our general framework of having, attributing, and updating beliefs can be extended to include these (explicit) beliefs.

Maintenance of the mental model In this work, we used a special case of Bayesian update that results in binary attributions of belief: given a particular choice c, the robot can think that c is a correct choice, or else it can think that c is not a correct choice, otherwise it has no belief about that choice. Still, we prefer representing these as probabilities 1 and 0: this allows for later grading the beliefs. The values can still be discretized in terms of confidence in a belief, e.g. in a manner similar to a Likert scale, from strongly disagree to strongly agree for “I think that c is a correct choice". Then, the robot can verbalize these linguistic representations of its belief, or compare two of its beliefs at different orders to see if they are similar or not, e.g. in order to raise discrepancies between its own beliefs and the beliefs it ascribed to the human.

Representation of and reasoning with the mental model We employ a state representation akin to finitely-nested I-POMDPs to recursively represent and maintain the robot's beliefs about the activity and the human learner: a full specification of the problem as an I-POMDP is likely to be excessive and impractical to solve as a planning problem for the robot; more importantly it might not be relevant for the interaction with a human, especially since it can elude the simpler or the core reasons why the robot decides on a specific action, to the extent that can be understood and reasoned with by the human.

Raising discrepancies is currently rule-based in our implementation: this and other rules encode logical reasoning that would follow within a logic-based framework. Meanwhile, the formalization provided by I-POMDPs encode the mechanistic aspects of the state transitions and available actions, and present a formal way to bound the recursion in the attribution of beliefs. Reasoning to solve an I-POMDP intrinsically includes an awareness of the actions, transitions, and the rewards: each step in the solving process consists of hypothetical, conditional reasoning with possible future states, in the form “if I take this action, it leads to this, ..., and we reach to an optimal solution": planning methods do this by computation. I-POMDPs can still allow for extracting logical descriptions, that can be used in logical reasoning. Thus, for future work, a more rigorous and principled approach that accounts immediately for phenomena such as discrepancies is a blend of logic, particularly DEL that can capture and reason with contradictions, and the formalization provided by I-POMDPs. Further research should investigate ways to further link the logical reasoning via DEL with planning problem specification like I-POMDPs to feed each other, and grow on the top of each other's shoulders.

9.2.3 Evaluation of Learning

To measure the learning effects, we adopted a pre-test and post-test design where the tests are administered immediately before and after the activity. We focused on assessment of the differential effects, by computing the learning gain from the pre-test and the post-test scores.

In Part I, the tests were in a context different than the collaborative activity, that required a far transfer from the task to the test. They were asked in a multiple-choice format. Multiple choice questions have their caveats, as they are not open to other solutions besides those that are given. The fact that only one option is correct motivates other types of reasoning, that include comparing and contrasting the options, which can have learning effects. Meanwhile, far transfer is the type of transfer we really want in learning: with the abstraction it brings, indicating a deeper understanding with the very ability to apply to different contexts. In Part II, we designed new tests to assess for a more conservative goal and measure near transfer, where the participant individually solves different instances of the problem in the same setting. This is easier to achieve; still, only slightly more than half of the children had positive learning gains in User Study 3: hence, it is not straightforward to achieve learning in this case as well.

Since the post-test was immediately after the activity, it could not measure if the knowledge gained is retained for later in the day, or some days or months after: what we really want. A *delayed test* could be administered to see the retention, as a better indicator for learning. In the tests, the participants answer several questions of the same type: this can result in learning due to doing the test itself, which is undesirable as we want to see the impact of our intervention, i.e. the collaboration with the robot; the tests are meant to be for assessment only. As the tests did not provide negative or positive feedback on the correctness of the answer, the possible learning effects as the participant is doing a test is mitigated as much as we could.

Our experiment via User Study 3 only compared different robot behaviors to each other, and did not make a comparison with a traditional/pen-and-paper approach by a teacher illustrating the concepts. This would be needed to show the effectiveness of using a robot over a human teacher. In this thesis, our focus was instead to validate if robot skills to produce credible dialogues, compare robot behaviors to see the differences in how the interactions evolve, and whether this has an effect on the learning outcomes. Thus, we rather assume the robot's presence; we ask what can we do with the robot, given that it is there, ready to engage.

9.2.4 Robot Embodiment

The mutual modeling skills that we developed do not require an embodiment for the robot: to assess the mutual understanding between humans, we focused more on the assessment of the human-human dialogue than its use by the robot. For building a mutual understanding, the mental modeling by the robot is more about the agent nature of the robot: instead of a robot, it could have been a virtually embodied agent, or even a non-embodied conversational agent. The robot being the physically embodied robot is based on our assumption that the robot is there, and our goal is to develop its mutual understanding skills in the world of humans.

In learning contexts, the physical embodiment is important, as it brings clear motivational aspects to encourage the child to participate in the activity, to make one's best effort towards achieving the task's goals. In this thesis, with the mental modeling abilities and behaviors driven by the model, we have explored ways the robot can contribute to the dialogue with the human as a partner, as a counterpart to the human; as more than a motivator.

We chose to use the humanoid robot QTrobot[I] that is designed for social interaction. We added several custom gestures to emphasize on the turn-taking aspects in the JUSThink human-robot scenario, such as a me-gesture to point to self while looking at the shared screen to indicate it is the robot's turn to take an action. We primarily used speech and pre-recorded gestures in the behaviors we developed, and did not make comparisons with other robots.

QTrobot was positively perceived by our target learner group of age 9 up to 16, as also measured by a self-perception questionnaire on items based on Godspeed, as reported in Chapter 3. It proved to be a very stable and reliable platform that can operate for longer hours into consecutive experiments with different learners, as it is fixed to a position with non-actuated legs: we faced no serious issues due to e.g. heating through weeks of experiments. The robot is designed to be quite safe, with a plastic cover and no sharp parts, and safety mechanisms to turn off motor torque if an external force is applied above a certain limit which is an important feature as the robot is operating with and near children. In addition, it is simple to program (via ROS), and has a behavior library of modest size that contains facial expressions, and specifically gestures that can be expanded on quite easily. Yet, QTrobot is limited in terms of physical manipulation capabilities, especially with the hands being non-actuated and the working space being more suited for gestures than manipulation of objects. As an alternative embodiment, we considered another humanoid robot Reachy[II]. Reachy allows a precise control of the arms and hands, that can be used to simulate very lifelike gestures of acting on the same screen as it works with a learner, as well as deictic gestures to point to and higlight certain parts of the problem.

We were curious if the robot can use its embodiment to physically take actions, like a human does in our context: by touching on the screen, instead of sending commands to the activity that are appearing on the screen. Thus, we postulated that the causal link of the robot taking the action will be easier to establish by the human, when the human observes the robot move its arm and physically take the action. This could induce a more evident and stronger attribution of the action to the robot, as the true actor of the action. In a collaborative effort with Elisa Bianchi, whose semester project was co-supervised by the author of this thesis, we explored to what extent the robot Reachy can use a touchscreen. To this end, we attached a touch pen to the hand of Reachy to be used on the touchscreens we employed in our user studies (by adding aluminum foil on the pen for it to hold enough capacitive charge to

[I] QTrobot by LuxAI SA, Luxembourg, https://luxai.com
[II] Reachy by Pollen Robotics SAS, France, https://www.pollen-robotics.com

trigger touch events). We adapted the activity to add connections by pressing the two nodes separately, instead of the original press-drag-release gesture as a dragging gesture requires a precise control. The results were promising with accuracy of presses about 90%, and a precision over 99%; however the calibration remained the most fragile and prone to error.

Further research can explore the trade-off brought by physical manipulation capabilities that bring additional complexities by requiring precise measurements (e.g. by calibration, or very reliable situation assessment for the exact positions of the objects) against our perceptions of action attributions to the robot, for the human to acknowledge the agency of the robot in taking the action. We did not observe this problem to occur in our situated (co-located and happening at the same time) human-robot interaction with the shared screen. Yet, this can become crucial in determining whether the human attributes a mental state to the robot or not; e.g. if the action is interpreted to be coming from something else than the robot.

9.3 Impact and New Perspectives

As we conclude this thesis, we revisit our hypothesis that drove our endeavors to design and develop the robot's social skills in collaborative learning situations. That is, firstly, a humanoid robot equipped with mutual modeling skills can monitor and *assess* the collaboration between humans: we showed the robot can expose nuances of and give insights about how their dialogue is evolving, from what they say and do in and about the activity. Secondly, the robot can *contribute* to the interaction and support the human's learning: we showed that robots that build and maintain a mental model about the activity and the human have the potential to hold genuine dialogues with a human, and they can tell from the way the human interacts whether the human is learning or not; which can be acted upon in the robots of the future.

Bringing in the conversational and situational context to the attention of the robot was crucial in its assessment of the mutual understanding between humans. The robot then becoming an agent that is aware of its own agency for its actions and recognizing the agency of the human were essential in building a mutual understanding with the human. These are necessary steps from having robots as motivational tools which are essentially scripted to induce behaviors and influence behavior change for the human, towards social machines that acknowledge and utilize the agency of their own and the human. We hope this perspective serves as a stepping stone to realize these robots' true potential as machines that try to understand us as *humans*.

A Algorithms for Detecting Behavioral Alignment

A.1 Recognizing Instructions

RECOGNIZE-INSTRUCTIONS, used in Chapter 5, automatically infers a sequence of instructions for an input utterance via a simple rule-based algorithm, as described in Algorithm A.2. The algorithm allows inference of partial instructions, i.e. that contain one node name only (e.g. "Let's go to Mount Montreux."). It uses RECOGNIZE-ENTITIES in Algorithm A.1 to find the edit instructions in an utterance in a simple way.

Algorithm A.1: RECOGNIZE-ENTITIES finds the edit entities in an utterance via a simple rule-based named entity recognition procedure.

Input: A sequence of tokens $U = \langle t_1, t_2, \ldots, t_n \rangle$ that make up an utterance U
Output: A sequence of entities $E = \langle e_1, e_2, \ldots, e_m \rangle$

$N \leftarrow \langle$'Montreux', 'Bern', …, 'Basel'$\rangle$ `// all node names`
$A \leftarrow \langle$'add', 'remove', 'build', 'connect', 'do', 'go', 'put'$\rangle$ `// add verbs`
$R \leftarrow \langle$'away', 'cut', 'delete', 'erase', 'remove', 'rub'$\rangle$ `// remove verbs`
$E \leftarrow$ an empty sequence `// for inferred entities`
for each token $t \in U$ **do**
    $l \leftarrow$ NIL
    if $t \in N$ **then**
        $l \leftarrow$ 'Node'
    else if $t \in A$ **then**
        $l \leftarrow$ 'Add'
    else if $t \in R$ **then**
        $l \leftarrow$ 'Remove'
    if $l \neq$ NIL **then**
        $e \leftarrow$ a new entity object ; e.token $\leftarrow t$; e.label $\leftarrow l$
        insert e into E
return E

Algorithm A.2: RECOGNIZE-INSTRUCTIONS finds the instructions in an utterance.

Input: A sequence of tokens $U = \langle t_1, t_2, \ldots, t_n \rangle$ that make up an utterance U
Output: A sequence of instructions $I = \langle i_1, i_2, \ldots, i_k \rangle$

```
E ← RECOGNIZE-ENTITIES(U)
I ← an empty sequence                                   // list for inferred instructions
i ← a new instruction object; i.verb ← NIL; i.u ← NIL; i.v ← NIL
for each entity e ∈ E do
    if e.label = 'Add' or e.label = 'Remove' then
        if i.verb ≠ NIL then                            // inferring an instruction
            if i.u ≠ NIL then                           // save the partial instruction
                insert i into I
            i.u ← NIL; i.v ← NIL
             // clear nodes of the instruction to start a new one
            i.verb ← e.label
    else if i.u = NIL then                              // that is, e.label = 'Node'
        i.u ← e.token
    else if i.v = e.token then
        if i.u ≠ e.token then                           // if not repeating node name
            i.v ← e.token
        if i.verb = NIL then                            // default to a verb if not detected
            if I.length = 0 then                        // no previous instruction: default to 'Add'
                i.verb ← 'Add'
            else                                        // default to previous instruction's verb if exists
                i.verb ← I[I.length − 1].verb
            insert i into I
            i ← a new instruction object (i.verb ← NIL, i.u ← NIL, i.v ← NIL)
    return I
```

A.2 Detecting Follow-up Actions of the Instructions

MATCH-INSTRUCTIONS-TO-ACTIONS, used in Chapter 5, pairs instructions with actions as matches or mismatches for a verbal and physical actions list A, as described in Algorithm A.4. It uses CHECK-MATCH in Algorithm A.3 to compare an instruction and and action to decide it it is a match or a mismatch.

Algorithm A.3: CHECK-MATCH checks if an instruction matches with the action. It allows partial matching for partially inferred instructions (i.e. only one of the node names is mentioned).

Input: An instruction i and an action a
Output: True if the intended action in i and action a match, False otherwise

```
// If different action, e.g. instruct removal vs. add action: mismatch.
if i.action ≠ a.verb then
    return False
u ← a.object.u          // first node in the edited edge
v ← a.object.v          // second node in the edited edge, sorted by the direction
// If only one node is inferred: try to match the instruction partially.
if i.v = NIL then
    // If the node matches either node of the edge: match.
    if i.u = u or i.u = v then
        return True
    else
        return False
// If both nodes match in either direction (undirected edge): match.
else if (i.u = u  or  i.u = v)  and  ( i.v = u  or  i.v = v) then
    return True
else
    return False
```

Algorithm A.4: MATCH-INSTRUCTIONS-TO-ACTIONS links a list of pending instructions to actions as matches or mismatches.

Input: A sequence of verbal and physical actions $A = \langle a_1, a_2, \ldots, a_k \rangle$
Output: A sequence of $M = \langle m_1, m_2, \ldots, m_k \rangle$ holding (mis)match info m_i for each a_i

```
P ← an empty sequence for pending instructions to be matched
M ← an empty sequence for (mis)match for each action in A
attempt ← 1            // submission no to clear the pending instructions list
turn ← 1               // turn no to clear the pending instructions list
foreach action a ∈ A do
    // Clear pending instructions if a new turn or attempt (submission).
    if a.turn = turn + 1 then
        clear P                        // remove all items in the sequence P
        turn ← a.turn                  // update the current episode (new turn)
    else if a.attempt = attempt + 1 then
        clear P                        // remove all items in the sequence P
        attempt ← a.attempt            // update the current episode (new attempt)
    // Say action: recognize instructions and update pending instructions.
    if a.verb = 'says' then
        I ← RECOGNIZE-INSTRUCTIONS(a.object) // a.object is the utterance
        foreach instruction i ∈ I do
            i.agent ← a.subject        // set the instructing agent
            insert i into P            // update the pending instructions
    // Do action: try to match with a pending instruction.
    else if a.verb = 'does' then
        I' ← {i : i ∈ I and i.agent ≠ a.subject} // filter for the other's instructions
        m ← a new matching object ; m.match ← NIL
        if I'.length > 0 then          // there is an instruction that may (mis)match
            // Try to match a pending instruction with the current action.
            foreach instruction i ∈ I' do
                if CHECK-MATCH(i, a) then
                    m.match ← True; m.instruction ← i; m.action ← a
            if m.match = NIL then                  // no matches, hence a mismatch
                i ← I'[I'.length − 1]      // get the last instruction by the other
                m.match ← False; m.instruction ← i; m.action ← a
            // Process the match (if matched or mismatched).
            if m.match ≠ NIL then              // Match is True and mismatch is False
                M[i] ← m                   // add the inference to (mis)matches list
                // remove matching instructions from pending instructions
                foreach instruction i ∈ P do
                    if CHECK-MATCH(i, a) then
                        remove i from P
return M
```

B Algorithms of the Robot Strategies in JUSThink Domain

The strategies presented below are used to illustrate ways the robot can reason with its mental model in Chapter 7, and constitute policies that determine the action selection procedures for the behavior of the optimal, sub-optimal and aligning robots in Chapter 8.

B.1 A Correct Strategy (Jarník's Algorithm)

POLICY-SINGLE-ZEROTH and POLICY-COLLABORATIVE-ZEROTH choose an action to guide the action selection of an agent in the JUSThink domain, that interacts with a single- or multi-agent collaborative world, respectively—see Algorithm B.1 and Algorithm B.2. The policies are correct, in leading to an optimal solution, if they use INTENTION-CORRECT in Algorithm B.4 to find equivalence classes for intended and remaining action choices, by Jarník's algorithm.

Algorithm B.1: POLICY-SINGLE-ZEROTH selects robot's action in a single-agent world of JUSThink domain based on its explicit zeroth-order beliefs about the correctness of choices in the world.

Input: current belief b about the world state, correctness of choices, and the current node
Output: decided action a,
updated belief b

$(I, R) \leftarrow$ INTENTION-CORRECT(b) // or INTENTION-INCORRECT(b), depending on strategy
$b \leftarrow$ BELIEF-UPDATE-ZEROTH(I, R, b)
// Select an action according to the beliefs, which entails an explanation
$s \leftarrow \arg\max(b.\text{world})$ // "I think the world is in state s."
// If there is no connection available: submit
if $I \cup R = \emptyset$ **then**
  $a \leftarrow$ submit
// Otherwise: pick a connection according to the current beliefs
else
  $e^* \leftarrow \arg\max_{e \in I \cup R}(b.\text{correct}[e])$ // e^* is among the most desired edges, if any
  $a \leftarrow \text{pick}(e)$ // Explanation: "I think e is a correct choice, because
  // <explanation based on the strategy>.": e.g.
  // Correct strategy: "it is the best from the ones that are connected."
  // Incorrect strategy: "it is better than connecting those in R."
return (a, b)

Algorithm B.2: POLICY-COLLABORATIVE-ZEROTH selects robot's action in a collaborative world of JUSThink domain based on its explicit zeroth-order beliefs about the choices in the world.

Input: current belief b about world state, correctness of choices, current node, and disagree counts, persistence criterion T_{dis} (the disagree threshold)
Output: decided action a, updated belief b

$(I, R) \leftarrow$ INTENTION-CORRECT(b) `// or INTENTION-INCORRECT(b), depends on strategy`
$b \leftarrow$ BELIEF-UPDATE-ZEROTH(I, R, b)

$e^* \leftarrow \arg\max_{e \in I \cup R}(b.\text{correct}[e])$ `// e* is among the most desired connections, if any`
$s \leftarrow \arg\max(b.\text{world})$ `// "I think the world is in state s."`

`// If there is no connection available: submit`
if $I \cup R = \emptyset$ **then**
    $a \leftarrow$ submit

`// Or else, if there is no connection suggested by the other: suggest`
else if $s.\text{suggested} = $ NIL **then**
    $a \leftarrow$ suggest(e^*) `// Explanation: "I think e* is a correct choice, because`
    `// it is the best from the ones that are connected."`

`// Otherwise (there is a suggested connection): agree or disagree`
else
    $e \leftarrow s.\text{suggested}$ `// "I think e is the suggested connection."`

    `// If the suggestion is equivalent to the intended action: agree`
    if $b.\text{correct}[e] = b.\text{correct}[e^*]$ **then**
        $a \leftarrow$ agree `// Explanation: "I agree with the suggestion, because`
        `// it is the best from the ones that are connected."`

    `// Or else, if the disagree threshold reached: agree`
    else if $b.\text{nDisagree}[e] \geq T_{dis}$ **then**
        $a \leftarrow$ agree `// Explanation: "Fine, since you insist so much!"`
        `// I still do not think it is a correct choice."`

    `// Otherwise: disagree`
    else
        $a \leftarrow$ disagree `// Explanation: "I disagree, I do not think it is correct:`
        `// it is not the best from the ones that are connected."`
        $b.\text{nDisagree}[e] \leftarrow b.\text{nDisagree}[e] + 1$ `// increment the disagree counter for e`

return (a, b)

Algorithm B.3: BELIEF-UPDATE-ZEROTH updates zeroth-order beliefs about a JUSThink world, from intended and remaining choices decided by a reasoning mechanism such as INTENTION-CORRECT.

Input: equivalence class of intended choices I,
set of remaining choices R,
current zeroth-order belief b about the choices
Output: updated beliefs b

```
// Update belief b about the intended available choices.
for each edge e ∈ I do
    b.correct[e] ← 1        // "I think (connecting) e is a correct choice."
// Update beliefs about the remaining available choices.
for each edge e ∈ R do
    b.correct[e] ← 0        // "I think e is not correct choice."
return b
```

Algorithm B.4: INTENTION-CORRECT presents a reasoning mechanism with zeroth-order beliefs about a JUSThink world, that prefers optimal choices by the Jarník's algorithm.

Input: belief b about the world state and the current node
Output: an equivalence class of intended choices I by Jarník's algorithm,
remaining choices R

```
s ← argmax(b.world)                          // "I think the world is in state s."
c ← argmax(b.current)                        // "I think the current node is c."
// Build a node set V of the selected nodes to pick edges from.
V ← {c}                                      // select the start node
for each edge (u, v) ∈ s.T do               // selected edges s.T at the state s if any
    V ← V ∪ {u, v}                           // add the nodes of the edge
// Make an edge set Q from selected to non-selected nodes, find min cost.
w* ← NIL; Q ← ∅
for each node u ∈ V do
    for each node v ∈ s.G.adj[u] do      // for each neighbor in background network s.G
        if v ∉ V then                            // if it is a non-selected node
            Q ← Q ∪ {(u, v)}
            if w* > s.w(u, v) or w* = NIL then     // update min from cost function s.w
                w* = s.w(u, v)
// Assign min-cost edges as the intended and remaining as not-intended.
I ← ∅; R ← ∅
for each edge (u, v) ∈ Q do
    if s.w(u, v) = w* then                       // if the edge has the minimum cost
        I ← I ∪ {(u, v)}                       // it is equivalent to the intended edges
    else
        R ← R ∪ {(u, v)}                       // it is a remaining edge with higher cost
return (I, R)
```

B.2 An Incorrect Strategy (A Locally Greedy Algorithm)

The policies POLICY-SINGLE-ZEROTH and POLICY-COLLABORATIVE-ZEROTH are incorrect, i.e. they will not lead to an optimal solution, if they use INTENTION-INCORRECT in Algorithm B.5 to find equivalence classes for intended and remaining action choices, by a locally greedy algorithm.

Algorithm B.5: INTENTION-INCORRECT presents a reasoning mechanism with zeroth-order beliefs about a JUSThink world, that prefers choices that can be sub-optimal by a greedy algorithm.

Input: belief b about the world state and the start node to build the equivalence class
Output: equivalence class of intended choices I by Jarník's algorithm,
remaining choices R

```
s ← argmax(b.world)                        // "I believe the world state is s."
c ← argmax(b.current)                      // "I believe the start node is c."
// Make an edge set Q from c to non-selected nodes and find the min cost.
w* ← NIL; Q ← ∅
for each node v ∈ s.G.adj[c] do            // for each neighbor of c in network s.G
    if v ∉ V then                          // if it is a non-selected node
        Q ← Q ∪ {(u, v)}
        if w* > s.w(u, v) or w* = NIL then // update min from cost function s.w
            w* = s.w(u, v)
// Assign min-cost edges as the intended and remaining as not-intended.
I ← ∅; R ← ∅
for each edge (u, v) ∈ Q do
    if s.w(u, v) = w* then                 // if the edge has the minimum cost
        I ← I ∪ {(u, v)}                   // it is equivalent to the intended edges
    else
        R ← R ∪ {(u, v)}                   // it is a remaining edge with higher cost
return (I, R)
```

B.3 A Mutual Modeling Strategy (An Aligning Algorithm)

POLICY-ALIGNING in Algorithm B.6 chooses an aligned action, or else if not available, an action chosen by the locally greedy algorithm to guide the action selection of the robot in the JUSThink domain. It uses BELIEF-UPDATE-ALIGNING in Algorithm B.7, instead of BELIEF-UPDATE-ZEROTH in Algorithm B.3, to update its first-order beliefs about the other as well as its zeroth-order beliefs in its reasoning, to refer to its first-order beliefs while revising its zeroth order beliefs.

Algorithm B.6: POLICY-ALIGNING selects robot's action in a collaborative world of JUSThink domain based on first- and zeroth-order beliefs about choices in the world.

Input: current zeroth-order beliefs b_0 about world state, choices, current node, and disagree counts
current first-order beliefs b_1 about the other's beliefs about choices
persistence criterion T_{dis} (the disagree threshold)

Output: decided action a,
updated zeroth-order belief b_0

```
(I, R) ← INTENTION-INCORRECT(b) // We experiment with the incorrect strategy.
b_0 ← BELIEF-UPDATE-ALIGNING(I, R, b_0, b_1)
E*_0 ← {argmax_{e∈I∪R}(b_0.correct[e])} // E*_0 is the set of desired connections by own
E*_1 ← {argmax_{e∈I∪R}(b_1.correct[e])} // E*_1 is the set of inferred intentions by other
s ← argmax(b.world)    // "I think the world is in state s."
// If there is no connection available: submit
if I ∪ R = ∅ then
    a ← submit
// Or else, if there is no connection suggested by the other: suggest
else if s.suggested = NIL then
    Pick the first e* in the set E*_0
    a ← suggest(e*)     // Explanation: "I think e* is a correct choice, because
                        // it is the best from the ones that are connected."
// Otherwise (there is a suggested connection): agree or disagree
else
    e ← s.suggested     // "I think e is the suggested connection."
    // If the suggestion is in the common ground: agree
    if e ∈ E*_0 ∩ E*_1 then
        a ← agree       // Explanation: "I agree with the suggestion, because
                        // we both think it is a correct choice."
    // If the suggestion is equivalent to intentions: agree
    if b_0.correct[e] = b_0.correct[e*] then
        a ← agree       // Explanation: "I agree with the suggestion, because
                        // it is better than connecting those in R."
    // Or else, if the disagree threshold reached: agree
    else if b_0.nDisagree[e] ≥ T_dis then
        a ← agree       // Expl.: "I agree, now we both think that it is correct,"
                        // although previously I did not think it was correct."
        b_0.correct[e] ← 1 // "I (now) think e is correct choice."
    // Otherwise: disagree
    else
        a ← disagree    // Explanation: "I disagree, I do not think it is correct:
                        // it is not the best from the ones that are connected."
        b_0.nDisagree[e] ← b_0.nDisagree[e] + 1 // increment the disagree counter for e
return (a, b_0)
```

Algorithm B.7: BELIEF-UPDATE-ALIGNING updates zeroth-order beliefs about the other in a JUSThink world, from intended and remaining choices in comparison to first-order beliefs.

Input: equivalence class of intended choices I
remaining choices R
current zeroth-order beliefs b_0 about the choices
current first-order beliefs b_1 about the other's beliefs about choices
Output: updated zeroth-order beliefs b_0

`// Update beliefs about the intended available choices.`
for each edge $e \in I$ **do**
  if b_0.correct[e] $\neq$ b_1.correct[e] **then**
    b_0.correct[e] $\leftarrow 1$ `// "I think (connecting) e is a correct choice."`

`// Update beliefs about the remaining available choices.`
for each edge $e \in R$ **do**
  if b_0.correct[e] $\neq$ b_1.correct[e] **then**
    b_0.correct[e] $\leftarrow 0$ `// "I think e is not correct choice."`
return b_0

Bibliography

Albrecht, S. V., & Stone, P. (2018). Autonomous agents modelling other agents: a comprehensive survey and open problems. *Artificial Intelligence*, *258*, 66–95. https://doi.org/10.1016/j.artint.2018.01.002. (Cited on pages 7, 18, 20, 98)

Allen, J. F., Schubert, L. K., Ferguson, G., Heeman, P., Hwang, C. L., Kato, T., Light, M., Martin, N., Miller, B., Poesio, M., & Traum, D. R. (1995). The TRAINS project: a case study in building a conversational planning agent. *Journal of Experimental & Theoretical Artificial Intelligence*, *7*(1), 7–48. https://doi.org/10.1080/09528139508953799. (Cited on page 4)

Anderson, A. H., Bader, M., Bard, E. G., Boyle, E., Doherty, G., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., Thompson, H. S., & Weinert, R. (1991). The HCRC map task corpus. *Language and Speech*, *34*(4), 351–366. https://doi.org/10.1177/002383099103400404. (Cited on pages 37, 64)

Anderson, J. R., Boyle, C. F., Corbett, A. T., & Lewis, M. W. (1990). Cognitive modeling and intelligent tutoring. *Artificial Intelligence*, *42*(1), 7–49. https://doi.org/10.1016/0004-3702(90)90093-F. (Cited on page 18)

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching and assessing: a revision of bloom's taxonomy of educational objectives*. New York: Addison Wesley Longman. (Cited on pages 29, 78)

Argall, B. D., Chernova, S., Veloso, M., & Browning, B. (2009). A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, *57*(5), 469–483. https://doi.org/10.1016/j.robot.2008.10.024. (Cited on page 12)

Aumann, R. J. (1999). Interactive epistemology I: knowledge. *International Journal of Game Theory*, *28*(3), 263–300. https://doi.org/10.1007/s001820050111. (Cited on pages 19, 21)

Aumann, R. J., & Dreze, J. H. (2008). Rational expectations in games. *American Economic Review*, *98*(1), 72–86. https://doi.org/10.1257/aer.98.1.72. (Cited on page 19)

Aumann, R. J., & Maschler, M. (1995). *Repeated games with incomplete information*. The MIT Press. (Cited on page 19)

Austin, J. L. (1962). *How to do things with words*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198245537.001.0001. (Cited on page 4)

Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, *1*, 0064. https://doi.org/10.1038/s41562-017-0064. (Cited on pages 1, 7, 105)

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, *113*(3), 329–349. https://doi.org/10.1016/j.cognition.2009.07.005. (Cited on page 105)

Baker, C. L., Saxe, R. R., & Tenenbaum, J. B. (2011). Bayesian theory of mind: modeling joint belief-desire attribution. *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*, *1*, 2469–2474. (Cited on page 105)

Baker, M. J., Schwarz, B. B., & Ludvigsen, S. R. (2021). Educational dialogues and computer supported collaborative learning: critical analysis and research perspectives. *International Journal of Computer-Supported Collaborative Learning*, *16*(4), 583–604. https://doi.org/10.1007/s11412-021-09359-1. (Cited on page 14)

Bangalore Kantharaju, R., Langlet, C., Barange, M., Clavel, C., & Pelachaud, C. (2020). Multimodal analysis of cohesion in multi-party interactions. *Proceedings of the 12th Language Resources and Evaluation Conference*, 498–507. (Cited on page 14)

Baron-Cohen, S. (1995). *Mindblindness: an essay on autism and theory of mind*. The MIT Press. (Cited on page 6)

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*(1), 37–46. https://doi.org/10.1016/0010-0277(85)90022-8. (Cited on page 1)

Bartlett, M. E., Edmunds, C. E. R., Belpaeme, T., Thill, S., & Lemaignan, S. (2019). What can you see? Identifying cues on internal states from the movements of natural social interactions. *Frontiers in Robotics and AI*, *6*, 49. https://doi.org/10.3389/frobt.2019.00049. (Cited on page 7)

Bartneck, C., Kanda, T., Mubin, O., & Al Mahmud, A. (2009). Does the design of a robot influence its animacy and perceived intelligence? *International Journal of Social Robotics*, *1*(2), 195–204. https://doi.org/10.1007/s12369-009-0013-7. (Cited on page 44)

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*(1), 71–81. https://doi.org/10.1007/s12369-008-0001-3. (Cited on page 34)

Bell, T., Witten, I. H., & Fellows, M. (2015). *Computer science unplugged: an enrichment and extension programme for primary-aged children*. University of Canterbury. Computer Science; Software Engineering. Retrieved August 12, 2022, from http://hdl.handle.net/10092/247. (Cited on page 29)

Bellman, R. E. (1957). *Dynamic programming*. Princeton University Press. (Cited on pages 7, 102)

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: a review. *Science Robotics*, *3*(21), eaat5954. https://doi.org/10.1126/scirobotics.aat5954. (Cited on pages 1, 12, 13)

Bernstein, D. S., Givan, R., Immerman, N., & Zilberstein, S. (2002). The complexity of decentralized control of Markov decision processes. *Mathematics of Operations Research*, *27*(4), 819–840. https://doi.org/10.1287/moor.27.4.819.297. (Cited on page 104)

Bolander, T., & Andersen, M. B. (2011). Epistemic planning for single- and multi-agent systems. *Journal of Applied Non-Classical Logics*, *21*(1), 9–34. https://doi.org/10.3166/jancl.21.9-34. (Cited on page 21)

Borge, M., & Rosé, C. (2021). Quantitative approaches to language in CSCL. In U. Cress, C. Rosé, A. F. Wise, & J. Oshima (Eds.), *International Handbook of Computer-Supported Collaborative Learning* (pp. 585–604). Springer International Publishing. https://doi.org/10.1007/978-3-030-65291-3_32. (Cited on page 14)

Borůvka, O. (1926a). "O jistém problému minimálním" [About a certain minimal problem]. *Práce Moravské Přírodovědecké Společnosti*, *3*(3), 37–58. Retrieved October 3, 2022, from http://dml.cz/dmlcz/500114. (Cited on pages 100, 101)

Borůvka, O. (1926b). "Přisěvek k řešení otázky ekonomické stavby elektrovodních sítí" [A contribution to the solution of a problem on the economical construction of power networks]. *Elecktronický obzor*, *15*, 153–154. (Cited on pages 100, 101)

Boutilier, C., Reiter, R., & Price, B. (2001). Symbolic dynamic programming for first-order MDPs. *IJCAI'01: Proceedings of the 17th international joint conference on Artificial intelligence*, *1*, 690–697. (Cited on page 22)

Bowling, M., Burch, N., Johanson, M., & Tammelin, O. (2015). Heads-up limit hold'em poker is solved. *Science*, *347*(6218), 145–149. https://doi.org/10.1126/science.1259433. (Cited on page 19)

Boyer, K., Ha, E. Y., Phillips, R., Wallis, M., Vouk, M., & Lester, J. (2010). Dialogue act modeling in a complex task-oriented domain. *Proceedings of SIGDIAL 2010: the 11th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 297–305. (Cited on page 15)

Bratman, M. (1987). *Intention, plans, and practical reason* (Vol. 10). Harvard University Press Cambridge, MA. (Cited on page 141)

Breazeal, C., Berlin, M., Brooks, A., Gray, J., & Thomaz, A. L. (2006). Using perspective taking to learn from ambiguous demonstrations. *Robotics and Autonomous Systems*, *54*(5), 385–393. https://doi.org/10.1016/j.robot.2006.02.004. (Cited on page 104)

Breazeal, C., Gray, J., & Berlin, M. (2009). An embodied cognition approach to mindreading skills for socially intelligent robots. *The International Journal of Robotics Research*, *28*(5), 656–680. https://doi.org/10.1177/0278364909102796. (Cited on page 21)

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental*

Psychology: Learning Memory and Cognition, *22*(6), 1482–1493. https://doi.org/10.1037/0278-7393.22.6.1482. (Cited on page 17)

Brentano, F. (2009). *Psychologie vom empirischen standpunkte* [Psychology from empirical standpoint] (O. Kraus & L. L. McAlister, Eds.; A. C. Rancurello, D. B. Terrell, & L. L. McAlister, Trans.). Taylor & Francis e-Library. (Original work published 1874). (Cited on page 6)

Brown, G. W. (1951). Iterative solution of games by fictitious play. *Activity Analysis of Production and Allocation*, *13*(1), 374–376. (Cited on page 19)

Brown, N., & Sandholm, T. (2018). Superhuman AI for heads-up no-limit poker: Libratus beats top professionals. *Science*, *359*(6374), 418–424. https://doi.org/10.1126/science.aao1733. (Cited on page 19)

Brown, N., & Sandholm, T. (2019). Superhuman AI for multiplayer poker. *Science*, *365*(6456), 885–890. https://doi.org/10.1126/science.aay2400. (Cited on page 19)

Buisan, G., Favier, A., Mayima, A., & Alami, R. (2022). HATP/EHDA: a robot task planner anticipating and eliciting human decisions and actions. *2022 International Conference on Robotics and Automation (ICRA)*, 2818–2824. https://doi.org/10.1109/ICRA46639.2022.9812227. (Cited on page 21)

Butera, F., Sommet, N., & Darnon, C. (2019). Sociocognitive conflict regulation: How to make sense of diverging ideas. *Current Directions in Psychological Science*, *28*(2), 145–151. https://doi.org/10.1177/0963721418813986. (Cited on pages 2, 62)

Campbell, M. S., & Marsland, T. A. (1983). A comparison of minimax tree search algorithms. *Artificial Intelligence*, *20*(4), 347–367. https://doi.org/10.1016/0004-3702(83)90001-2. (Cited on page 19)

Carmel, D., & Markovitch, S. (1996). Learning models of intelligent agents. *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, *1*, 62–67. (Cited on page 20)

Caruana, N., Moffat, R., Blanco, A. M., & Cross, E. S. (2022). Perceptions of intelligence & sentience shape children's interactions with robot reading companions: a mixed methods study. https://doi.org/10.31234/osf.io/7t2w9. (Cited on page 13)

Chan, L., Hadfield-Menell, D., Srinivasa, S., & Dragan, A. (2019). The assistive multi-armed bandit. *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 354–363. https://doi.org/10.1109/HRI.2019.8673234. (Cited on page 20)

Chase, C. C., Chin, D. B., Oppezzo, M. A., & Schwartz, D. L. (2009). Teachable agents and the protégé effect: increasing the effort towards learning. *Journal of Science Education and Technology*, *18*(4), 334–352. https://doi.org/10.1007/s10956-009-9180-4. (Cited on page 12)

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on Socially Shared Cognition* (pp. 127–149). American Psychological Association. https://doi.org/10.1037/10096-006. (Cited on page 15)

Clark, H. H., & Marshall, C. R. (1981). Definite knowledge and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of Discourse Understanding* (pp. 10–63). Cambridge, UK: Cambridge University Press. (Cited on page 105)

Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science*, *13*(2), 259–294. https://doi.org/10.1016/0364-0213(89)90008-6. (Cited on pages 3, 4)

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*(1), 1–39. https://doi.org/10.1016/0010-0277(86)90010-7. (Cited on page 3)

Claus, C., & Boutilier, C. (1998). The dynamics of reinforcement learning in cooperative multiagent systems. *Proceedings of the Fifteenth National Conference on Artificial Intelligence and Tenth Innovative Applications of Artificial Intelligence Conference (AAAI '98/IAAI '98)*, 746–752. (Cited on page 19)

Cohen, P. R., & Perrault, C. R. (1979). Elements of a plan-based theory of speech acts. *Cognitive Science*, *3*(3), 177–212. https://doi.org/10.1016/S0364-0213(79)80006-3. (Cited on page 18)

Cohen, P. R., Perrault, C. R., & Allen, J. F. (1981). Beyond question answering. *Strategies for Natural Language Processing, 245274*. (Cited on page 103)

Conover, W. J. (1999). *Practical nonparametric statistics* (3rd ed.). John Wiley & Sons. (Cited on pages 89, 121)

Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to algorithms* (3rd ed.). MIT Press. (Cited on page 28)

Crane, T. (1998). Intentionality as the mark of the mental. *Royal Institute of Philosophy Supplements*, *43*, 229–251. https://doi.org/10.1017/S1358246100004380. (Cited on page 6)

De Raedt, L., & Kimmig, A. (2015). Probabilistic (logic) programming concepts. *Machine Learning*, *100*(1), 5–47. https://doi.org/10.1007/s10994-015-5494-z. (Cited on page 22)

Dennett, D. C. (1971). Intentional systems. *The Journal of Philosophy*, *68*(4), 87–106. https://doi.org/10.2307/2025382. (Cited on pages 6, 98)

Dennett, D. C. (1987). *The intentional stance*. MIT Press. (Cited on pages 1, 6, 20, 98)

Dennett, D. C. (1988). Précis of the intentional stance. *Behavioral and Brain Sciences*, *11*(3), 495–546. https://doi.org/10.1017/S0140525X00058611. (Cited on page 6)

Deublein, A., Pfeifer, A., Merbach, K., Bruckner, K., Mengelkamp, C., & Lugrin, B. (2018). Scaffolding of motivation in learning using a social robot. *Computers & Education*, *125*(100), 182–190. https://doi.org/10.1016/j.compedu.2018.06.015. (Cited on page 13)

DeVries, R. (2000). Vygotsky, Piaget, and education: a reciprocal assimilation of theories and educational practices. *New Ideas in Psychology*, *18*(2), 187–213. https://doi.org/10.1016/S0732-118X(00)00008-8. (Cited on page 2)

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, *1*(1), 269–271. https://doi.org/10.1007/BF01386390. (Cited on page 100)

Dillenbourg, P. (1999). What do you mean by 'collaborative learning'? In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and Computational Approaches* (pp. 1–19). Oxford: Elsevier. (Cited on pages 1, 2)

Dillenbourg, P. (2015). *Orchestration graphs: modeling scalable education*. EPFL Press. (Cited on page 79)

Dillenbourg, P., Järvelä, S., & Fischer, F. (2009). The evolution of research on computer-supported collaborative learning. In N. Balacheff, S. Ludvigsen, T. de Jong, A. Lazonder, & S. Barnes (Eds.), *Technology-Enhanced Learning: Principles and Products* (pp. 3–19). Springer Netherlands. https://doi.org/10.1007/978-1-4020-9827-7_1. (Cited on page 2)

Dillenbourg, P., & Traum, D. (2006). Sharing solutions: persistence and grounding in multimodal collaborative problem solving. *Journal of the Learning Sciences*, *15*(1), 121–151. https://doi.org/10.1207/s15327809jls1501_9. (Cited on pages 4, 15)

Dinkar, T. (2022). *Computational models of disfluencies: fillers and discourse markers in spoken language understanding* (Doctoral dissertation). Institut Polytechnique de Paris. Palaiseau, France. Retrieved August 10, 2022, from https://www.theses.fr/2022IPPAT001. (Cited on page 48)

Dissing, L., & Bolander, T. (2020). Implementing theory of mind on a robot using dynamic epistemic logic. *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI 2020)*, *2*, 1615–1621. https://doi.org/10.24963/ijcai.2020/224. (Cited on page 21)

D'Mello, S., & Graesser, A. (2013). Design of dialog-based intelligent tutoring systems to simulate human-to-human tutoring. In A. Neustein & J. A. Markowitz (Eds.), *Where Humans Meet Machines: Innovative Solutions for Knotty Natural-Language Problems* (pp. 233–269). Springer. https://doi.org/10.1007/978-1-4614-6934-6_11. (Cited on pages 14, 15)

Doise, W., & Mugny, G. (1984). *Le développement social de l'intelligence* [The social development of the intellect] (A. S. James-Emler & N. Emler, Trans.; 1st ed., Vol. 10). Pergamon Press. (Original work published 1981). (Cited on pages 2, 62)

Dönmez, P., Rosé, C., Stegmann, K., Weinberger, A., & Fischer, F. (2005). Supporting CSCL with automatic corpus analysis technology. *CSCL '05: Proceedings of the 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years!*, 125–134. https://doi.org/10.3115/1149293.1149310. (Cited on page 14)

Doshi, P., Gmytrasiewicz, P., & Durfee, E. (2020). Recursively modeling other agents for decision making: a research perspective. *Artificial Intelligence*, *279*, 103202. https://doi.org/10.1016/j.artint.2019.103202. (Cited on pages 19, 20, 98)

Doshi, P., Qu, X., & Goodie, A. (2014). Decision-theoretic planning in multi-agent settings with application to behavioral modeling. In G. Sukthankar, C. Geib, H. H. Bui, D. V. Pynadath, & R. P. Goldman (Eds.), *Plan, Activity, and Intent Recognition: Theory and Practice* (pp. 205–224). Elsevier. (Cited on page 21)

Doshi, P., Qu, X., Goodie, A., & Young, D. (2010). Modeling recursive reasoning by humans using empirically informed interactive POMDPs. *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1 - Volume 1*, 1223–1230. (Cited on page 21)

Dubuisson Duplessis, G., Clavel, C., & Landragin, F. (2017). Automatic measures to characterise verbal alignment in human-agent interaction. *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue*, 71–81. https://doi.org/10.18653/v1/w17-5510. (Cited on pages 17, 18, 54)

Dubuisson Duplessis, G., Langlet, C., Clavel, C., & Landragin, F. (2021). Towards alignment strategies in human-agent interactions based on measures of lexical repetitions. *Language Resources and Evaluation*, *55*(2), 353–388. https://doi.org/10.1007/s10579-021-09532-w. (Cited on pages 17, 18, 54, 55)

Engwall, O., & Lopes, J. (2020). Interaction and collaboration in robot-assisted language learning for adults. *Computer Assisted Language Learning*, *0*(0), 1–37. https://doi.org/10.1080/09588221.2020.1799821. (Cited on page 13)

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: a three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864–886. https://doi.org/10.1037/0033-295X.114.4.864. (Cited on pages 6, 44)

Errattahi, R., El Hannani, A., & Ouahmane, H. (2018). Automatic speech recognition errors detection and correction: a review. *Procedia Computer Science*, *128*, 32–37. https://doi.org/10.1016/j.procs.2018.03.005. (Cited on page 52)

Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences*. Brooks/Cole Publishing Company. (Cited on pages 43, 56, 92)

Ezen-Can, A., & Boyer, K. (2014). Combining task and dialogue streams in unsupervised dialogue act models. *Proceedings of the 15th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*, 113–122. https://doi.org/10.3115/v1/W14-4316. (Cited on page 15)

Ezen-Can, A., & Boyer, K. E. (2015a). A tutorial dialogue system for real-time evaluation of unsupervised dialogue act classifiers: exploring system outcomes. In C. Conati, N. Heffernan, A. Mitrovic, & M. F. Verdejo (Eds.), *Artificial Intelligence in Education* (pp. 105–114). Springer International Publishing. https://doi.org/10.1007/978-3-319-19773-9_11. (Cited on page 15)

Ezen-Can, A., & Boyer, K. E. (2015b). Understanding student language: an unsupervised dialogue act classification approach. *Journal of Educational Data Mining*, *7*(1), 51–78. https://doi.org/10.5281/zenodo.3554708. (Cited on page 15)

Fikes, R. E., & Nilsson, N. J. (1971). STRIPS: a new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, *2*(3), 189–208. https://doi.org/https://doi.org/10.1016/0004-3702(71)90010-5. (Cited on page 21)

Foerster, J., Chen, R. Y., Al-Shedivat, M., Whiteson, S., Abbeel, P., & Mordatch, I. (2018). Learning with opponent-learning awareness. *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS 2018)*, 122–130. (Cited on page 20)

Forbus, K. D. (1980). Spatial and qualitative aspects of reasoning about motion. *Proceedings of the First AAAI Conference on Artificial Intelligence*, 170–173. (Cited on page 20)

Friedberg, H., Litman, D., & Paletz, S. B. F. (2012). Lexical entrainment and success in student engineering groups. *2012 IEEE Spoken Language Technology Workshop (SLT)*, 404–409. https://doi.org/10.1109/SLT.2012.6424258. (Cited on page 17)

Fry, P. S. (1976). Success, failure, and self-assessment ratings. *Journal of Consulting and Clinical Psychology*, *44*(3), 413–419. https://doi.org/10.1037/0022-006X.44.3.413. (Cited on page 45)

Fusaroli, R., & Tylén, K. (2016). Investigating conversational dynamics: interactive alignment, interpersonal synergy, and collective task performance. *Cognitive Science*, *40*(1), 145–171. https://doi.org/10.1111/cogs.12251. (Cited on page 3)

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: a study in conceptual and semantic co-ordination. *Cognition*, *27*(2), 181–218. https://doi.org/10.1016/0010-0277(87)90018-7. (Cited on page 3)

Gauvain, M. (2020). Vygotsky's sociocultural theory. In J. B. Benson (Ed.), *Encyclopedia of Infant and Early Childhood Development* (2nd ed., pp. 446–454). Elsevier. https://doi.org/10.1016/B978-0-12-809324-5.23569-4. (Cited on page 2)

Gillies, D. (2000). *Philosophical theories of probability* (1st ed.). Routledge. (Cited on page 101)

Gmytrasiewicz, P., & Adhikari, S. (2019). Optimal sequential planning for communicative actions: a Bayesian approach. *AAMAS '19: Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, 1985–1987. (Cited on page 7)

Gmytrasiewicz, P., & Doshi, P. (2005). A framework for sequential planning in multi-agent settings. *Journal of Artificial Intelligence Research*, *24*(1), 49–79. https://doi.org/10.1613/jair.1579. (Cited on pages 7, 20, 104–106)

Gmytrasiewicz, P. J., & Durfee, E. H. (2000). Rational coordination in multi-agent environments. *Autonomous Agents and Multi-Agent Systems*, *3*(4), 319–350. https://doi.org/10.1023/A:1010028119149. (Cited on pages 20, 104)

Goodie, A. S., Doshi, P., & Young, D. L. (2012). Levels of theory-of-mind reasoning in competitive games. *Journal of Behavioral Decision Making*, *25*(1), 95–108. https://doi.org/10.1002/bdm.717. (Cited on page 21)

Goodman, B. A., Linton, F. N., Gaimari, R. D., Hitzeman, J. M., Ross, H. J., & Zarrella, G. (2005). Using dialogue features to predict trouble during collaborative learning. *User Modeling and User-Adapted Interaction*, *15*(1), 85–134. https://doi.org/10.1007/s11257-004-5269-x. (Cited on page 15)

Goodwin, R. (1995). Formalizing properties of agents. *Journal of Logic and Computation*, *5*(6), 763–781. https://doi.org/10.1093/logcom/5.6.763. (Cited on page 97)

Griffiths, T. L., Kemp, C., & Tenenbaum, J. B. (2008). Bayesian models of cognition. In R. Sun (Ed.), *The Cambridge Handbook of Computational Psychology* (pp. 59–100). Cambridge University Press. https://doi.org/10.1017/CBO9780511816772.006. (Cited on page 105)

Grosz, B. J., & Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, *12*(3), 175–204. (Cited on page 18)

Gulz, A., Haake, M., & Silvervarg, A. (2011). Extending a teachable agent with a social conversation module – effects on student experiences and learning. In G. Biswas, S. Bull, J. Kay, & A. Mitrovic (Eds.), *Artificial Intelligence in Education* (pp. 106–114). Springer. https://doi.org/10.1007/978-3-642-21869-9_16. (Cited on page 13)

Halpern, J. Y., & Moses, Y. (1990). Knowledge and common knowledge in a distributed environment. *Journal of the ACM*, *37*(3), 549–587. https://doi.org/10.1145/79147.79161. (Cited on page 19)

Hamscher, W. C. (1991). Modeling digital circuits for troubleshooting. *Artificial Intelligence*, *51*(1), 223–271. https://doi.org/10.1016/0004-3702(91)90112-W. (Cited on page 20)

Harsanyi, J. C. (1967). Games with incomplete information played by "Bayesian" players. *Management Science*, *14*(3), 159–182. https://doi.org/10.1287/mnsc.14.3.159. (Cited on page 19)

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, *57*(2), 243. https://doi.org/10.2307/1416950. (Cited on page 6)

Herzig, A., Lang, J., & Marquis, P. (2003). Action representation and partially observable planning using epistemic logic. *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (IJCAI-03)*, 1067–1072. (Cited on page 22)

Hong, Z.-W., Huang, Y.-M., Hsu, M., & Shen, W.-W. (2016). Authoring robot-assisted instructional materials for improving learning performance and motivation in EFL classrooms. *Journal of Educational Technology & Society*, *19*(1), 337–349. (Cited on page 13)

Hough, J., & Schlangen, D. (2017a). It's not what you do, it's how you do it: grounding uncertainty for a simple robot. *HRI '17: Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, 274–282. https://doi.org/10.1145/2909824.3020214. (Cited on page 5)

Hough, J., & Schlangen, D. (2017b). A model of continuous intention grounding for HRI. *"The Role of Intentions in Human-Robot Interaction" Workshop in Conjunction with the 12th ACM / IEEE International Conference on Human-Robot Interaction (HRI 2017)*. Retrieved July 1, 2022, from https://pub.uni-bielefeld.de/record/2909062. (Cited on page 5)

Howley, I. K., & Rose, C. P. (2016). Towards careful practices for automated linguistic analysis of group learning. *Journal of Learning Analytics*, *3*(3), 239–262. https://doi.org/10.18608/jla.2016.33.12. (Cited on page 15)

Hsiao, H.-S., Chang, C.-S., Lin, C.-Y., & Hsu, H.-L. (2015). "iRobiQ": the influence of bidirectional interaction on kindergarteners' reading motivation, literacy, and behavior. *Interactive Learning Environments*, *23*(3), 269–292. https://doi.org/10.1080/10494820.2012.745435. (Cited on page 13)

Jacq, A., Geist, M., Paiva, A., & Pietquin, O. (2019). Learning from a learner. In K. Chaudhuri & R. Salakhutdinov (Eds.), *Proceedings of the 36th International Conference on Machine Learning (ICML 2019)* (pp. 2990–2999). PMLR. (Cited on page 20)

Jacq, A. D. (2020). *Mutual understanding in educational human-robot collaborations* (Doctoral dissertation). EPFL. Lausanne. https://doi.org/10.5075/epfl-thesis-10144. (Cited on page 104)

Jacq, A. D., Johal, W., Dillenbourg, P., & Paiva, A. (2016). Cognitive architecture for mutual modelling. *Proceedings of the 2nd Workshop on Cognitive Architectures for Social Human-Robot Interaction 2016 (CogArch4sHRI 2016)*. https://doi.org/10.48550/arXiv.1602.06703. (Cited on page 104)

Jarník, V. (1930). "O jistém problému minimálním" [About a certain minimal problem]. *Práce Moravské Přírodovědecké Společnosti*, *6*(4), 57–63. Retrieved October 3, 2022, from http://dml.cz/dmlcz/500726. (Cited on page 100)

Jarrett, D., Hüyük, A., & Schaar, M. V. D. (2021). Inverse decision modeling: learning interpretable representations of behavior. *Proceedings of the 38th International Conference on Machine Learning (ICML 2021)*, 4755–4771. (Cited on page 20)

Jensen, M. H. (2014). *Epistemic and doxastic planning* (Doctoral dissertation). Technical University of Denmark. Kgs. Lyngby. (Cited on page 21)

Jermann, P., Mullins, D., Nüssli, M.-A., & Dillenbourg, P. (2011). Collaborative gaze footprints: correlates of interaction quality. *Connecting Computer-Supported Collaborative Learning to Policy and Practice: CSCL2011 Conference Proceedings, Volume I - Long Papers*, 184–191. (Cited on page 14)

Jermann, P., & Nüssli, M.-A. (2012). Effects of sharing text selections on gaze cross-recurrence and interaction quality in a pair programming task. *CSCW'12: Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, 1125–1134. https://doi.org/10.1145/2145204.2145371. (Cited on pages 14, 16)

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: a survey. *Journal of Artificial Intelligence Research*, *4*, 237–285. https://doi.org/10.1613/jair.301. (Cited on page 21)

Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, *101*(1), 99–134. https://doi.org/10.1016/S0004-3702(98)00023-X. (Cited on pages 7, 20, 102)

Kalai, E., & Lehrer, E. (1993). Rational learning leads to Nash equilibrium. *Econometrica*, *61*(5), 1019–1045. https://doi.org/10.2307/2951492. (Cited on page 19)

Kapur, M. (2008). Productive failure. *Cognition and Instruction*, *26*(3), 379–424. https://doi.org/10.1080/07370000802212669. (Cited on pages 3, 77)

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)*. https://doi.org/10.1109/ARSO.2015.7428217. (Cited on page 12)

Kass, R., & Finin, T. (1988). Modeling the user in natural language systems. *Computational Linguistics*, *14*(3), 5–22. (Cited on page 18)

Kennedy, J., Baxter, P., Senft, E., & Belpaeme, T. (2015). Higher nonverbal immediacy leads to greater learning gains in child-robot tutoring interactions. In A. Tapus, E. André, J.-C. Martin, F. Ferland, & M. Ammi (Eds.), *Social Robotics* (pp. 327–336). Springer International Publishing. https://doi.org/10.1007/978-3-319-25554-5_33. (Cited on page 12)

King, A. (2007). Scripting collaborative learning processes: a cognitive perspective. In F. Fischer, I. Kollar, H. Mandl, & J. M. Haake (Eds.), *Scripting Computer-Supported Collaborative Learning: Cognitive, Computational and Educational Perspectives* (pp. 13–37). Springer US. https://doi.org/10.1007/978-0-387-36949-5_2. (Cited on page 2)

King, W. J., & Ohya, J. (1996). The representation of agents: anthropomorphism, agency, and intelligence. *CHI '96: Conference Companion on Human Factors in Computing Systems*, 289–290. https://doi.org/10.1145/257089.257326. (Cited on page 44)

Koiso, H., Horiuchi, Y., Tutiya, S., Ichikawa, A., & Den, Y. (1998). An analysis of turn-taking and backchannels based on prosodic and syntactic features in Japanese map task dialogs. *Language and Speech*, *41*(3-4), 295–321. (Cited on page 51)

Kopp, S., & Krämer, N. (2021). Revisiting human-agent communication: the importance of joint co-construction and understanding mental states. *Frontiers in Psychology*, *12*, 580955. https://doi.org/10.3389/fpsyg.2021.580955. (Cited on page 1)

Krämer, N. C., & Bente, G. (2010). Personalizing e-learning. the social effects of pedagogical agents. *Educational Psychology Review*, *22*(1), 71–87. https://doi.org/10.1007/s10648-010-9123-x. (Cited on page 12)

Kruskal, J. B. (1956). On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical Society*, *7*(1), 48–50. https://doi.org/10.1090/S0002-9939-1956-0078686-7. (Cited on page 100)

Kuhn, D. (2015). Thinking together and alone. *Educational Researcher*, *44*, 46–53. https://doi.org/10.3102/0013189X15569530. (Cited on page 3)

Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: a meta-analytic review. *Review of Educational Research*, *86*(1), 42–78. https://doi.org/10.3102/0034654315581420. (Cited on page 12)

Kumar, R., Ai, H., Beuth, J. L., & Rosé, C. P. (2010). Socially capable conversational tutors can be effective in collaborative learning situations. In V. Aleven, J. Kay, & J. Mostow (Eds.), *Intelligent Tutoring Systems* (pp. 156–164). Springer. https://doi.org/10.1007/978-3-642-13388-6_20. (Cited on page 13)

Le Grézause, E. (2017). *Um and uh, and the expression of stance in conversational speech* (Thesis). University of Washington. Retrieved August 3, 2022, from http://hdl.handle.net/1773/40619. (Cited on page 55)

Leite, I. M. d. S. C. (2013). *Long-term interactions with empathic social robots* (Doctoral dissertation). Instituto Superior Técnico. Lisboa, Portugal. (Cited on page 13)

Lemaignan, S., & Dillenbourg, P. (2015). Mutual modelling in robotics: inspirations for the next steps. *HRI '15: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 303–310. https://doi.org/10.1145/2696454.2696493. (Cited on pages 1, 3, 21)

Lemaignan, S., Jacq, A., Hood, D., Garcia, F., Paiva, A., & Dillenbourg, P. (2016). Learning by teaching a robot: the case of handwriting. *IEEE Robotics and Automation Magazine*, *23*(2), 56–66. https://doi.org/10.1109/MRA.2016.2546700. (Cited on page 12)

Lemaignan, S., Garcia, F., Jacq, A., & Dillenbourg, P. (2016). From real-time attention assessment to "with-me-ness" in human-robot interaction. *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 157–164. https://doi.org/10.1109/HRI.2016.7451747. (Cited on page 17)

Levitan, S., Maredia, A., & Hirschberg, J. (2018). Acoustic-prosodic indicators of deception and trust in interview dialogues. *19th Annual Conference of the International Speech Communication Association (INTERSPEECH 2018)*, 416–420. https://doi.org/10.21437/Interspeech.2018-2443. (Cited on page 17)

Levitan, S. I., Xiang, J., & Hirschberg, J. (2018). Acoustic-prosodic and lexical entrainment in deceptive dialogue. *Proceedings of the 9th International Conference on Speech Prosody*, 532–536. https://doi.org/10.21437/SpeechProsody.2018-108. (Cited on page 17)

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, *22 140*, 55–55. (Cited on page 34)

Litman, D. J., & Allen, J. F. (1984). A plan recognition model for clarification subdialogues. *Proceedings of the 10th International Conference on Computational Linguistics and 22nd Annual Meeting on Association for Computational Linguistics*, 302–311. https://doi.org/10.3115/980491.980554. (Cited on page 18)

Litman, D. J., Rosé, C. P., Forbes-Riley, K., VanLehn, K., Bhembe, D., & Silliman, S. (2004). Spoken versus typed human and computer dialogue tutoring. In J. C. Lester, R. M. Vicari, & F. Paraguaçu (Eds.), *Intelligent Tutoring Systems* (pp. 368–379). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-30139-4_35. (Cited on page 15)

Loberman, H., & Weinberger, A. (1957). Formal procedures for connecting terminals with a minimum total wire length. *Journal of the ACM*, *4*(4), 428–437. https://doi.org/10.1145/320893.320896. (Cited on page 100)

Louvan, S., & Magnini, B. (2020). Recent neural methods on slot filling and intent classification for task-oriented dialogue systems: a survey. *Proceedings of the 28th International Conference on Computational Linguistics*, 480–496. https://doi.org/10.18653/v1/2020.coling-main.42. (Cited on page 15)

Lu, C., Willi, T., Witt, C. A. S. D., & Foerster, J. (2022). Model-free opponent shaping. *Proceedings of the 39th International Conference on Machine Learning (ICML 2022)*, 14398–14411. (Cited on page 20)

Lubold, N., & Pon-Barry, H. (2014). Acoustic-prosodic entrainment and rapport in collaborative learning dialogues. *Proceedings of the 2014 ACM workshop on Multimodal Learning Analytics Workshop and Grand Challenge*, 5–12. https://doi.org/10.1145/2666633.2666635. (Cited on pages 16, 17)

Lubold, N., Pon-Barry, H., & Walker, E. (2015). Naturalness and rapport in a pitch adaptive learning companion.

2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU), 103–110. https://doi.org/10.1109/ASRU.2015.7404781. (Cited on page 16)

Lubold, N., Walker, E., Pon-Barry, H., & Ogan, A. (2018). Automated pitch convergence improves learning in a social, teachable robot for middle school mathematics. In C. Penstein Rosé, R. Martínez-Maldonado, H. U. Hoppe, R. Luckin, M. Mavrikis, K. Porayska-Pomsta, B. McLaren, & B. du Boulay (Eds.), *Artificial Intelligence in Education* (pp. 282–296). Springer International Publishing. https://doi.org/10.1007/978-3-319-93843-1_21. (Cited on pages 13, 16)

Marchesi, S., Ghiglino, D., Ciardo, F., Perez-Osorio, J., Baykara, E., & Wykowska, A. (2019). Do we adopt the intentional stance toward humanoid robots? *Frontiers in Psychology*, *10*, 450. https://doi.org/10.3389/fpsyg.2019.00450. (Cited on page 6)

Marge, M., Espy-Wilson, C., Ward, N. G., Alwan, A., Artzi, Y., Bansal, M., Blankenship, G., Chai, J., Daumé, H., Dey, D., Harper, M., Howard, T., Kennington, C., Kruijff-Korbayová, I., Manocha, D., Matuszek, C., Mead, R., Mooney, R., Moore, R. K., . . . Yu, Z. (2022). Spoken language interaction with robots: recommendations for future research. *Computer Speech & Language*, *71*, 101255. https://doi.org/10.1016/j.csl.2021.101255. (Cited on page 5)

Matarić, M. J. (2017). Socially assistive robotics: human augmentation versus automation. *Science Robotics*, *2*(4), eaam5410. https://doi.org/10.1126/scirobotics.aam5410. (Cited on page 13)

Matarić, M. J., Eriksson, J., Feil-Seifer, D. J., & Winstein, C. J. (2007). Socially assistive robotics for post-stroke rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, *4*(1), 5. https://doi.org/10.1186/1743-0003-4-5. (Cited on page 13)

Matarić, M. J., & Scassellati, B. (2016). Socially assistive robotics. In B. Siciliano & O. Khatib (Eds.), *Springer Handbook of Robotics* (pp. 1973–1994). Springer International Publishing. https://doi.org/10.1007/978-3-319-32552-1_73. (Cited on page 13)

Mazzoni, E., & Benvenuti, M. (2015). A robot-partner for preschool children learning english using socio-cognitive conflict. *Journal of Educational Technology & Society*, *18*(4), 474–485. (Cited on page 13)

McCalla, G., Vassileva, J., Greer, J., & Bull, S. (2000). Active learner modelling. In G. Gauthier, C. Frasson, & K. VanLehn (Eds.), *Intelligent Tutoring Systems* (pp. 53–62). Springer. https://doi.org/10.1007/3-540-45108-0_9. (Cited on page 18)

McKelvey, R. D., & Palfrey, T. R. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior*, *10*(1), 6–38. https://doi.org/10.1006/game.1995.1023. (Cited on page 21)

Mead, G. H. (1934). *Mind, self, and society: from the standpoint of a social behaviorist* (C. W. Morris, Ed.; 1st ed.). University of Chicago Press. (Cited on page 2)

Menon, D., Sowmya, B. P., Romero, M., & Viéville, T. (2020). Going beyond digital literacy to develop computational thinking in K-12 education. In L. Daniela (Ed.), *Epistemological Approaches to Digital Learning in Educational Contexts* (1st ed., p. 18). Routledge. https://doi.org/10.4324/9780429319501-2. (Cited on page 28)

Mertens, J.-F., & Zamir, S. (1985). Formulation of Bayesian analysis for games with incomplete information. *International Journal of Game Theory*, *14*(1), 1–29. https://doi.org/10.1007/BF01770224. (Cited on page 19)

Milliez, G., Warnier, M., Clodic, A., & Alami, R. (2014). A framework for endowing an interactive robot with reasoning capabilities about perspective-taking and belief management. *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, 1103–1109. https://doi.org/10.1109/ROMAN.2014.6926399. (Cited on page 21)

Moro, C., Lin, S., Nejat, G., & Mihailidis, A. (2019). Social robots and seniors: a comparative study on the influence of dynamic social features on human–robot interaction. *International Journal of Social Robotics*, *11*(1), 5–24. https://doi.org/10.1007/s12369-018-0488-1. (Cited on page 44)

Morris, A., Maier, V., & Green, P. (2004). From WER and RIL to MER and WIL: improved evaluation measures for connected speech recognition. *NTERSPEECH 2004 - ICSLP, 8th International Conference on Spoken Language Processing*. https://doi.org/10.21437/Interspeech.2004-668. (Cited on page 53)

Mu, J., Stegmann, K., Mayfield, E., Rosé, C., & Fischer, F. (2012). The ACODEA framework: developing segmentation and classification schemes for fully automatic analysis of online discussions. *International Journal of*

Computer-Supported Collaborative Learning, *7*, 285–305. https://doi.org/10.1007/s11412-012-9147-y. (Cited on page 14)

Mubin, O., Stevens, C. J., Shahid, S., Mahmud, A. A., & Dong, J.-J. (2013). A review of the applicability of robots in education. *Technology for Education and Learning*, *1*, 1–7. https://doi.org/10.2316/Journal.209.2013.1.209-0015. (Cited on page 12)

Mugny, G., & Doise, W. (1978). Socio-cognitive conflict and structure of individual and collective performances. *European Journal of Social Psychology*, *8*(2), 181–192. https://doi.org/10.1002/ejsp.2420080204. (Cited on page 2)

Nash, J. F. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, *36*(1), 48–49. https://doi.org/10.1073/pnas.36.1.48. (Cited on page 19)

Nasir*, J., **Norman***, **U.**, Bruno, B., & Dillenbourg, P. (2020). When positive perception of the robot has no effect on learning. *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 313–320. https://doi.org/10.1109/RO-MAN47096.2020.9223343. **Contributed equally to this work.* (Cited on pages 8, 25)

Nasir, J. (2022). *Introducing productive engagement for social robots supporting learning* (Doctoral dissertation). EPFL. Lausanne. https://doi.org/10.5075/epfl-thesis-9781. (Cited on page 84)

Nasir, J., Bruno, B., Chetouani, M., & Dillenbourg, P. (2022a). A speech-based productive engagement metric for real-time human-robot interaction in collaborative educational contexts. *IEEE Transactions on Affective Computing*. Retrieved July 19, 2022, from https://infoscience.epfl.ch/record/294035. (Cited on pages 26, 40, 72)

Nasir, J., Bruno, B., Chetouani, M., & Dillenbourg, P. (2022b). What if social robots look for productive engagement? *International Journal of Social Robotics*, *14*, 55–71. https://doi.org/10.1007/s12369-021-00766-w. (Cited on pages 26, 40, 68, 72, 77)

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. https://doi.org/10.1007/s11412-021-09358-2. (Cited on pages 26, 40, 72, 82)

Ndousse, K. K., Eck, D., Levine, S., & Jaques, N. (2021). Emergent social learning via multi-agent reinforcement learning. *Proceedings of the 38th International Conference on Machine Learning (ICML 2021)*, 7991–8004. (Cited on page 20)

Nederhof, A. J. (1985). Methods of coping with social desirability bias: a review. *European Journal of Social Psychology*, *15*(3), 263–280. https://doi.org/10.1002/ejsp.2420150303. (Cited on page 44)

Nenkova, A., Gravano, A., & Hirschberg, J. (2008). High frequency word entrainment in spoken dialogue. *Proceedings of ACL-08: HLT, Short Papers*, 169–172. (Cited on page 17)

Nešetřil, J., Milková, E., & Nešetřilová, H. (2001). Otakar Borůvka on minimum spanning tree problem: translation of both the 1926 papers, comments, history. *Discrete Mathematics*, *233*(1), 3–36. https://doi.org/10.1016/S0012-365X(00)00224-7. (Cited on page 101)

Neumann, M. M. (2020). Social robots and young children's early language and literacy learning. *Early Childhood Education Journal*, *48*(2), 157–170. https://doi.org/10.1007/s10643-019-00997-7. (Cited on page 13)

Ng, A. Y., & Russell, S. J. (2000). Algorithms for inverse reinforcement learning. *Proceedings of the Seventeenth International Conference on Machine Learning (ICML 2000)*, 663–670. (Cited on page 20)

Norman, **U.**, Bruno, B., & Dillenbourg, P. (2021). Mutual modelling ability for a humanoid robot: How can it improve my learning as we solve a problem together? *Robots for Learning Workshop in 16th Annual IEEE/ACM Conference on Human-Robot Interaction (HRI 2021)*. Retrieved July 1, 2022, from http://infoscience.epfl.ch/record/283614. (Cited on page 8)

Norman, **U.**, Chin, A., Bruno, B., & Dillenbourg, P. (2022). Efficacy of a 'misconceiving' robot to improve computational thinking in a collaborative problem solving activity: a pilot study. *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 1413–1420. https://doi.org/10.1109/RO-MAN53752.2022.9900775. (Cited on pages 8, 76)

Norman*, **U.**, Dinkar*, T., Bruno, B., & Clavel, C. (2022). Studying alignment in a collaborative learning activity via automatic methods: the link between what we say and do. *Dialogue & Discourse*, *13*(2), 1–48. https://doi.org/10.5210/dad.2022.201. **Contributed equally to this work.* (Cited on pages 8, 48, 61, 62)

Oertel, C., Castellano, G., Chetouani, M., Nasir, J., Obaid, M., Pelachaud, C., & Peters, C. (2020). Engagement in human-agent interaction: an overview. *Frontiers in Robotics and AI*, *7*, 92. https://doi.org/10.3389/frobt.2020.00092. (Cited on page 16)

Okita, S. Y., Ng-Thow-Hing, V., & Sarvadevabhatla, R. (2009). Learning together: ASIMO developing an interactive learning partnership with children. *RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication*, *4*, 1125–1130. https://doi.org/10.1109/ROMAN.2009.5326135. (Cited on pages 12, 13)

Papadopoulos, I., Lazzarino, R., Miah, S., Weaver, T., Thomas, B., & Koulouglioti, C. (2020). A systematic review of socially assistive robots in pre-tertiary education. *Computers & Education*, *155*, 103924. https://doi.org/10.1016/j.compedu.2020.103924. (Cited on page 13)

Papakostas, G. A., Sidiropoulos, G. K., Papadopoulou, C. I., Vrochidou, E., Kaburlasos, V. G., Papadopoulou, M. T., Holeva, V., Nikopoulou, V.-A., & Dalivigkas, N. (2021). Social robots in special education: a systematic review. *Electronics*, *10*(12), 1398. https://doi.org/10.3390/electronics10121398. (Cited on page 13)

Park, Y., Patwardhan, S., Visweswariah, K., & Gates, S. C. (2008). An empirical analysis of word error rate and keyword error rate. *Proceedings of the 9th Annual Conference of the International Speech Communicatiuon Association (INTERSPEECH 2008)*, 2070–2073. https://doi.org/10.21437/Interspeech.2008-537. (Cited on page 53)

Perez-Osorio, J., & Wykowska, A. (2019). Adopting the intentional stance towards humanoid robots. In J.-P. Laumond, E. Danblon, & C. Pieters (Eds.), *Wording Robotics: Discourses and Representations on Robotics* (pp. 119–136). Springer International Publishing. https://doi.org/10.1007/978-3-030-17974-8_10. (Cited on page 1)

Perez-Osorio, J., & Wykowska, A. (2020). Adopting the intentional stance toward natural and artificial agents. *Philosophical Psychology*, *33*(3), 369–395. https://doi.org/10.1080/09515089.2019.1688778. (Cited on page 6)

Perkins, D., & Salomon, G. (1992). Transfer of learning. In T. Husen & N. Postlethwaite (Eds.), *International Encyclopedia of Education* (2nd ed., pp. 6452–6457). Pergamon Press. (Cited on pages 43, 79)

Piaget, J. (1971). *L'epistémologie génétique* [Genetic epistemology] (E. Duckworth, Trans.). W. W. Norton & Company, Inc. (Original work published 1970). (Cited on page 2)

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(2), 169–225. https://doi.org/10.1017/s0140525x04000056. (Cited on pages 3, 4, 18, 47, 54, 136)

Pickering, M. J., & Garrod, S. (2006). Alignment as the basis for successful communication. *Research on Language and Computation*, *4*(2), 203–228. https://doi.org/10.1007/s11168-006-9004-0. (Cited on pages 3, 4, 47, 136)

Popper, K. (2002). *"logik der forschung"* [The logic of scientific discovery] (2nd ed.). Routledge. (Original work published 1935). (Cited on page 101)

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526. https://doi.org/10.1017/S0140525X00076512. (Cited on page 1)

Prim, R. C. (1957). Shortest connection networks and some generalizations. *The Bell System Technical Journal*, *36*(6), 1389–1401. https://doi.org/10.1002/j.1538-7305.1957.tb01515.x. (Cited on page 100)

Pu, L., Moyle, W., Jones, C., & Todorovic, M. (2019). The effectiveness of social robots for older adults: a systematic review and meta-analysis of randomized controlled studies. *The Gerontologist*, *59*(1), e37–e51. https://doi.org/10.1093/geront/gny046. (Cited on page 13)

Rahimi, Z., Kumar, A., Litman, D., Paletz, S., & Yu, M. (2017). Entrainment in multi-party spoken dialogues at multiple linguistic levels. *Proc. Interspeech 2017*, 1696–1700. https://doi.org/10.21437/Interspeech.2017-1568. (Cited on pages 17, 18)

Ramachandran, A., Litoiu, A., & Scassellati, B. (2016). Shaping productive help-seeking behavior during robot-child tutoring interactions. *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 247–254. https://doi.org/10.1109/HRI.2016.7451759. (Cited on page 12)

Ramponi, G., Drappo, G., & Restelli, M. (2020). Inverse reinforcement learning from a gradient-based learner. *Proceedings of the 33rd International Conference on Neural Information Processing Systems (NeurIPS 2020)*, *33*, 2458–2468. (Cited on page 20)

Rasouli, A., Lanillos, P., Cheng, G., & Tsotsos, J. K. (2020). Attention-based active visual search for mobile robots. *Autonomous Robots*, *44*(2), 131–146. https://doi.org/10.1007/s10514-019-09882-z. (Cited on page 18)

Reitter, D., & Moore, J. D. (2014). Alignment and task success in spoken dialogue. *Journal of Memory and Language*, *76*, 29–46. https://doi.org/10.1016/j.jml.2014.05.008. (Cited on page 16)

Richman, W. L., Kiesler, S., Weisband, S., & Drasgow, F. (1999). A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology*, *84*(5), 754–775. https://doi.org/10.1037/0021-9010.84.5.754. (Cited on page 44)

Robinson, H., MacDonald, B., & Broadbent, E. (2014). The role of healthcare robots for older people at home: a review. *International Journal of Social Robotics*, *6*(4), 575–591. https://doi.org/10.1007/s12369-014-0242-2. (Cited on page 13)

Robinson, J. (1951). An iterative method of solving a game. *Annals of Mathematics*, *54*(2), 296–301. https://doi.org/10.2307/1969530. (Cited on page 19)

Rogoff, B. (1990). *Apprenticeship in thinking: cognitive development in social context* (1st ed.). Oxford University Press. (Cited on page 2)

Rohlfing, K. J., Altvater-Mackensen, N., Caruana, N., van den Berghe, R., Bruno, B., Tolksdorf, N. F., & Hanuliková, A. (2022). Dialogical roles of social robots for supporting children's learning of language and literacy—a review and analysis of innovative roles [submitted]. (Cited on page 12)

Romano, J., Kromrey, J. D., Coraggio, J., Skowronek, J., & Devine, L. (2006). Exploring methods for evaluating group differences on the NSSE and other surveys: Are the t-test and Cohen's d indices the most appropriate choices? *Annual Meeting of the Southern Association for Institutional Research*, 1–51. (Cited on pages 43, 89)

Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving (C. O'Malley, Ed.). *Computer Supported Collaborative Learning*, *128*, 69–97. https://doi.org/10.1007/978-3-642-85098-1_5. (Cited on page 2)

Rosé, C., Wang, Y.-C., Cui, Y., Arguello, J., Stegmann, K., Weinberger, A., & Fischer, F. (2008). Analyzing collaborative learning processes automatically: exploiting the advances of computational linguistics in computer-supported collaborative learning. *International Journal of Computer-Supported Collaborative Learning*, *3*(3), 237–271. https://doi.org/10.1007/s11412-007-9034-0. (Cited on pages 14, 15)

Rubin, J., & Watson, I. (2011). Computer poker: a review. *Artificial Intelligence*, *175*(5-6), 958–987. https://doi.org/10.1016/j.artint.2010.12.005. (Cited on page 19)

Rueben, M., Rothberg, E., Tang, M., Inzerillo, S., Kshirsagar, S. S., Manchanda, M., Dudley, G., Fraune, M. R., & Matarić, M. J. (2022). The robot olympics: estimating and influencing beliefs about a robot's perceptual capabilities. *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 1405–1412. https://doi.org/10.1109/RO-MAN53752.2022.9900796. (Cited on page 106)

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), 68–78. https://doi.org/10.1037/0003-066X.55.1.68. (Cited on page 34)

Sadigh, D., Landolfi, N., Sastry, S. S., Seshia, S. A., & Dragan, A. D. (2018). Planning for cars that coordinate with people: leveraging effects on human actions for planning and active information gathering over human internal state. *Autonomous Robots*, *42*(7), 1405–1426. https://doi.org/10.1007/s10514-018-9746-1. (Cited on page 18)

Sallami, Y., Lemaignan, S., Clodic, A., & Alami, R. (2019). Simulation-based physics reasoning for consistent scene estimation in an HRI context [ISSN: 2153-0866]. *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 7834–7841. https://doi.org/10.1109/IROS40897.2019.8968106. (Cited on page 21)

Sandholm, T. (2010). The state of solving large incomplete-information games, and application to poker. *AI Magazine*, *31*(4), 13–32. https://doi.org/10.1609/aimag.v31i4.2311. (Cited on page 19)

Sangin, M., Molinari, G., Nüssli, M.-A., & Dillenbourg, P. (2011). Facilitating peer knowledge modeling: Effects of a knowledge awareness tool on collaborative learning outcomes and processes. *Computers in Human Behavior*, *27*(3), 1059–1067. https://doi.org/10.1016/j.chb.2010.05.032. (Cited on page 41)

Sarthou, G., Clodic, A., & Alami, R. (2019). Ontologenius: a long-term semantic memory for robotic agents [ISSN:

1944-9437]. *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 1–8. https://doi.org/10.1109/RO-MAN46459.2019.8956305. (Cited on page 21)

Scassellati, B. M. (2002). Theory of mind for a humanoid robot. *Autonomous Robots*, *12*(1), 13–24. https://doi.org/10.1023/A:1013298507114. (Cited on page 1)

Schellen, E., & Wykowska, A. (2019). Intentional mindset toward robots—open questions and methodological challenges. *Frontiers in Robotics and AI*, *5*, 139. https://doi.org/10.3389/frobt.2018.00139. (Cited on page 6)

Scheuer, O., Loll, F., Pinkwart, N., & McLaren, B. M. (2010). Computer-supported argumentation: a review of the state of the art. *International Journal of Computer-Supported Collaborative Learning*, *5*(1), 43–102. https://doi.org/10.1007/s11412-009-9080-x. (Cited on page 15)

Schneider, B., Worsley, M., & Martinez-Maldonado, R. (2021). Gesture and gaze: multimodal data in dyadic interactions. In U. Cress, C. Rosé, A. F. Wise, & J. Oshima (Eds.), *International Handbook of Computer-Supported Collaborative Learning* (pp. 625–641). Springer International Publishing. https://doi.org/10.1007/978-3-030-65291-3_34. (Cited on page 14)

Schroeder, N. L., Adesope, O. O., & Gilbert, R. B. (2013). How effective are pedagogical agents for learning? A meta-analytic review. *Journal of Educational Computing Research*, *49*(1), 1–39. https://doi.org/10.2190/EC.49.1.a. (Cited on page 12)

Schwartz, D. L. (1995). The emergence of abstract representations in dyad problem solving. *Journal of the Learning Sciences*, *4*(3), 321–354. https://doi.org/10.1207/s15327809jls0403_3. (Cited on page 2)

Searle, J. R. (1983). *Intentionality: an essay in the philosophy of mind*. Cambridge University Press. (Cited on page 6)

Sharma, K., Jermann, P., & Dillenbourg, P. (Eds.). (2014). "With-me-ness": a gaze-measure for students' attention in MOOCs. In *Proceedings of 11th International Conference of the Learning Sciences (ICLS)*. (Cited on page 16)

Shoham, Y. (1993). Agent-oriented programming. *Artificial Intelligence*, *60*(1), 51–92. https://doi.org/10.1016/0004-3702(93)90034-9. (Cited on page 6)

Silver, D., & Veness, J. (2010). Monte-Carlo planning in large POMDPs. In J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, & A. Culotta (Eds.), *Advances in Neural Information Processing Systems 23* (pp. 2164–2172). Curran Associates, Inc. (Cited on page 106)

Sinclair, A. J., & Schneider, B. (2021). Linguistic and gestural coordination: do learners converge in collaborative dialogue? *Proceedings of the 14th International Conference on Educational Data Mining (EDM 2021)*, 431–437. (Cited on page 16)

Sinha, T., & Cassell, J. (2015). We click, we align, we learn: impact of influence and convergence processes on student learning and rapport building. *Proceedings of the 1st Workshop on Modeling INTERPERsonal SynchrONy And infLuence*, 13–20. https://doi.org/10.1145/2823513.2823516. (Cited on pages 16, 17)

Strijbos, J.-W., Martens, R. L., Prins, F. J., & Jochems, W. M. G. (2006). Content analysis: What are they talking about? *Computers & Education*, *46*(1), 29–48. https://doi.org/10.1016/j.compedu.2005.04.002. (Cited on page 14)

Szymański, P., Żelasko, P., Morzy, M., Szymczak, A., Ży\la-Hoppe, M., Banaszczak, J., Augustyniak, L., Mizgajski, J., & Carmiel, Y. (2020). WER we are and WER we think we are. *Findings of the Association for Computational Linguistics: EMNLP 2020*, 3290–3295. https://doi.org/10.18653/v1/2020.findings-emnlp.295. (Cited on page 53)

Tabrez, A., Luebbers, M. B., & Hayes, B. (2020). A survey of mental modeling techniques in human–robot teaming. *Current Robotics Reports*, *1*(4), 259–267. https://doi.org/10.1007/s43154-020-00019-0. (Cited on page 18)

Tanaka, F., & Matsuzoe, S. (2012). Children teach a care-receiving robot to promote their learning: field experiments in a classroom for vocabulary learning. *Journal of Human-Robot Interaction*, *31*(1), 78–96. https://doi.org/10.5898/JHRI.1.1.Tanaka. (Cited on page 12)

Tchounikine, P., Rummel, N., & McLaren, B. M. (2010). Computer supported collaborative learning and intelligent tutoring systems. In R. Nkambou, J. Bourdeau, & R. Mizoguchi (Eds.), *Advances in Intelligent Tutoring Systems* (pp. 447–463). Springer. https://doi.org/10.1007/978-3-642-14363-2_22. (Cited on page 15)

Thomason, J., Nguyen, H. V., & Litman, D. (2013). Prosodic entrainment and tutoring dialogue success. In H. C.

Lane, K. Yacef, J. Mostow, & P. Pavlik (Eds.), *Artificial Intelligence in Education* (pp. 750–753). Springer. https://doi.org/10.1007/978-3-642-39112-5_104. (Cited on page 16)

Traum, D. R. (1999). Computational models of grounding in collaborative systems. *Proceedings of the AAAI Fall Symposium on Psychological Models of Communication*, 124–131. (Cited on page 4)

Traum, D. R., & Allen, J. F. (1992). A "speech acts" approach to grounding in conversation. *Proceedings of International Conference on Spoken Language Processing (ICSLP'92)*, 137–140. (Cited on page 4)

Traum, D. R., & Hinkelman, E. A. (1992). Conversation acts in task-oriented spoken dialogue. *Computational Intelligence*, *8*(3), 575–599. https://doi.org/10.1111/j.1467-8640.1992.tb00380.x. (Cited on page 4)

Trausan-Matu, S., & Slotta, J. D. (2021). Artifact analysis. In U. Cress, C. Rosé, A. F. Wise, & J. Oshima (Eds.), *International Handbook of Computer-Supported Collaborative Learning* (pp. 551–567). Springer International Publishing. https://doi.org/10.1007/978-3-030-65291-3_30. (Cited on page 14)

Tur, G., & De Mori, R. (2011). *Spoken language understanding: systems for extracting semantic information from speech*. John Wiley & Sons. (Cited on page 15)

van den Berghe, R., Verhagen, J., Oudgenoeg-Paz, O., van der Ven, S., & Leseman, P. (2019). Social robots for language learning: a review. *Review of Educational Research*, *89*(2), 259–295. https://doi.org/10.3102/0034654318821286. (Cited on page 13)

van Ditmarsch, H., van Der Hoek, W., & Kooi, B. (2007). *Dynamic epistemic logic* (Vol. 337). Springer Science & Business Media. (Cited on page 21)

Verbrugge, R. (2009). Logic and social cognition. *Journal of Philosophical Logic*, *38*(6), 649–680. https://doi.org/10.1007/s10992-009-9115-9. (Cited on page 1)

Vogt, P., van den Berghe, R., de Haas, M., Hoffman, L., Kanero, J., Mamus, E., Montanier, J.-M., Oranc, C., Oudgenoeg-Paz, O., Garcia, D. H., Papadopoulos, F., Schodde, T., Verhagen, J., Wallbridgell, C. D., Willemsen, B., de Wit, J., Belpaeme, T., Goksun, T., Kopp, S., … Pandey, A. K. (2019). Second language tutoring using social robots: a large-scale study. *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 497–505. https://doi.org/10.1109/HRI.2019.8673077. (Cited on page 12)

Vygotsky, L. S. (1978). *Mind in society: the development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press. https://doi.org/10.2307/j.ctvjf9vz4 (Original work published 1930–1934). (Cited on page 2)

Vygotsky, L. S. (1986). *Myshlenie i rech* [Thought and language] (A. Kozulin, Trans.). The MIT Press. (Original work published 1934). (Cited on pages 2, 77)

Walker, E., Rummel, N., & Koedinger, K. R. (2011a). Designing automated adaptive support to improve student helping behaviors in a peer tutoring activity. *International Journal of Computer-Supported Collaborative Learning*, *6*(2), 279–306. https://doi.org/10.1007/s11412-011-9111-2. (Cited on page 15)

Walker, E., Rummel, N., & Koedinger, K. R. (2011b). Using automated dialog analysis to assess peer tutoring and trigger effective support. In G. Biswas, S. Bull, J. Kay, & A. Mitrovic (Eds.), *Artificial Intelligence in Education* (pp. 385–393). Springer. https://doi.org/10.1007/978-3-642-21869-9_50. (Cited on page 15)

Wandzel, A., Oh, Y., Fishman, M., Kumar, N., Wong, L. L., & Tellex, S. (2019). Multi-object search using object-oriented POMDPs. *2019 International Conference on Robotics and Automation (ICRA)*, 7194–7200. https://doi.org/10.1109/ICRA.2019.8793888. (Cited on page 106)

Wang, C., & Schmolze, J. (2005). Planning with POMDPs using a compact, logic-based representation. *17th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'05)*. https://doi.org/10.1109/ICTAI.2005.96. (Cited on page 22)

Wang, Y.-Y., Acero, A., & Chelba, C. (2003). Is word error rate a good indicator for spoken language understanding accuracy. *IEEE Workshop on Automatic Speech Recognition and Understanding*, 577–582. https://doi.org/10.1109/ASRU.2003.1318504. (Cited on page 53)

Ward, A., & Litman, D. (2007a). Automatically measuring lexical and acoustic/prosodic convergence in tutorial dialog corpora. *SLaTE Workshop on Speech and Language Technology in Education*. Retrieved June 25, 2022, from http://d-scholarship.pitt.edu/id/eprint/23210. (Cited on page 16)

Ward, A., & Litman, D. (2007b). Dialog convergence and learning. *Proceedings of the 2007 Conference on Artificial Intelligence in Education: Building Technology Rich Learning Contexts That Work*, 262–269. (Cited on page 16)

Watkins, C. J. C. H., & Dayan, P. (1992). Q-learning. *Machine Learning*, *8*(3), 279–292. https://doi.org/10.1007/BF00992698. (Cited on page 101)

Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & Education*, *46*(1), 71–95. https://doi.org/10.1016/j.compedu.2005.04.003. (Cited on page 14)

Werfel, J. (2013). Embodied teachable agents: learning by teaching robots. *Proceedings of the 13th International Conference on Intelligent Autonomous Systems*. Retrieved July 15, 2022, from http://people.seas.harvard.edu/~jkwerfel/nrfias14.pdf. (Cited on page 12)

Wiggins, G. P., & McTighe, J. (2005). *Understanding by design* (2nd ed.). Association for Supervision; Curriculum Development. (Cited on page 28)

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103–128. https://doi.org/10.1016/0010-0277(83)90004-5. (Cited on page 1)

Wooldridge, M., & Jennings, N. R. (1995). Intelligent agents: theory and practice. *The Knowledge Engineering Review*, *10*(2), 115–152. https://doi.org/10.1017/S0269888900008122. (Cited on page 98)

Zambak, A. F. (2009). *Can a machine think? A methodological approach to artificial intelligence* (Doctoral dissertation). Katholieke Universiteit Leuven. (Cited on pages 18, 98)

Zayats, V., Tran, T., Wright, R. A., Mansfield, C., & Ostendorf, M. (2019). Disfluencies and human speech transcription errors. *INTERSPEECH*. https://doi.org/10.21437/interspeech.2019-3134. (Cited on page 55)

Zhao, T., & Kawahara, T. (2019). Joint dialog act segmentation and recognition in human conversations using attention to dialog context. *Computer Speech & Language*, *57*, 108–127. https://doi.org/10.1016/j.csl.2019.03.001. (Cited on page 16)

Zheng, K., & Tellex, S. (2020). Pomdp_py: a framework to build and solve POMDP problems. *ICAPS 2020 Workshop on Planning and Robotics (PlanRob)*. https://doi.org/10.48550/arXiv.2004.10099. (Cited on page 97)

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162–185. https://doi.org/10.1037/0033-2909.123.2.162. (Cited on page 4)

Utku Norman

norman.utku@gmail.com ◇ https://orcid.org/0000-0002-6802-1444
EPFL IC IINFCOM CHILI ◇ RLC D1 740 (Rolex Learning Center) ◇ Station 20 ◇ Lausanne, VD CH-1015 Switzerland

EDUCATION

Ph.D. in Computer & Communication Sciences September 2018 – present
EPFL (Lausanne, Switzerland)

· As a Marie Skłodowska-Curie Fellow, developed computational models to equip social robots with mutual understanding skills to support collaborative learning, evaluated the effects via user studies at schools. Advised by Prof. Pierre Dillenbourg.

M.Sc. in Computer Engineering September 2016 – July 2018
Bilkent University (Ankara, Turkey)

· Developed a spatio-temporal data mining algorithm to predict new risk genes for autism spectrum disorder (ASD), through dynamic functional clustering of genes in gene expression networks. Advised by Assoc. Prof. A. Ercüment Çiçek.

B.Sc. in Electrical and Electronics Engineering *(Specialization: Computers)* September 2011 – June 2016
Middle East Technical University (Ankara, Turkey)

· Co-lead a team of five to build a collaborative maze-solving robot for Capstone Design. Advised by Assoc. Prof. Elif Vural.
· Co-created mobile robots to navigate an augmented space for Mechatronics Design. Advised by Assoc. Prof. Buğra Koku.

B.A. in Philosophy *(Double Major)* September 2012 – June 2018
Middle East Technical University (Ankara, Turkey)

Erasmus+ Exchange Student August 2014 – January 2015
Technical University of Denmark (Copenhagen, Denmark)

PROFESSIONAL EXPERIENCE

Doctoral Assistant June 2018 – December 2022
Computer-Human Interaction Lab for Learning & Instruction (CHILI), EPFL (Lausanne, Switzerland)

· Introduced a mutual modeling framework for artificial agents to recursively build a mental model of a human partner
· Developed AI / ML algorithms to assess structures of alignment in dialogue and action data with NLP, NLU, Cloud APIs
· Co-lead the JUSThink Project (https://go.epfl.ch/justhink) to build intelligent autonomous social robots in order to develop children's computational thinking skills
· Co-directed HRI studies with 250+ children at 8 schools; published open datasets, data analysis results, and software tools
· Created a robot-mediated collaborative activity using Python and ROS, that elicits dialogue on how humans collaborate
· Co-supervised a summer intern that lead to [1], and 6 semester project students to develop robot skills, leading to [5]
· Assisted courses: Robotics Practicals (Spring '21), Visual Computing (Spring '19, '20) and Linear Algebra (Fall '20, '21)

Visiting Researcher November 2019 – December 2019
Télécom Paris, hosted at Sorbonne University (Paris, France)

· Designed algorithms to analyze linguistic alignment in dialogue and action data, collaborating with Prof. Chloé Clavel [2]

Graduate Research and Teaching Assistant September 2016 – July 2018
CICEKLAB, Bilkent University (Ankara, Turkey)

· Developed ML algorithms for spatio-temporal clustering by designing network-based models of gene expression data
· Assisted: Advanced Algorithms (Spring '18), Digital Design (Fall '16, '17), Programming (Spring '16), AI Summer School ('17)

Summer Research Intern August 2016 – September 2016
Neuroscience Lab, Hacettepe University (Ankara, Turkey)

· Created an image processing tool to evaluate biological activity in image data in MATLAB, lab lead by Prof. Turgay Dalkara

Summer Engineering Intern August 2015 – September 2015
ASELSAN (Ankara, Turkey)

· Prepared a knowledge representation and reasoning tool to integrate business rule logic as validation constraints in Java

Research Assistant June 2014 – August 2014
Artificial Intelligence Lab, Koç University (Istanbul, Turkey)

- Worked on data augmentation to generate natural language navigation instructions in Java, lab lead by Prof. Deniz Yüret

Summer Engineering Intern June 2013 – July 2013
UDEA Wireless Technologies (Ankara, Turkey)

- Implemented a time-series data acquisition system with real-time visualization in Visual C# on PC and C on TI MSP430

HONORS AND AWARDS

Marie Skłodowska-Curie Research Fellowship September 2018 – June 2022

- Early Stage Researcher as part of the EU's Horizon 2020 ANIMATAS Project (MSCA – ITN – 2017 – 765955)
- Co-organized the ANIMATAS Symposium (https://sites.google.com/view/animatas-symposium) on human-machine interaction: perception, social learning, and adaptation in educational settings; and the EPFL workshop on educational technologies
- Participated in 6 technical / transferable-skills workshops in academia & industry in 4 countries within this training network

TÜBİTAK Undergraduate Scholarship (#2205) February 2015 – June 2016

PUBLICATIONS

[1] **U. Norman**, A. Chin, B. Bruno, and P. Dillenbourg, "Efficacy of a 'misconceiving' robot to improve computational thinking in a collaborative problem solving activity: a pilot study," in *2022 31st IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*, Naples, Italy, Aug. 2022, pp. 1413–1420. **Finalist for Best Student Paper Award**. doi: 10.1109/RO-MAN53752.2022.9900775

Made the code and analysis publicly available on GitHub: utku-norman/justhink-preexp-analysis
and the tools to represent and visualize the activities: utku-norman/justhink_world
and the tools to govern the interaction with the robot: utku-norman/justhink-ros

[2] **U. Norman***, T. Dinkar*, B. Bruno, and C. Clavel, "Studying alignment in a collaborative learning activity via automatic methods: the link between what we say and do," in *Dialogue & Discourse*, vol. 13, no. 2, pp. 1–48, Aug. 2022. doi: 10.5210/dad.2022.201

Made the dataset & tools publicly available at doi: 10.5281/zenodo.4627104 and doi: 10.5281/zenodo.6974562

[3] **U. Norman**, B. Bruno, and P. Dillenbourg, "Mutual modelling ability for a humanoid robot: How can it improve my learning as we solve a problem together?," in *Robots for Learning Workshop in 16th annual IEEE/ACM Conference on Human-Robot Interaction (HRI 2021)*, Online, Mar. 2021. Available: http://infoscience.epfl.ch/record/283614

[4] J. Nasir*, **U. Norman***, B. Bruno, and P. Dillenbourg, "When positive perception of the robot has no effect on learning," in *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, Online, Aug. 2020, pp. 313–320. doi: 10.1109/RO-MAN47096.2020.9223343

[5] R. Maure*, E. A. Wengle*, **U. Norman**, D. C. Tozadore, and B. Bruno, "A case for the design of attention and gesture systems for social robots," in *Social Robotics*, Florence, Italy, Dec. 2022.

[6] J. Nasir*, **U. Norman***, W. Johal, J. Olsen, S. Shahmoradi, and P. Dillenbourg, "Robot analytics: What do human-robot interaction traces tell us about learning?," in *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, New Delhi, India, Oct. 2019, doi: 10.1109/RO-MAN46459.2019.8956465

[7] S. Shahmoradi, J. K. Olsen, S. Haklev, W. Johal, **U. Norman**, J. Nasir, and P. Dillenbourg, "Orchestration of robotic activities in classrooms: challenges and opportunities," in *EC-TEL 2019: Transforming Learning with Meaningful Technologies*, pp. 640–644, 2019, doi: 10.1007/978-3-030-29736-7_57

[8] **U. Norman** and A. E. Cicek, "ST-Steiner: a spatio-temporal gene discovery algorithm," in *Bioinformatics*, vol. 35, no. 18, pp. 3433–3440, Sep. 2019, doi: 10.1093/bioinformatics/btz110

Also presented as a poster in *26th Annual Conference on Intelligent Systems for Molecular Biology (ISMB)*, Chicago, IL, United States, Jul. 2018.

[9] F. K. Satterstrom, [and 189 others, including **U. Norman**], "Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism," *Cell*, vol. 180, no. 3, pp. 568–584.e23, Feb. 2020, doi: 10.1016/j.cell.2019.12.036

**Contributed equally to this work.*