

Robot Analytics: What Do Human-Robot Interaction Traces Tell Us About Learning?

Jauwairia Nasir^{*1}, Utku Norman^{*1}, Wafa Johal^{1,2}, Jennifer K. Olsen¹, Sina Shahmoradi¹,
and Pierre Dillenbourg¹

Abstract—In this paper, we propose that the data generated by educational robots can be better used by applying learning analytics methods and techniques which can lead to a deeper understanding of the learners’ apprehension and behavior as well as refined guidelines for roboticists and improved interventions by the teachers. As a step towards this, we put forward analyzing behavior and task performance at team and/or individual levels by coupling robot data with the data from conventional methods of assessment through quizzes. Classifying learners/teams in the behavioral feature space with respect to the task performance gives insight into the behavior patterns relevant for high performance, which could be backed by feature ranking. As a use case, we present an open-ended learning activity using tangible haptic-enabled *Cellulo* robots in a classroom-level setting. The pilot study, spanning over approximately an hour, is conducted with 25 children in teams of two that are aged between 11-12. A linear separation is observed between the high and low performing teams where two of the behavioral features, namely number of distinct attempts and the visits to the destination, are found to be important. Although the pilot study in its current form has limitations, e.g. its low sample size, it contributes to highlighting the potential of the use of learning analytics in educational robotics.

Keywords—educational robotics, learning analytics, computational thinking, path planning.

I. INTRODUCTION

The potential of robots is increasingly being explored in educational settings across the globe. When used in a classroom or group setting, they introduce a range of sensors that would not otherwise be available. Currently, these sensors are primarily being used to mediate and guide educational activities where the system benefits from being situated in the real world to help the learners gain understanding in various domains as diverse as physics [1], chemistry [2], maths [3], handwriting [4], [5], reading [6], storytelling [7], vocabulary [8], chess [9], meteorology [10] etc.

Two types of intervention of robots for learners have been identified [11]: i) robot as a tool, ii) robot as an agent. When used as a tool, researchers have explored using robots to teach robotics or other STEM curricular content [12]. As an agent, scientists focused their research in creating engaging and personalized scenarios for learners interacting with the robots [13]. While performances of students have been evaluated in both types of interventions, the analysis is often

limited to learning outcomes and statistics. We believe that robot sensors and interaction logs can be used to get deeper insights in the learning process of students during the activity with robots. These insights can be useful for designing new experiments, refining guidelines and enhancing the research value of studies in robots for learning.

Within the field of education, analytics is termed as *learning analytics* and is a relatively new trend. Learning analytics provides a lens for analysis and representation of data collected during a learning activity to improve learning [14]. Through the collection, measurement, and analysis of behavioral data, we can develop a deeper understanding of learning. Through the dissemination of this information back to teachers and students with visualizations and interventions, we can impact learning [15], [16]. To this end, we believe that learning analytics methods like *prediction* algorithms, *structure discovery*, and *relationship mining* can bridge the aforementioned gap as robots generate an abundance of data that can give useful information about learners’ behavior.

In this paper, we present a study featuring a constructivist learning activity, the *Cellulo City*, targeted to improve the computational thinking skills of children aged 11 to 12 through a path planning problem. The activity makes use of tangible haptic-enabled *Cellulo* robot [17]. Using this study, we demonstrate with learning analytics applied to the robots’ sensor and interaction logs that we could get a better understanding of learning strategies of students in link with their learning outcomes. Specifically, we observe patterns of interaction with the robot that are relevant for high performance by performing feature ranking, clustering and classification techniques on the collected data.

II. RELATED WORK

A. Use of Robots in Education

In STEM, for programming applications, the Thymio robot [18] has served as an environment to learn programming; however, a logging capability is not readily available in the associated software development environments. The reported experiments with Thymio consist mainly of workshops to gain programming experience [19] or recently with case studies for experimenting with swarm robotics [20], where they provide no analysis for learning. In [21], LEGO Mindstorms was used as the physical instances corresponding to objects in object-oriented programming paradigm.

For non-programming STEM activities, LEGO Mindstorms robot kit in [22] served as a platform to provide

^{*}The first two authors contributed equally to this work.

¹Computer-Human Interaction in Learning and Instruction (CHILI) Lab, Swiss Federal Institute of Technology in Lausanne (EPFL), Switzerland

²Biorobotics Laboratory (BioRob), Swiss Federal Institute of Technology in Lausanne (EPFL), Switzerland

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 765955.

students with physical robot building experience to understand mechanical concepts. Then, pretest, post-test and student notes on a template were analyzed for conceptual understanding and development. Furthermore, we increasingly see active tangible robots (most prominently *Cellulo*) being employed as experimental tools to enhance the learning *experience* for various educational concepts like wind formation [10] and states of matter [2]. The analyses in these two activities seem to be closest to our proposed approach. In [10], in addition to the post-test, robot logs such as pose of the robot, grasp and release events are collected to analyze more complex metrics like exploration similarity and entropy. The authors used these metrics to explain what each of them meant in terms of behavior and if statistically each individual metric relates to performance. Similarly, in [2], apart from the pretest and post-test, the authors statistically tie back the learning outcome to metrics generated from robot logs. However, in both of the studies, we do not see any prediction or structure discovery techniques to classify/cluster high and low performing teams with respect to a multi-dimensional behavioral space.

For other non-STEM disciplines, the letter handwriting activity [4] is similar to the aforementioned activities with *Cellulo*. Mostly, in this domain, we find activities such as story telling [7] and language learning [23], where social robots are used as either peers/tutors to assist in learning [13]. In addition to this, more recently, social robots are utilized for measuring high level constructs like engagement, trust and attention of the learners [24], [25], [26] that are assumed to be linked with performance.

B. Learning Analytics

When students are working on a technology-enhanced learning activity, they create data traces that can be used to analyze and support their learning. The most common modality of collected data consists of logs from the students' interactions with personal computing devices [27]. However, with the advances in technologies, learning analytics is shifting towards integrating multimodal data taking into account text, speech, visual, physiological cues by making use of data devices like Kinect sensor, microphones, eye-trackers, cameras, and wrist-bands/wearables [28]. These different data streams can be used to provide additional information about the learning state of the students, such as their current level of attention [29]. Multimodal data collection and analysis techniques (termed as multimodal learning analytics) can be utilized for looking into the learning process of the students in open-ended, complex learning environments where assessment is particularly difficult: in doing so, it would be possible to "generate distinctive insights into what happens when students create unique solution paths to problems, interact with peers, and act in both the physical and digital worlds" [30]. Specifically, in tasks where learners have to build functioning artifacts, learning analytics can help "to make rich inferences about learning and learners" [31]. For instance, [32] applied learning analytics to associate multimodal data with the experimental condition in a hands-on "making" activity (of

constructing a complete operational object) and correlate the student behavior with two learning outcomes, success in the task and learning. On gesture (hand/wrist movement), audio and bio-physiological data, the authors identify common student behaviors, and then compare how behaviors differ as a process in time and separately in frequency: the results support the usefulness of separating success, learning and process as well as considering multiple modalities in order to differentiate between effective behaviors.

Data captured by robots serves as a beneficial source and adds to the multimodality of the data. Robots that are used in the classroom can produce continuous behavioral data around how the students are interacting with the robot. As the data collected for analysis becomes more diverse, the methods within learning analytics used to address this data also grows. To better understand how these student behaviors relate to their learning, structure discovery methods and prediction algorithms can be used to find which behaviors are productive for learning. On the other hand, to understand which behaviors often proceed other behaviors, sequential pattern mining can be used. For instance, data mining has been applied to data from the interaction with the programming environment in order to classify [33] students' interaction with the educational robotics environment *Lego Mindstorms*, and predict [34] the learning progress of students to help teachers facilitate this progress. Once these relationships are better understood, interventions can be put into place. By applying learning analytics methods to data collected from a robot, we can develop a better understanding of the learning strategies that the students use as they engage with the robot.

Briefly, with regards to the contributions, this paper: i) highlights the potential of analyzing learners' behaviors by coupling robot data with the data from conventional methods of assessment through quizzes in educational settings with a pilot study, ii) introduces a novel open-ended learning activity using tangible haptic-enabled *Cellulo* robots in a classroom-level setting for the pilot study, and iii) showcases classification of learners/teams in the behavioral feature space with respect to the task performance giving insight into the relevant behavior patterns for high performance in the use case.

III. LEARNING ACTIVITY DESIGN

A. Path Planning and the Notion of Cost

Path planning can simply be defined as developing a strategy aiming to optimize certain constraints of navigating from one point to another. In order to find the best path when multiple paths exist, one would need to define a *criterion* to determine how good a path is. For instance, considering the layout of a city as in the *muddy city* problem¹, a criterion could be the number of paving stones connecting two houses or the number of houses that you pass to reach a specific house. This criterion is basically what is termed as the notion of *cost* in the field of path planning, and the goal is to minimize this cost.

¹<https://csunplugged.org/minimal-spanning-trees/>

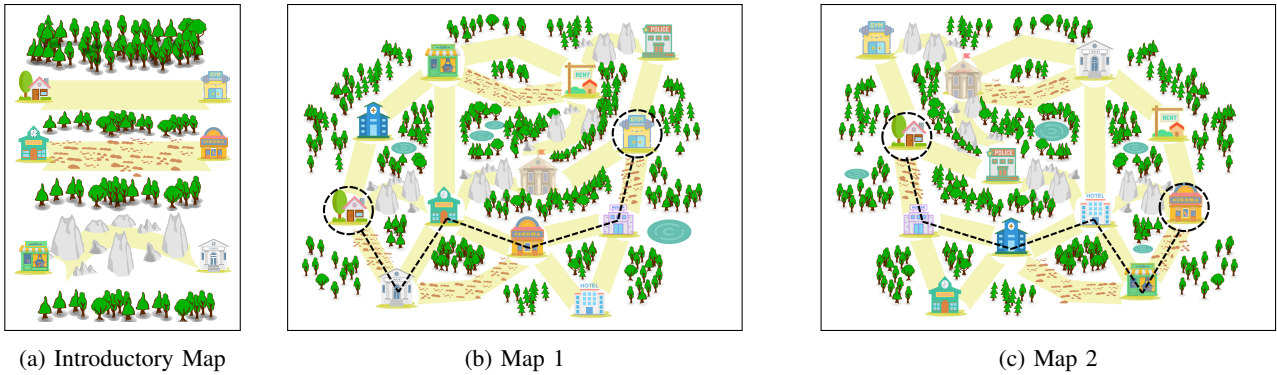


Fig. 1: (a) The introduction sheet to familiarize children with Cellulo, specifically how to move a Cellulo robot and to distinguish haptic feedback given by the robot. (b, c) Map 1 and Map 2 used in the learning activity, where the goal on a map is to find the optimal path from *home* represented with a clip-art of a house to (b) *gym* and (c) *cinema*. The source and destination nodes, and the optimal paths are highlighted here with dashed circles and dashed lines respectively.

TABLE I: Content breakdown for path planning

Concept	Description
node	the idea of a node as an entity
edge	the idea of an edge as a connection or link between two entities
path	the idea of a path as a sequence of edges connecting nodes
reachability of a node	the existence of a path between two nodes
multiplicity of paths	that there can be multiple distinct paths between two nodes, rather than a single path
path length	the length of a path as the number of edges on that path
edge weight	the idea of a weight function on edges
path distance	interpreting the length of a path as the sum of costs on the edges of that path
optimality of a path	that a path can be better (more preferable) or worse by some cost criterion

We highlight some of the core concepts involved in path planning in Table I. The idea is to design a learning activity in a manner that helps in extracting some part of a concept by exploratory behavior; hence, serving as an advance organizer [35] to a conventional lesson or even a stand alone session with some modifications.

Path planning is useful in applications as diverse as computer networks, telephone networks, oil pipelines, airline routes, social networks, etc. All in all, the inspiration for the subject material around path planning is the idea that it can help children acquire some of the computational thinking skills such as *analysis* and *pattern generalization*. The motivation behind choosing a constructivist activity design involving path planning as a use case shares the same spirit as proposed by [36] that “a constructionism-based problem-solving learning environment, with information processing, scaffolding and reflection activities, could be designed to foster computational practices and computational perspectives”.

B. Learning Activity for the Notion of Cost

We formulate our open ended learning activity on the lines of helping a robot go from a source location *home* to a destination by using as little battery as possible. The learning outcome for this activity, i.e. what the learners are expected to learn at the end, is the notion of *cost*. What are the learners able to do at the end, that they are not able to do at the beginning, is that they are able to choose the best path among many possible paths by considering the cost of each segment.

To set the task space for the *Cellulo City* problem, we

design maps as given in Fig. 1, and utilize the Cellulo robotic platform where a tangible, handheld, haptic-enabled Cellulo robot operates on a printed paper sheet [17]. The sheet is overlaid with Anoto pattern [37] by utilizing an open-source 2D localization library *libdots*². A Cellulo robot is equipped with a camera below with the help of which it can infer its position and orientation on the printed document.

1) *Dimensions of the cost criterion*: In order to represent the cost of traversing an edge, each edge is assigned a weight that is associated with two dimensions: the type of the edge and the length of the edge. For the type dimension, we distinguish between three types of edges—plain, stony, mountainous as shown in Fig. 1a. For the length dimension, an edge is either of one or two unit lengths, where the cost of an edge is twice if it is two unit lengths. These relationships constitute as the specific instance of the notion of cost in the learning activity that is to be discovered by the learners.

2) *Feedback modalities*: The first modality is used to visualize the remaining pretended battery level through the six RGB LEDs located on top of a Cellulo robot. When a Cellulo is placed at home, the pretended battery is charged to the full value of 24 and displayed by illuminating all six of the robot’s LEDs with the color green, which represent six full batteries—see Fig. 2a for a Cellulo placed at home. As the robot is moved on an edge, the battery value decreases by the cost on that edge from which a learner can infer the total cost of the path taken. Every time the robot is moved to home, the battery level is reset to six full batteries.

The second modality is the haptic feedback that matches with the type of an edge: a Cellulo robot gives little resistance and vibrates on stony edges, gives higher resistance with no vibration on mountainous edges, and assists the movement on plain edges.

IV. PILOT STUDY

A. Setup

The pilot study was conducted with a class of twenty five children that were aged 11 to 12 years (15 females: $M = 11.9$, $SD = 0.34$; 10 males: $M = 11.9$, $SD = 0.47$). The learners

²<https://chili.epfl.ch/libdots>

TABLE II: Activity pipeline of our pilot study. Visual refers to the visual representation of the remaining battery level through the LEDs on top of the robot.

Name	What are the learners asked to do?	Feedback from the Robot	Level	Duration
Introduction	hold and move the Cellulo robot and feel the haptic feedback associated with the type of the road to get familiar with the robot	- haptic	team	2 min
Phase 1 (Complete Feedback Phase)	find the best path from the source to the destination	- haptic - visual	team	15-20 min
Phase 2 (Delayed Feedback Phase)	find the best path from the source to the destination on a different map	- haptic - visual at home and destination	team	15-20 min
Individual Quiz	draw the best path on a set of three maps	not available (no robot)	individual	5 min
Collaborative Quiz	answer a list of questions that were designed to gather the learners' understanding about the underlying notion of cost	not available (no robot)	team	5 min
Time for Telling	listen to a brief explanation on what makes a path an optimal path, what is notion of cost, and link them to the activity at hand	not available (no robot)	all	5 min



(a) A team of two children

(b) Several teams

Fig. 2: Photos of (a) a single team while answering the collaborative quiz and (b) several teams participating concurrently in the learning activity in our pilot study

participated in teams of two (with the exception of one team, team 1, that consisted of three children) and each team was given one Cellulo robot. The study was run in parallel with all teams going through the activity pipeline described in Section IV-B synchronously. There were four supervisors, one per three teams, that instructed and managed the teams. The learners had no prior experience with the Cellulo robot platform used in this study, and reported by their teacher to have been coding occasionally for three years. The study was in the form of a one hour long session, and took place in a school in Switzerland. See Fig. 2a and Fig. 2b for one team and several teams in parallel respectively, as they participate in the learning activity.

B. Activity Pipeline

See Table II for an overview of the activities that the learners went through. In Phase 1 and Phase 2, the teams were asked to find the best path from home to the destination. In Phase 1, the remaining pretended battery level of the robot is shown throughout the map, whereas in Phase 2, the battery level is shown only at the goal node and not on the intermediate nodes or edges. In both phases, learners are allowed to repeat going from home to destination as many times as they want for a limited time. The teams were not notified whether a path they took was the best or not. Each team was given a blank note-taking sheet (with a printed suggestion to use the paper to keep track of the battery) that was collected at the end of the phase. Afterwards, the teams were asked to report the best path on their given map and draw the path on a copy of the map. In Phase 1, the first six teams were given Map 1 and the remaining six teams were given Map 2 which were then swapped in Phase 2—see Fig. 1b and Fig. 1c for the maps, and the best paths to be found and reported by the teams.

After the two phases in which the children worked in teams, in individual quiz, each learner was asked to draw the best path on paper for a set of maps similar to the maps in the phases. Then, in collaborative quiz, each team was asked to answer a list of questions that were designed to gather the learners' understanding about the underlying notion of cost in the context of this activity, that connects to the learning outcome of the activity. Finally, we showed the teams their performance as collected through robots on an orchestration teacher dashboard, which is out of scope of this paper. We explained the teams together the notion of cost and what makes a path an optimal path as a general case, as well as by linking it to the activity at hand. A detailed description is available online³.

C. Data Collection

We recorded the poses (position and orientation) of all robots on the paper maps in Introduction, Phase 1 and Phase 2. Every time an event with a pose-change (either position or orientation or both) occurred, it was recorded. We also collected the note-taking sheets given to the teams in the two phases, the best path reporting sheets that were distributed to the teams at the end of each phase, and the individual and the collaborative quizzes.

V. RESULTS AND DISCUSSION

We discuss the performance and behavioral analysis in the following subsections.

A. Analysis of the Task Performance

We do not measure the *learning gain* itself on the concept of the notion of *cost* per se but we measure: if the children are able to report the best path correctly i) for the maps in the two phases, ii) for the maps in the individual quiz; and iii) to what extent they understand the notion of *cost* in *this* context, which is evaluated through the collaborative quiz.

1) *Success in the Phases*: A team in a phase is labeled as “Pass” if it successfully reported the optimal path in the end of the phase, and as “Fail” otherwise. For each of the two phases, 10 out of 12 teams were able to find the best path, and there exist no teams that failed both of the phases.

³<https://github.com/chili-epfl/robot-analytics>

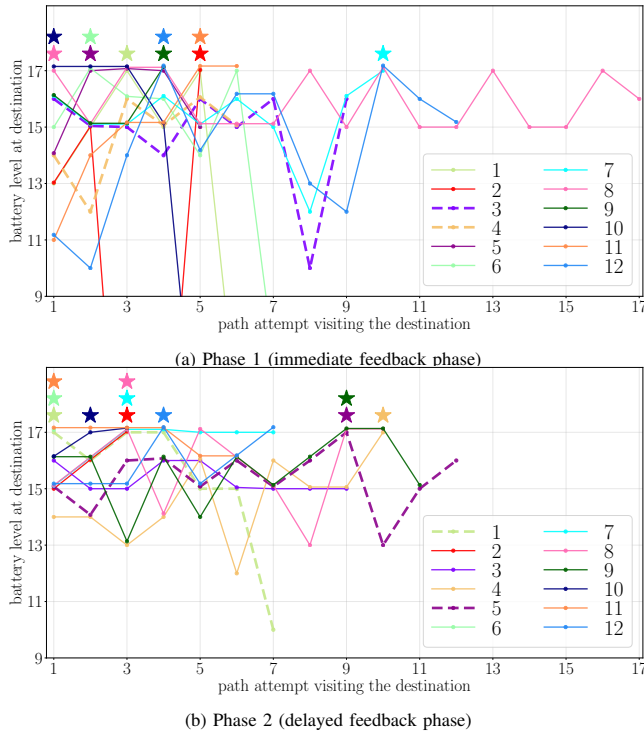


Fig. 3: Robot battery level at destination vs. path number for each team. A star marks the first time a team traverses on the optimal path. Dashed line indicates that the team has failed to report the optimal path. The best path is unique and has the cost of 17.

Fig. 3a and Fig. 3b give the battery level at the destination for each path that a team traversed in Phase 1 and Phase 2, respectively. We see that the two teams (Team 3 and 4) that were not able to report the best path correctly in Phase 1 were the only ones that never came across it during the exploration of the map. However in Phase 2, the two teams that failed did come across the best path during the exploration stage, but were unable to report it at the end. Note that most teams kept attempting even after reaching the best score, since they were not notified about the path being best.

2) *Quiz and Overall Performance*: From the quizzes, we calculate for each team i) individual quiz score of the team, by averaging the individual quiz scores of the team members as estimated by the fraction of correct answers per individual; and ii) collaborative quiz score, by averaging the Jaccard index used to measure the similarity between a team’s answer with the correct answer for each multiple-choice multiple-answers question in the collaborative quiz.

Consider the distribution of i) average quiz score in Fig. 4b, calculated by averaging collaborative quiz score and individual quiz score for each team; and ii) overall score

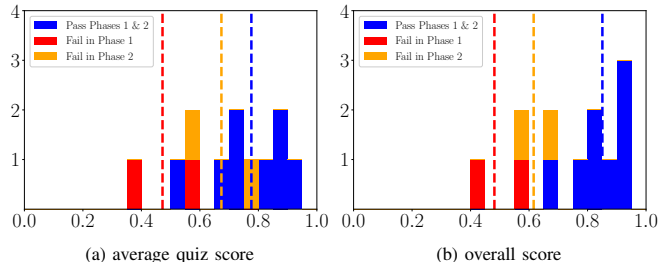


Fig. 4: Stacked bar charts that show the distribution of team-level task performance. The averages of phase success groups are marked by the dashed lines.

TABLE III: Exploration metrics and their ranks in Phase 1 and 2

Exploration Metric	Rank in Phase 1	Rank in Phase 2
number of attempts	2	2
number of distinct attempts	1	4
number of destination attempts	5	1
average variation of attempts	3	3
number of pauses	4	6
number of backtracks	6	5

in Fig. 4a, estimated by averaging collaborative quiz score, individual quiz score, success in Phase 1, and success in Phase 2, where success is one if the team reports the best path and zero otherwise. We see that the teams that passed both phases score higher. The teams that were not able to find the best path in Phase 1 score the lowest followed by the ones that failed Phase 2. The teams failing in Phase 1 did not transfer their understanding well to the following phase, since they failed despite the immediate feedback in Phase 1.

B. Analysis of the Teams’ Exploration Behavior

In this section, we analyze the behavior of teams to see if some behaviors taken by learners lead to better performance outcomes.

1) *Exploration Metrics and Their Ranking*: With the purpose of quantifying the exploration behavior of a team, we consider and explore six basic features and omit their transformations. The features are defined as follows:

- *Number of attempts* is the number of paths taken by a learner, where a path begins at home but does not necessarily visit the destination.
- *Number of distinct attempts* is the number of unique paths taken.
- *Number of destination attempts* is the number of paths that go through the destination node.
- *Average variation of attempts* is the averaged graph edit distance of all subpaths to the best path, where a subpath is only the portion from source to destination of an attempt. The edit distance is calculated with all edit costs (node or edge deletion/insertion) as 1.
- *Number of pauses* is calculated by summing pause counts at three category of places: i) source node, ii) destination node, and iii) the remaining i.e. all of the edges and intermediate nodes. We define a pause at a node or an edge in a category as spending more than one standard deviation time from the time spent in that specific category in both Phase 1 and Phase 2 averaged for all teams.
- *Number of backtracks* is the sum of backtrack counts, where a backtrack is defined as moving a robot backward from one node to another node on the x-axis (further from the goal and closer to the home), and specifically into the home and out of the goal for the neighbors of source and goal.

Among all the features, we notice that not all of them are equally informative for performance in the given scenario. In Table III, exploration metrics are ranked using a Support Vector Machine (SVM) classifier with linear kernel, with a large penalty parameter $C = 1000$, by using the absolute value of SVM weights as ranking criterion [38]. We observe

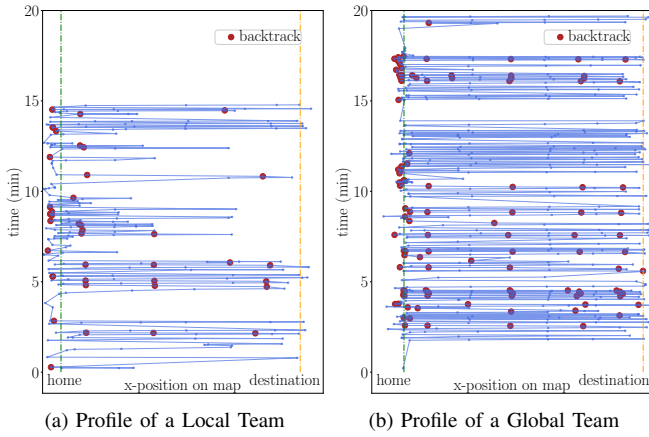


Fig. 5: Dynamics of robot movement on x-axis for two teams which exhibit (a) a local attitude and (b) global attitude in their exploration of the map in Phase 1. The blue lines indicate the position of the Cellulo on the x-axis, which roughly aligns the axis from the home to the destination. The red dots indicate where a backtrack is made.

that some of the features are less predictive of performance: For instance, *number of pauses* and *number of backtracks* are less informative, whereas *number of attempts* and *average variation of attempts* seem to be good predictors. Another interesting observation is that the highest ranked feature for each phase is among the lowest ranked in the other phase which indicates how a certain behavior is important the first time but not in the delayed feedback phase and vice versa.

2) *Local vs Global Exploration Attitude*: Using the two features *number of destination attempts* and *number of attempts*, we observe two kinds of attitudes in the way a team explores a map: a team that prefers a more *global* approach going until the destination, or a more *local* approach. Fig. 5 illustrates how the Cellulo robot is moved between home and the destination. Note that the exploration attitude is not an indication of the performance as can be seen in the Table IV that gives performance of the teams.

3) *Linking Exploration Metrics with Task Performance*: As an attempt to cluster the teams for patterns of behavior in the exploration metrics space, we selected the top three ranked features in Phase 2 and ran K-Means algorithm, which resulted in two clusters for optimizing the average silhouette. In Fig. 6, we do not see the clusters correlating with the performance. Hence, as a step further, we employ an SVM classifier with linear kernel in the same space: it is interesting to observe the feature *number of destination attempts* allows for the linear separability to low and high performing teams of the given dataset. In Phase 2, the teams that visited destination less seem to have performed better. This could be because of a better transfer of the underlying cost in Phase 1 while for the teams that did more destination attempts, it could be because of a lack of understanding

TABLE IV: The relevant exploration and performance metrics for a local and a global team in Phase 1, where the attitude can be seen at the fraction of destination attempts

Metric	Local Team	Global Team
number of backtracks	40	112
number of destination attempts / number of attempts	21.7%	54.8%
Phase 1	Pass	Pass
Phase 2	Fail	Pass
Quiz Average	59%	68%
Overall	57%	79%

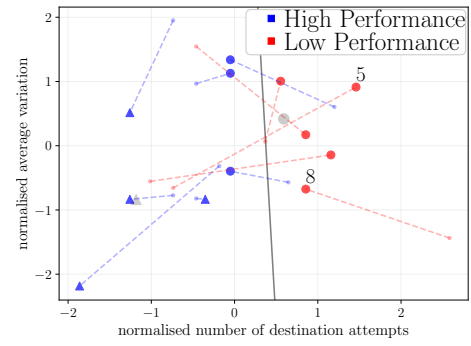


Fig. 6: Visualization of the teams in Phase 2, as drawn in the feature space spanned by the three highest ranked metrics in Phase 2 and projected onto two. High performing teams ($> 70\%$ average quiz score) are in blue, and low performing are in red. A line that linearly separates the two classes is drawn using a linear SVM classifier trained on the three features, with penalty $C = 1000$. The marker shape indicates which cluster the data point belongs to in two cluster K-Means, where cluster centers are indicated in gray, and the number of clusters is optimal for maximizing the average silhouette. Dashed lines are drawn from positions of the teams in Phase 1 to 2 and illustrate how a team moves in the feature space. The numbers on the data points indicate the team number. The metrics are normalized to zero mean and unit variance.

which lead them to try to solve the problem at hand (to succeed in the phase) by brute force. This behavior could also be suggestive of a thought process that considers the problem in Phase 2 as a completely new problem while it was just a reverse problem of Phase 1—see Fig. 1.

We notice that the teams with high scores lie in the similar range of exploration metrics in both the phases. One interesting thing to note in Fig. 6 is that two of the teams that failed in Phase 1 moved through the decision boundary closer to the low scoring teams. In other words, we see teams that kept the same behavior in both phases (i.e. lie on the same side of the boundary in the feature space) and teams that change their behavior (i.e. move in the feature space to the other side of the boundary in Phase 2). Although, with a low sample size, it may not be wise to generalize; however, it can be said that certain behaviors can be good indicators for high performance. This information can then be utilized by the roboticists/teachers to refine the learning design or focus on certain aspects of the learners' behavior more than the others. One hypothesis that could be tested in a larger study is to see whether interventions made early in Phase 1 would lead to learners getting a better understanding, and hence transfer well to Phase 2 and later. Another hypothesis that we can test is teams visiting destination more in Phase 2 would perform worse.

VI. CONCLUSION AND FUTURE WORK

In this paper, we presented a pilot study involving a novel open-ended learning activity with a robot to showcase the potential that lies in the data captured by robots to inform learning, especially, in constructivist activities that are more challenging to quantify. This can help the roboticists to design robot's behavior for learners, and to have effective interventions, improved learning design and orchestration that could eventually benefit teachers as well. In an experiment carried out at a school in Switzerland, we observe how a classification technique discriminates between high and low performing students based on behavioral features extracted

from the data collected through robots. However, in the current design, there are limitations that can be improved.

As such, for future work, we plan to refine the design by adding a path planner that can help the system evaluate how good a path that is being taken by a learner is at a given point in time for appropriate interventions. In addition to this, it may be a nice idea to integrate heuristics from [39] to have a more robust and plausible activity design. Also, we feel it may be interesting to increase the complexity of the map to discourage brute force approaches. This will hopefully promote meta-cognitive behavior like reflection which can improve learning. Furthermore, certain landmarks can be defined inside the activity for interventions. For instance, in the immediate feedback phase, going through the best path at least once in a certain time limit during the exploration can be a possible landmark.

ACKNOWLEDGMENT

We would like to thank Patrick Jermann for helping us to perform a pre-experiment. We are grateful to the Swiss National Science Foundation for supporting this project through the National Centre of Competence in Research Robotics.

REFERENCES

- [1] D. C. Williams, Y. Ma, L. Prejean, M. J. Ford, and G. Lai, "Acquisition of physics content knowledge and scientific inquiry skills in a robotics summer camp," *JRTE*, vol. 40, no. 2, pp. 201–216, 2007.
- [2] A. Özgür, "Cellulo: Tangible Haptic Swarm Robots for Learning," 2018.
- [3] R. Mitnik, M. Nussbaum, and A. Soto, "An autonomous educational mobile robot mediator," *Autonomous Robots*, vol. 25, no. 4, pp. 367–382, 2008.
- [4] T. Asselborn, A. Guneysoy, K. Mrini, E. Yadollahi, A. Ozgur, W. Johal, and P. Dillenbourg, "Bringing Letters to Life: Handwriting with Haptic-Enabled Tangible Robots," in *IDC*, pp. 219–230, 2018.
- [5] S. Lemaignan, A. Jacq, D. Hood, F. Garcia, A. Paiva, and P. Dillenbourg, "Learning by Teaching a Robot: The Case of Handwriting," *IEEE RA Magazine*, vol. 23, no. 2, pp. 56–66, 2016.
- [6] E. Yadollahi, W. Johal, A. Paiva, and P. Dillenbourg, "When deictic gestures in a robot can harm child-robot collaboration," in *IDC*, pp. 195–206, 2018.
- [7] J. Kory Westlund, "Telling stories with green the dragonbot: A showcase of children's interactions over two months," in *HRI EA*, (New York, NY, USA), pp. 263–263, ACM, 2015.
- [8] J. M. Kory, S. Jeong, and C. L. Breazeal, "Robotic learning companions for early language development," in *ICMI*, pp. 71–72, 2013.
- [9] A. Pereira, C. Martinho, I. Leite, and A. Paiva, "iCat, the chess player: the influence of embodiment in the enjoyment of a game," *AAMAS*, pp. 1253–1256, 2008.
- [10] A. Özgür, W. Johal, F. Mondada, and P. Dillenbourg, "Windfield: Learning Wind Meteorology with Handheld Haptic Robots," *HRI*, pp. 156–165, 2017.
- [11] O. Mubin, C. J. Stevens, S. Shahid, A. A. Mahmud, and J.-J. Dong, "A review of the applicability of robots in education," *Technology for Education and Learning*, vol. 1, pp. 1–7, 2013.
- [12] M. E. Karim, S. Lemaignan, and F. Mondada, "A review: Can robots reshape K-12 STEM education?," *ARSO*, 2015.
- [13] T. Belpaeme, J. Kennedy, A. Ramachandran, B. Scassellati, and F. Tanaka, "Social robots for education: A review," *Science Robotics*, vol. 3, no. 21, p. eaat5954, 2018.
- [14] D. Clow, "An overview of learning analytics," *Teaching in Higher Education*, vol. 18, no. 6, pp. 683–695, 2013.
- [15] R. Baker and G. Siemens, "Educational Data Mining and Learning Analytics," in *CHLS* (R. K. Sawyer, ed.), pp. 253–272, Cambridge: Cambridge University Press, 2012.
- [16] D. Gašević, S. Dawson, and G. Siemens, "Let's not forget: Learning analytics are about learning," *TechTrends*, vol. 59, no. 1, pp. 64–71, 2015.
- [17] A. Özgür, S. Lemaignan, W. Johal, M. Beltran, M. Briod, L. Pereyre, F. Mondada, and P. Dillenbourg, "Cellulo: Versatile Handheld Robots for Education," *HRI*, pp. 119–127, 2017.
- [18] F. Mondada, M. Bonani, F. Riedo, M. Briod, L. Pereyre, P. Retornaz, and S. Magnenat, "Bringing Robotics to Formal Education: The Thymio Open-Source Hardware Robot," *IEEE RA Magazine*, vol. 24, no. 1, pp. 77–85, 2017.
- [19] F. Riedo, M. Chevalier, S. Magnenat, and F. Mondada, "Thymio II, a robot that grows wiser with children," in *ARSO*, pp. 187–193, 2013.
- [20] A. Vitanza, P. Rossetti, F. Mondada, and V. Trianni, "Robot swarms as an educational tool: The Thymio's way," *IJARS*, vol. 16, no. 1, pp. 1–13, 2019.
- [21] P. B. Lawhead, M. E. Duncan, C. G. Bland, M. Goldweber, M. Schep, D. J. Barnes, and R. G. Hollingsworth, "A Road Map for Teaching Introductory Programming Using LEGO® Mindstorms Robots," *ITiCSE-WGR*, vol. 35, no. 2, pp. 191–201, 2002.
- [22] J. M. Chambers and H. Murray, "Developing conceptual understanding of mechanical advantage through the use of Lego robotic technology," *Australas. J. Educ. Technol.*, vol. 24, no. 4, pp. 387–401, 2008.
- [23] P. Vogt, R. van den Berghe, M. de Haas, L. Hoffmann, J. Kanero, E. Mamus, J.-M. Montanier, O. Oudgenoeg-Paz, D. H. Garcia, F. Papadopoulos, T. Schodde, J. Verhagen, C. D. Wallbridge, B. Willemsen, J. de Wit, T. Belpaeme, T. Göksun, S. Kopp, E. Krahmer, A. C. Küntay, P. Leseman, and A. K. Pandey, "Second Language Tutoring using Social Robots. A Large-Scale Study," in *HRI*, 2019.
- [24] G. Castellano, I. Leite, A. Pereira, C. Martinho, A. Paiva, and P. W. McOwan, "Affect recognition for interactive companions: Challenges and design in real world scenarios," *JMUI*, vol. 3, no. 1, pp. 89–98, 2009.
- [25] L. J. Corrigan, C. Peters, and G. Castellano, "Identifying task engagement: Towards personalised interactions with educational robots," *ACII*, pp. 655–658, 2013.
- [26] L. V. Brown, R. Kerwin, and A. M. Howard, "Applying behavioral strategies for student engagement using a robotic educational agent," *SMC*, pp. 4360–4365, 2013.
- [27] M. Khalil and M. Ebner, "What is learning analytics about? a survey of different methods used in 2013-2015," in *Proceedings of smart learning conference*, pp. 1–12, UAE, 2016.
- [28] P. Blikstein, "Multimodal learning analytics," in *LAK*, pp. 102–106, ACM, 2013.
- [29] K. Sharma, H. S. Alavi, P. Jermann, and P. Dillenbourg, "A gaze-based learning analytics model: in-video visual feedback to improve learner's attention in MOOCs," in *LAK*, pp. 417–421, ACM, 2016.
- [30] P. Blikstein and M. Worsley, "Multimodal Learning Analytics and Education Data Mining: Using Computational Technologies to Measure Complex Learning Tasks," *JLA*, vol. 3, no. 2, pp. 220–238, 2016.
- [31] M. Berland, R. S. Baker, and P. Blikstein, "Educational data mining and learning analytics: Applications to constructionist research," *Technology, Knowledge and Learning*, vol. 19, no. 1-2, pp. 205–220, 2014.
- [32] M. Worsley and P. Blikstein, "A Multimodal Analysis of Making," *IJAIED*, vol. 28, no. 3, pp. 385–419, 2018.
- [33] I. Jormanainen and E. Sutinen, "Using data mining to support teacher's intervention in a robotics class," *DIGITEL*, pp. 39–46, 2012.
- [34] I. Jormanainen and E. Sutinen, "Role blending in a learning environment supports facilitation in a robotics class," *Educational Technology and Society*, vol. 17, no. 1, pp. 294–306, 2013.
- [35] D. P. Ausubel, "The use of advance organizers in the learning and retention of meaningful verbal material," *Journal of Educational Psychology*, vol. 51, no. 5, pp. 267–272, 1960.
- [36] S. Y. Lye and J. H. L. Koh, "Review on teaching and learning of computational thinking through programming: What is next for K-12?," *Computers in Human Behavior*, vol. 41, pp. 51–61, 2014.
- [37] M. P. Pettersson, "Method and device for decoding a position-coding pattern," Dec. 5 2006. US Patent 7,145,556.
- [38] I. Guyon, J. Weston, and S. Barnhill, "Gene Selection for Cancer Classification using Support Vector Machines," *Machine Learning*, vol. 46, pp. 389–422, 2002.
- [39] C. Giang, A. Piatti, and F. Mondada, "Heuristics for the Development and Evaluation of Educational Robotics Systems," *IEEE Transactions on Education*, pp. 1–10, 2019.