

Is There ‘ONE way’ of Learning? A Data-driven Approach

Jauwairia Nasir
CHILI Lab, EPFL
Lausanne, Switzerland
jauwairia.nasir@epfl.ch

Barbara Bruno
CHILI Lab, EPFL
Lausanne, Switzerland
barbara.bruno@epfl.ch

Pierre Dillenbourg
CHILI Lab, EPFL
Lausanne, Switzerland
pierre.dillenbourg@epfl.ch

Abstract

Intelligent Tutoring Systems (ITS) are required to intervene in a learning activity while it is unfolding, to support the learner. To do so, they often rely on performance of a learner, as an approximation for engagement in the learning process. However, in learning tasks that are exploratory by design, such as constructivist learning activities, performance in the task can be misleading and may not always hint at an engagement that is conducive to learning. Using the data from a robot mediated collaborative learning task in an out-of-lab setting, tested with around 70 children, we show that data-driven clustering approaches, applied on behavioral features including interaction with the activity, speech, emotional and gaze patterns, not only are capable of discriminating between high and low learners, but can do so better than classical approaches that rely on performance alone. First experiments reveal the existence of at least two distinct multi-modal behavioral patterns that are indicative of high learning in constructivist, collaborative activities.

CCS Concepts: • Human-centered computing → User studies; Collaborative and social computing design and evaluation methods.

Keywords: collaborative learning, constructivist classrooms, engagement, productive engagement, learning analytics, multi-modal data

ACM Reference Format:

Jauwairia Nasir, Barbara Bruno, and Pierre Dillenbourg. 2020. Is There ‘ONE way’ of Learning? A Data-driven Approach. In *Companion Publication of the 2020 International Conference on Multi-modal Interaction (ICMI '20 Companion)*, October 25–29, 2020, Virtual event, Netherlands. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3395035.3425200>

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

ICMI '20 Companion, October 25–29, 2020, Virtual event, Netherlands

© 2020 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-8002-7/20/10.

<https://doi.org/10.1145/3395035.3425200>

1 Introduction

Two of the most important components in an ITS include a *student model* and a *pedagogical model* [2, 9]. The pedagogical model is responsible for making appropriate intervention[s] in the activity (i.e., interventions that have a positive effect on the student’s learning), knowing the details of the learning activity and being informed by the student model about the student’s status. Bayesian Knowledge Tracing (BKT) [4] is one of the most widely used approaches to model student knowledge [5, 11, 13]. One of the assumptions in BKT is that at each step, the “student can either succeed or fail the task”, i.e. there is a straightforward, binary mapping between performance in the task and learning, which makes the approach most “relevant for tutors that use exercises and scaffolding as the main vehicle for learning” [5]. However, there is an increasing emphasis towards incorporating more open-ended/constructivist learning activities that encourage the awareness of the knowledge construction process, e.g. by promoting, among other things, *Problem Based Learning* that requires the learners to devise a solution to a real world problem together, and/or *Cooperative Learning* in which interdependence among group members is needed to solve a problem; thus, violating the requirement of a chain of binary right/wrong steps towards the goal [3, 12]. The learners rather “become engaged by applying their existing knowledge and real-world experience, learning to hypothesize, testing their theories, and ultimately drawing conclusions from their findings” [10]. As a consequence, a constructivist learning environment makes it non-trivial to accurately identify when and if a learner is *productively engaged* in the task, i.e., engaged in a way that is conducive to learning. Some of the latest research, for example L2TOR¹ and EMOTE² projects, has looked into advancing adaptive tutoring systems; however, not much work is found in the domain of ITS to model student knowledge in constructivists environments.

Here, we utilize data from *JUSTThink*³: an open-ended activity in which two learners cooperate as a team to solve a semi-guided algorithmic reasoning problem [7, 8]; hence,

¹<http://www.l2tor.eu/>

²<http://www.emote-project.eu/>

³This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 765955 (ANIMATAS). We are also grateful to the SNSF for supporting this project through the National Centre of Competence in Research Robotics

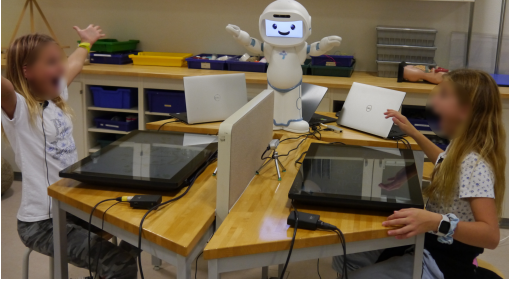


Figure 1. A team interacting with the JUSThink platform: how can observable multi-modal behaviors be linked with learning?

following both the *Problem Based Learning* and *Cooperative Learning* paradigms. In a previous analysis on these data, we observed that performance of the learners in the task was not directly correlated to learning gains, while it affected how the learners evaluate their competence, task engagement and mutual understanding [7]. We postulate the use of data-driven approaches on multi-modal behavioural features to make subtle learning-related behavioural cues emerge, and thus allow for the identification of high learners in a way which is independent from performance and human annotations, hence paving a way for novel ITS and, more broadly, more effective learning-supporting systems and interventions. It must be noted that here performance and learning refers to two very different things. By performance, we mean failing or succeeding at the task; while by learning we mean how the learners score at their pre- and post- tests.

2 Setup

We utilize the dataset generated from a study done with the JUSThink platform [7] where 68 children (in teams of two i.e., 34 teams) aged 9 to 12 years interact with a collaborative learning platform consisting of two screens and a QRobot acting as a guide and a mediator (see Fig. 1). For post analysis of the collaborative behavior, we look at: 1) **Multi-modal Behaviors** and 2) **Learning and Performance Measures**. For the former, we look at the children's interactions with the system, affective states, gaze and speech patterns that are extracted automatically using various sensors [7]. In total, there are 28 features where we use averages and frequencies for various behavioral features. For the latter, we take into account the team level learning measures such as absolute, relative, and joint absolute learning gain (A_LG, R_LG, J_A_LG, respectively); and the performance measure as the last error (L_E) in their solution. These measures are discussed and elaborated in an earlier work [7].

3 Research Question and Our Approach

To explore whether *the way in which children learn can be identified from their multi-modal behaviors in a collaborative exploratory activity*; we propose the technique graphically summarized in Figure 2. Under approach A, the teams are

Table 1. Clusters in each approach with mean values for learning gains (LG) as well as the last error

Cluster no.	L_E	A_LG	R_LG	J_A_LG	N
Approach A					
1	0.166	0.433	0.438	0.357	6
2	0.681	0.720	0.748	0.742	10
3	0.836	0.340	0.168	0.314	5
4	0.132	0.700	0.733	0.769	13
Approach B					
1	0.461	0.678	0.693	0.714	14
2	0.393	0.616	0.604	0.607	12
3	0.393	0.383	0.348	0.428	6

Table 2. Classification Results

Classifier	k-fold cross-validation		test-set	
	Accuracy	F1-score	Accuracy	F1-score
Approach A				
SVM	0.28	0.24	0.33	0.23
RF	0.49	0.29	0.44	0.46
Approach B				
SVM	0.80	0.74	1	1
RF	0.80	0.77	1	1

clustered according to their learning as well as performance (A-1) and the resulting cluster labels (denoted *Type 1 labels*) are used as ground truth (A-2) for a classifier trained on the teams' multi-modal behavioural features, to discriminate between high and low learners based on their behaviour (A-3).

Intuitively, this approach assumes that teams in each type of cluster w.r.t. learning and performance would exhibit a unique set of behaviors. *But what if learning, rather than one, is associated to two or more distinct sets of behaviours?* Designing a method allowing for testing the validity of this hypothesis is the focus of this work.

Approach B highlights a possible paradigm shift to deal with such a scenario. In this approach, teams are first clustered according to their multi-modal behaviors (B-1) and the resulting clusters are compared in terms of their average learning and performance metrics (B-2). If the above hypothesis is true, i.e., learning is associated to two or more distinct sets of behaviors, and we observe that: (1) the learning and performance profiles of at least some of these clusters show significant differences and (2) it is possible to define a one-to-many, comprehensible mapping between the clusters obtained in approach A and those obtained under approach B, in terms of the teams they consist of; then in this case, the clusters obtained with approach B would represent *distinct variants of high and low learning behaviours*. These corresponding labels (denoted *Type 2 labels*) could then be used

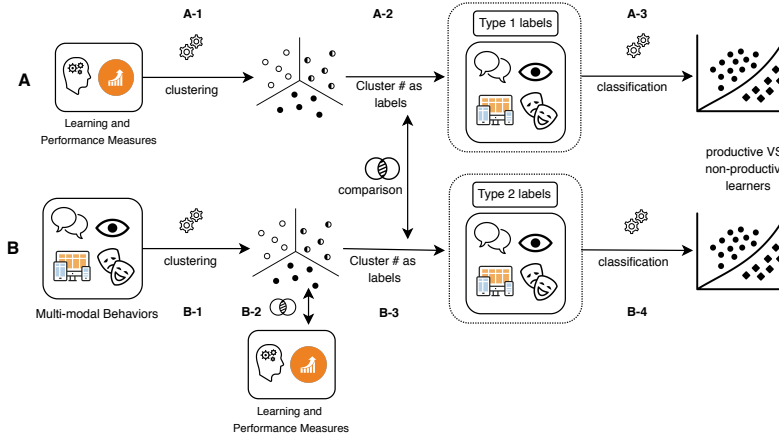


Figure 2. The two approaches we adopt and compare to investigate the relation between learning and behavior⁴

as ground truth (B-3) for a classifier trained on the teams' multi-modal behavioural features, to discriminate between high and low learners based on their behaviour (B-4). To summarize, in *A*, everything builds on learning as well as performance metrics from the start. In *B*, the multimodal behaviors come first, and the team-based learning and performance metrics are considered only after this initial clustering.

4 Findings

In both approaches, we rely on k-means as clustering method, using entropy analysis to properly set the parameter k , and both SVM and Random Forest as classifiers, similar to [1, 6] in steps A-3 and B-4. For both classifiers we rely on k-fold cross validation and parameters setting via grid-search.

The clustering step (A-1) in approach *A* identifies four different clusters, representing the four different combinations of high/low learning gains/last error (see Table 1).

As the upper half of Table 2 shows, using these *Type 1 labels* results in a very low accuracy and F1 score with both the SVM ($C = 1$, $kernel = poly$, $class_weight = balanced$) and the RF classifier ($n_estimators = 200$, $max_features = sqrt$, $class_weight = balanced$).

Moving to approach *B*, the clustering step (B-1) identifies three different clusters (see Table 1) which means that there are three sets of distinct behaviors observed among the participants. We note that the average performance for each of these clusters is similar while the average learning gains are not (B-2); hence, once again highlighting that performance may not always represent the actual engagement in the learning process. Furthermore, once verified the link of these clusters with learning, we further elaborate that when comparing these clusters with clusters generated in approach *A*: two of these behavioral clusters largely contain teams with higher learning gains. Some of these learning gains are

significantly different (p-values of $p = 0.016$, $p = 0.025$, and $p = 0.026$ obtained by Kruskal-Wallis test) from the learning gains exhibited by the third behavioral cluster, which is associated more with low learning teams (see Table 1). This means that condition (1) and (2), highlighted in section 3, are met.

Hence, these two conditions being met allow for a perspective shift as mentioned in the previous section: using the cluster numbers generated on multi-modal data as labels for supervised learning models where we distinguish productive from non-productive learners. As can be seen in the bottom half of table 2, *Type 2 labels* allow for achieving very high accuracy and F1 scores both with SVM ($C = 1000$, $kernel = rbf$, $\gamma = 0.001$, $class_weight = balanced$) and RF ($n_estimators = 100$, $max_features = auto$, $class_weight = balanced$).

5 Conclusion

The initial findings highlight a number of research hypotheses for ITS in constructivist designs that are worth exploring: 1) performance is not always a reliable indicator of engagement in the learning process, as hinted by the analysis conducted in step B-2; 2) learners having similar learning gain and performance profiles can have multiple set[s] of multi-modal behaviors (as supported by the two clusters of high-learning teams identified with approach *B*). This could also be the reason why the *Type 1 labels* did not perform well for classification; and 3) the use of behavioural labels, *proven solidly linked with learning*, seems to allow for a better discrimination between high and low learners than the direct use of learning labels, as suggested by the classification results of the two approaches reported in Table 2.

Thorough analyses investigating the above-listed research directions will give a deeper understanding of which multi-modal behavioural data and associated processing methods allow for effectively (and autonomously) separating low learners from high learners, to lead the way to effective interventions.

⁴The symbols in the image can be attributed to the Noun Project and FAVPNG

References

- [1] 2017. Fully Automatic Analysis of Engagement and Its Relationship to Personality in Human-Robot Interactions. *IEEE Access* (2017). <https://doi.org/10.1109/ACCESS.2016.2614525>
- [2] Alaa N Akkila, Abdelbaset Almasri, Adel Ahmed, Naser Masri, Yousef Abu Sultan, Ahmed Y Mahmoud, Ihab Zaout, and Samy S Abu-naser. 2019. Survey of Intelligent Tutoring Systems Up To the End of 2017. *International Journal of Academic Information Systems Research* 3, 3 (2019), 71–81. www.ijeais.org/ijaisr
- [3] John M. Brooks and Michael Brooks. 1993. In Search of Understanding: The Case for Constructivist Classrooms.
- [4] A T Corbett and John R Anderson. 1995. Knowledge tracing: Modeling the acquisition of student knowledge. , 253–278 pages. <http://act-r.psy.cmu.edu/papers/893/CorbettAnderson1995.pdf>
- [5] Michel C. Desmarais and Ryan S.J.D. Baker. 2012. A review of recent advances in learner and skill modeling in intelligent learning environments. *User Modeling and User-Adapted Interaction* 22, 1-2 (2012), 9–38. <https://doi.org/10.1007/s11257-011-9106-8>
- [6] Iolanda Leite, Marissa McCoy, Daniel Ullman, Nicole Salomons, and Brian Scassellati. 2015. Comparing Models of Disengagement in Individual and Group Interactions. *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (2015), 99–105. <https://doi.org/10.1145/2696454.2696466>
- [7] Jauwairia Nasir, Utku Norman, Barbara Bruno, and Pierre Dillenbourg. 2020. When Positive Perception of the Robot Has No Effect on Learning. In *29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*.
- [8] Jauwairia Nasir, Utku Norman, Barbara Bruno, and Pierre Dillenbourg. 2020. You Tell, I Do, and We Swap until we Connect All the Gold Mines! *ERCIM News* 2020, 120 (2020).
- [9] Hyacinth S. Nwana. 1990. Intelligent tutoring systems: an overview. *Artificial Intelligence Review* 4, 4 (1990), 251–277. <https://doi.org/10.1007/BF00168958>
- [10] Steve Olusegun. 2015. Constructivism Learning Theory: A Paradigm for Teaching and Learning. *IOSR Journal of Research Method in Education Ver. I* 5, 6 (2015), 2320–7388. <https://doi.org/10.9790/7388-05616670>
- [11] Jennifer Sabourin, Scott McQuiggan, and Andre De Waal. 2016. SAS tools for educational data mining. *CEUR Workshop Proceedings* 1633, 1 (2016), 85–106. <https://doi.org/10.3102/1076998616666808>
- [12] Paige L Schulte. 1996. A definition of constructivism. *Science Scope* 20, 3 (1996), 25–27.
- [13] George Siemens and Ryan S. J. d. Baker. 2012. Learning analytics and educational data mining. (2012), 252. <https://doi.org/10.1145/2330601.2330661>