Augmenting Teachers with Data Science Powers: Joining Human and Artificial Intelligence to Assist Students

Ranilson Paiva, Júlio Cesar de Holanda, Myron David, João Pedro

[1]Computing Institute – Federal University of Alagoas (UFAL)
Maceió, AL – Brazil

`{ranilsonpaiva, jcfsh, mdlcp, jpsv}@ic.ufal.br`

Abstract. *Technology has the potential to influence the processes and outcomes of teaching and learning. Researchers notice a high demand for online education, but these environments generate educational data in high quantity, diversity and velocity, which requires the use of technology. However, putting too much emphasis on artificial intelligence is not bringing the expected results. It is time to position teachers as the principal decision-makers in the "classroom" and use AI as an assistant. Our proposal is a tool (called T-Partner) to help online teachers: (1) search for relevant patterns in the educational data; (2) created visualisations to present the results in a more teacher-friendly way; (3) recommend study plans, based on students' needs. For that, we applied data science techniques on these data and processed the results, making it easier for teachers to understand them. We evaluated the tool, asking participants to use it to complete some tasks that simulate a teacher analysing data to make pedagogical decisions. We evaluated time, correctness, number of doubts and the participants' perceptions. The results show that most participants were able to complete the tasks in an appropriate amount of time and considered the T-Partner useful and easy to use in an online learning context.*

Keywords: *Augmented Intelligence, Data Science and Online Education.*

1. Introduction

Technology has the potential to influence the processes and outcomes of teaching and learning. A global belief in this potential is making countries invest in new and emerging technologies to support teaching and learning [Mouza 2018]. One of these technologies are the Online Learning Environments. The number of courses offered via this media increases constantly [Seaman et al. 2018]. Some countries, like Brazil, are passing laws to regulate online learning[1], others are investing considerable amounts to stimulate its use[2] and overcome some of its challenges [Li and Irby 2008]. These facts attest the high demand for online learning and that it is not just a trend, but a viable approach to propagate and democratize education.

---

[1]Available at: `https://bit.ly/2zM3JKp`
[2]Available at: `https://bit.ly/2FX7saW`

However, these environments offer some challenges to teachers and learners. One such challenge is to help teachers manage their courses. The difficulty lies in the fact that these environments store educational data in considerable volume, variety. These data are needed to provide teachers (almost) real-time results and feedbacks [Boettcher and Conrad 2016, Stephenson 2020]. According to [Kopcha 2012], the effective use of technology depends on the ways they are used in practice, highlighting the paramount importance of teachers' development to guarantee this effectiveness.

Besides that, we notice a recent change in focus: from artificial intelligence to human intelligence [Chou et al. 2011, of Education. 2014, Baker 2016]. This change, according to [Baker 2016], reflects the unfulfilled promises made by the Artificial Intelligence in Education field, repositioning teachers as the main decision-maker in the "classroom" [Burden 2014, Mouza 2018]. However, they do not have the necessary support and/or training [of Education. 2014, Dağ et al. 2014]

These information indicate the salient need, not only, to help teachers deal with the vast and diverse amount of educational data, but also provide them means to act based on these data and check the consequences of their actions [Paiva and Bittencourt 2017a, Paiva and Bittencourt 2017c]. These facts reinforce the latent necessity teachers have for direction and support in the technological field [Mandinach and Jackson 2012].

This way, we propose an authoring tool to assist online teachers to: (1) search for relevant pedagogical situations in the learning environment (using educational data); (2) extract and visualize patterns and trends, from these data, in order to understand what is happening with their students/groups; (3) create (personalized) study plans and other (personalized) pedagogical recommendations, delivering them to the target students; (4) evaluate the impact of these plans and recommendations, i.e., check whether they helped the students or not.

We implemented this authoring tool and evaluated an early version of this tool (called T-Partner), asking 24 participants to complete tasks related to what online teachers do when managing their courses. The results show the participants were able to properly complete most tasks and had very positive perceptions about the T-Partner, considering it easy to use, helpful and interesting to use.

2. Proposal

We intend to assist teachers from online learning environments. To do that, we use technology to augment their capacity to manage their courses, using educational data to help them make better pedagogical decisions. We implemented a tool named Teachers' Partner (T-Parnter), which follows a process called Pedagogical Decision-Making Process (PDMP). The PDMP and the T-Partner, together, constitute our proposal to give teachers the 4 capabilities listed in the previous paragraph. In the next subsections, we describe both the process and the tool.

2.1. The Pedagogical Decision-Making Process (PDMP)

The PDMP was created to augment teachers' (and tutors') decision-making capability, helping them use educational data from online learning environments to: (1) search for relevant pedagogical situations in the learning environment (using

educational data); (2) extract and visualize patterns and trends, from these data, in order to understand what is happening with their students/groups; (3) create (personalized) study plans and other (personalized) pedagogical recommendations, delivering them to the target students; (4) evaluate the impact of these plans and recommendations, i.e., check whether they helped the students or not. The PDMP [Paiva and Bittencourt 2017a] is a cyclical process, consisting of two phases: the construction phase and the execution phase. The construction phase is where the collaborative interaction between human and computational intelligences happen, according to the following steps:

1. Teachers (human intelligence) set the characteristics of the (pedagogical) situation they wish to search in the learning environment. The T-Partner (artificial intelligence) filters the educational data and sets the way these data are analysed (in the next step), based on these user-defined characteristics. For example: a teacher may wish to evaluate students' interactions with the videos-classes in the previous 15 days.

   - Teachers must also define (choose specific values) if the results are below, around or above the expected. The T-Partner labels the data, in preparation for the classification task. For example: the teacher, from the previous example, defines that, for the previous 15 days, if a student watched 2 video-classes, or less, it is an amount of interaction below the expected (low interaction); if a student watched between 3 and 5 video-classes, it is an acceptable amount of interaction (moderate interaction); and if a student watched more than 5 video-classes, it is an acceptable amount of interaction (high interaction).

2. The T-Partner applies a classification algorithm[3] on the filtered and labeled data, returning the result as a visualisation. Teachers interact with the visualisation in order to understand the information displayed in it to guide their decisions. For example: the teacher may correlate performance (higher grades in the course) with interactions with video-classes and noticed that those with higher grades were the students who properly watched the video-classes.

3. Teachers create activities for the students in the low, moderate or high interactions group. These activities contain 1 or more resources from the online learning environment. The T-Partner delivers these activities to the targeted (group of) students. For example: For the low interaction group, the teacher recommends them to watch, at least, 4 more video-classes and solve 8 related problems in the next 5 days. For the moderate interaction group, the teacher recommends them to watch, at least, 2 more video-classes and solve 4 related problems in the next 3 days. For the high interaction group, the teacher challenges them recommending some difficult problems.

4. Teachers define the criteria to consider the recommended activities successful. The T-Partner registes the students' previous performance (on the topic the activity is based) and monitoring the deadline to finish the it. For exemple: the teachers defines a minimum adherence of 66% (i.e. 2/3 of the students

---

[3]The algorithm will depend on the data and the objective of the analysis.

should do what they were recommended to do) and an increase of, at least, 30% in the students' performance regarding the new problems solved.

- When the activity reaches its due date, the T-Partner compares the performance of the student before and after receiving the activity/recommendation, in order to check whether the situation is solved or not.

We successfully applied the PDMP to: (1) evaluate the effectiveness of gamification elements in educational environments [Paiva et al. 2015a]; (2) Measure the differences of male and female students' interactions in an OLE [Sivaldo J. de Santana 2016]; (3) improve students' interactions in an OLE [Paiva et al. 2015b]; [Paiva and Bittencourt 2017b]; Among other applications. Manually following the PDMP requires proficiency with both pedagogy and technology. For this reason, we created an authoring tool called T-Partner (Teachers' Partner) that implements it, making it easier for teachers to follow the process.

2.2. The Teachers' Partner (T-Partner)

The T-Partner is a tool that implements the PDMP. It must be integrated to a learning environment and granted access to its data (user-user, user-content and user-environment interactions), which is used to support pedagogical decision-making. The integration between T-Partner and a learning environment, basically, happens according to the following steps: (1) learners interact with an online learning environment (OLE); (2) these interactions generate educational data that are stored in the OLE's database; (3) these data are retrieved and processed by the T-Partner; (4) the processing results are used to inform teachers about pedagogical situations occurring in the OLE; (5) instructors use this information to make pedagogical decisions; (6) these decisions use the educational resources from the OLE; (7) the decisions should consider the OLE's interface capabilities; (8) the decisions are sent to the specific (groups of) students.

2.3. Search for Pedagogical Situations

In this step teachers choose and detail the kind of situation they wish to search on the educational data. In the used version of the T-Partner, the available options were: (1) analyse Interactions: the objective is to measure if students are interacting with particular resources as expected (by the teacher); (2) Overall Data: the objective is to have an overall idea of the course's resources and interactions; (3) Correlation: the objective is to measure the the impact of an educational resource on the students' results; (4) Response Time: the objective is to check if students are solving the problem below (guessing?) or above (doubts?) the average response time.

Figure 1 shows how teachers define the data they want in the "Analyse Interactions" option. Teachers need to select the course and group, a period of time and the resources they wish to consider in the analysis. These options will personalize the data extraction and the data processing. In this case, the teacher chose to analyse the students' interactions from a Geography course (group named 2M2), from 01/01/2017 until 07/02/2019, considering the following resources: texts read, videos watched and problems solved.

Figure 1. The options to select the data to analyse the students' interactions.

2.4. Extract and Visualize Patterns and Trends

Figure 2a shows the educational data selected by the teacher (see Figure 1). The results are coloured as follows: 0 interactions with a resource are coloured *red*; 1 to 15 interactions are coloured *yellow*; More than 15 interactions are coloured *green* (based the RAG colours technique [Alexander et al. 2014]). Teachers can update the expected values and/or apply filters on the data (for example, they can hide registries with missing values). A bar graph below shows the results grouped by resources. These results suggest that students are solving the problems, but they are not interacting with the texts and videos.
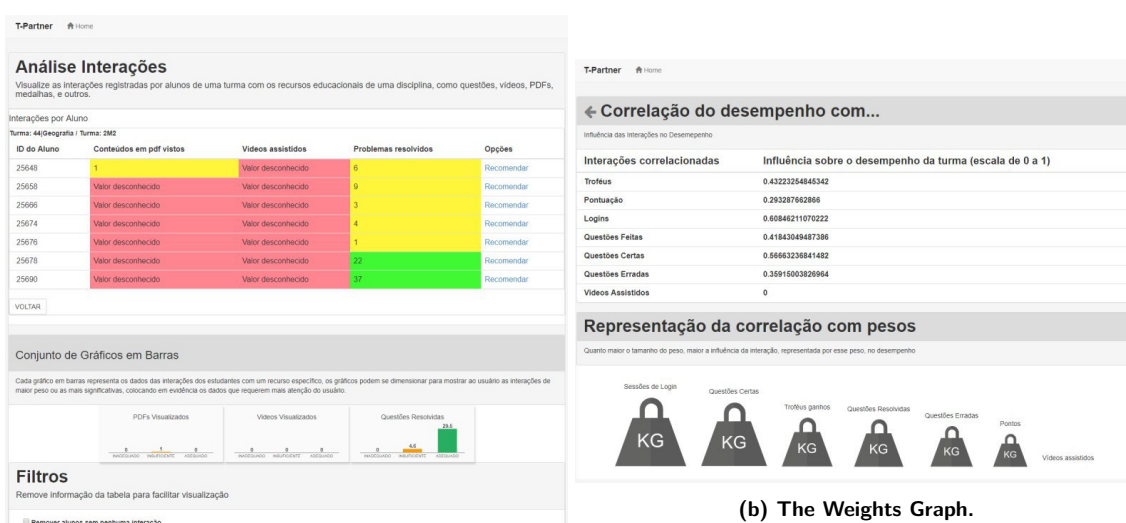
(a) RAG Colors and the Bar Graph.

(b) The Weights Graph.

Figure 2. Visualizations of the students' interactions.

Other visualizations used are the line graph (Figure 3a), that displays the

compound amount of interactions the students made during a period of time. The weights visualization that shows the correlation between the interactions with some resources and the students' results.
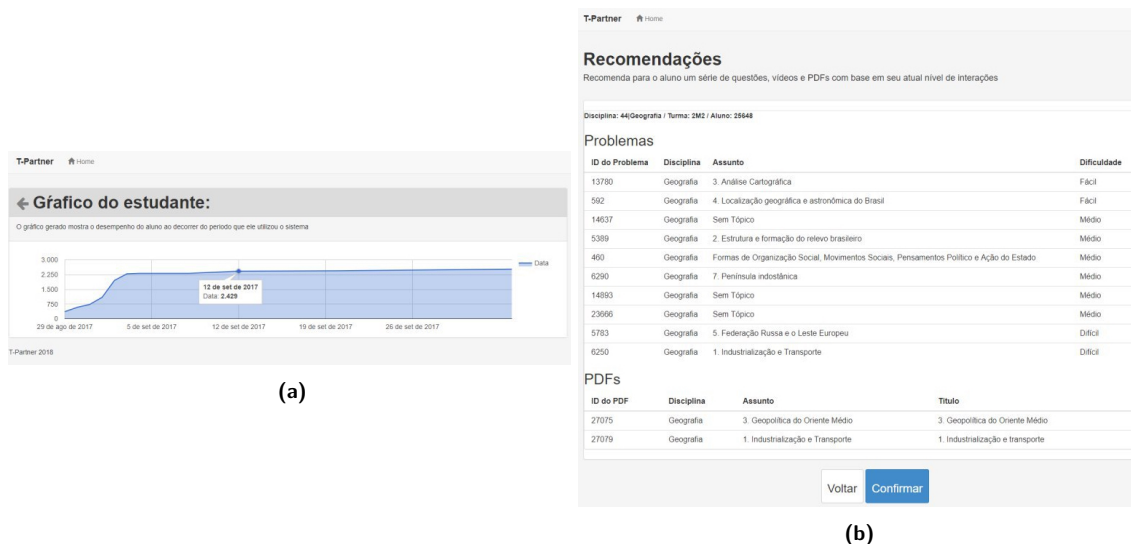

Figure 3. The Line Graph and a Study Plan Created by the teacher.

2.5. Create Study Plans

Figure 3b shows a study plan created by a teacher. In this case, a particular student (#25648, the first student in Figure 2a) was recommended to solve 10 specific problems and then read two texts.

Regarding the evaluation whether these recommendations helped their students or not, we were not able to check it, because we used data from a course that was already over. This way, there was no active students to follow the study plans.

3. Experimentation

The experiment was based on the silent dog protocol [Hayes et al. 1998], where participants were asked to use the T-Partner, acting as teachers managing their courses. They should talk aloud while performing nonverbal tasks [Arntzen et al. 2009], while guided by an assistant, whose function was to verify if the participants understood the tasks[4]. The objective was to check if T-Partner assisted users to: (1) extract relevant information from educational data; (2) understand what is happening with the students and (3) recommend a series of activities to help students (a study plan), based on what teachers detected was happening. We used data from a year-long online course held in 2017 with 10573 students from Brazil.

The tasks simulated the use of T-Partner for the analysis of the students' interactions[5]. After that, participants answered a questionnaire, stating their perceptions, in a Likert scale[6], about T-Partner.

---

[4]After the participants spoke out loud their perceptions and what they should do in each task.

[5]Interaction is an action or influence, either mutual or reciprocal (Merriam-Webster dictionary). In an online learning environment, these interactions can be: between users, between user and content, and between user and the environment [Rossi et al. 2011].

[6]0 = No Answer; 1 = Totally Disagree; 2 = Disagree; 3 = Indifferent; 4 = Agree and 5 =

4. Results and Discussion

We collected the following data from the 24 participants: (1) time to complete a task; (2) Doubts about a task; (3) Correctness of a task and; (4) Perceptions about the T-Partner. The objective was to analyse: (1) whether the T-Partner was easy to use or not; (2) whether the tasks were clear and understandable or not; (3) the participants' perceptions about the T-Partner, in order to check if the tool was useful and easy to use, and check any need for improvement.

4.1. Tasks Analysis

For each task, we measured the time to complete a task (the average, the median and the standard deviation). Small differences between average and median, and a small standard deviation indicate little variation in participants' time to complete a task, suggesting its completion was intuitive for most participants. We also measured the percentage of participants who had doubts and the percentage of participants who correctly completed each task. A high percentage of doubts with a low percentage of completion, indicates the task was difficult to do, difficult to understand or the T-Partner did not provide the appropriate assistance. In Table 1, we show these measures and we notice that:

- All participants were able to complete tasks 01, 02, 05, 07, 11 and 12. With a few of them showing doubts in task 12 (12.50%). This suggests these tasks were clear and appropriate, and the T-Partner provided the necessary assistance to complete them.
- At least 90% of the participants were able to complete tasks 08, 09, 14, 16 and 17. Just a few of them had doubts (16.67%) in task 14. This suggests these tasks are appropriate, and the T-Partner assisted to complete them. However, it is necessary to investigate if the tasks were not clear or if the T-Partner needs improvement.
- The results for tasks 03, 04, 06, 10, 13, 15 and 18 showed a high percentage of doubts or a low percentage of completion (or both). Tasks 03, 06 and 10 deal with the colour scheme used. Tasks 04, 13, 15 and 18 deal with interpreting the results in the visualisations presented by the T-Partner. These results echo the participants' perceptions analysis, strongly suggesting the visualisations need to be improved.

4.2. Participants' Perceptions Analysis

In the end, 20 (twenty) questionnaires were completed answered. We analysed the participants' perceptions regarding the use of the T-Partner with the following statements: (1) the application interface was intuitive; (2) The use of red, yellow and green colours was interesting and appropriate; (3) I would trust a student's analysis generated by this tool; (4) T-Partner appears useful; (5) The features have good design and style; (6) The features are easy to use; (7) The bar graph visualisations is easy to analyse; (8) The weights visualisation is easy to analyse; (9) The bar graph visualisation is clear and comprehensible; (10) The weights visualisation are clear and comprehensible; (11) It would be easy to make decisions based on the

---

Totally Agree

Task	Time (Avg)	Time (Med)	Time (StD)	Doubts	Completion
Task 01	3.75	3	2.47	0%	100%
Task 02	55.12	54.5	12.07	0%	100%
Task 03	73.05	65	34.97	50%	70.83%
Task 04	56.29	51.5	44.84	41.67%	79.17%
Task 05	25.62	18	33.36	8.33%	100%
Task 06	41.29	31.5	24.34	29.17%	54.17%
Task 07	6.29	6	2.71	0%	100%
Task 08	18.37	14.5	12.72	4.17%	91.67%
Task 09	7.78	7	4.08	0%	95.83%
Task 10	39.87	37	19.22	16.67%	66.67%
Task 11	8.71	3.5	11.66	4.17%	100%
Task 12	30.21	28.5	13.51	12.50%	100%
Task 13	40.09	30	25.18	12.50%	75%
Task 14	41.25	30.5	34.25	16.67%	95.83%
Task 15	65.33	68.5	34.51	50%	66.67%
Task 16	70.39	67	34.91	8.33%	95.83%
Task 17	33.71	30.5	30.09	8.33%	95.83%
Task 18	67.54	51	52.24	25%	83.33%

Table 1. The analysis of the tasks.

bar graphs visualisation; (12) It would be easy to make decisions based on weights visualisation; (13) I Would use this tool if it were available; (14) It was interesting to do what was proposed to do using the tool. The results for the participants' perceptions are summarised below:

- All participants agree or totally agree that it was interesting to use the T-Partner. This suggests the T-Partner was well-accepted by the participants.
- At least 80% of the participants agree or totally agree that: (1) the application interface was intuitive; (2) they would trust the analysis generated by the T-Partner; (3) the T-Partner appears to be useful; (4) the features were easy to use; (5) the bar graph was easy to analyse; (6) the bar graph was clear and comprehensible; (7) the bar graph helped them making decisions and; (8) they would use the T-Partner if it was available. This suggests the T-Partner was perceived as useful, helpful and, its analysis, trustworthy.
- At least 70% of the participants agree or totally agree that: (1) the features had good design/style; (2) the weights visualisation was easy to analyse and; (3) the weights visualisation was clear and comprehensible. This suggests the weights visualisation is easy to use, but there is space for improvement.
- At least 60% of the participants agree or totally agree that: (1) the color scheme (to show the students' results) was interesting and appropriate and; (2) the weights visualisation (Figure 2b) helped them making decisions. These results are consistent with the tasks analysis, where we noticed that tasks involving interpreting the color scheme, the weights visualisation and the line chart (Figure 3a) took a longer time and generated more doubts.

5. Conclusion

There is an increasing demand for online education, which are offered via online learning environments. These environments generate data in vast amounts and diversity. Teachers must use these data to provide the quick feedback students require, but they lack the abilities to quickly, and proficiently, deal with these data. However, many years of experience start to point to a new direction where Artificial Intelligence and Human Intelligence should complement each other. In order to "join these intelligences", we proposed an authoring tool where data science is used to augment the teachers' capability to deal with these data, allowing them to make better and informed decisions, tailored to assist their groups/students.

We evaluated T-Partner with 24 participants, using tasks that are related to what online teachers would need to manage their courses. We noticed participants had trouble on the tasks that deal with the colour scheme used (RAG colours) and those to interpret some visualisation. For all the other tasks, participants were able to complete them with none or a few doubts and in an appropriate amount of time, suggesting the T-Partner assisted them completing these tasks. Regarding their perceptions, participants stated lower perceptions for the colour scheme and the weights visualisation and the line chart, which signals these parts needs improvement. However, the T-Partner was well-accepted and was perceived as useful, helpful and its analysis (the results it presents to teachers) was trustworthy.

We believe the number of participants (24) may be a limiting factor for the generalisation of the results. Another fact is that we were not able to evaluate the impact of the recommendations (study plans) in the students' performance. This was due to the fact that we used data from a completed course. So, there was no active students to follow these study plans.

References

Alexander, M., Yuk, M., and Diamond, S. (2014). *Data Visualization for Dummies*. John Wiley & Sons.

Arntzen, E., Halstadtro, L.-B., and Halstadtro, M. (2009). The "silent dog" method: Analyzing the impact of self-generated rules when teaching different computer chains to boys with autism. *The Analysis of verbal behavior*, 25(1):51–66.

Baker, R. S. (2016). Stupid tutoring systems, intelligent humans. *International Journal of Artificial Intelligence in Education*, 26(2):600–614.

Boettcher, J. V. and Conrad, R.-M. (2016). *The online teaching survival guide: Simple and practical pedagogical tips*. John Wiley & Sons.

Burden, P; Fuller, A. (2014). The teacher as a decision maker. In *Methods for Effective Teaching Meeting the Needs of All Students*.

Chou, C.-Y., Huang, B.-H., and Lin, C.-J. (2011). Complementary machine intelligence and human intelligence in virtual teaching assistant for tutoring program tracing. *Computers & Education*, 57(4):2303–2312.

Dağ, F., Durdu, L., and Gerdan, S. (2014). Evaluation of educational authoring tools for teachers stressing of perceived usability features. *Procedia-Social and Behavioral Sciences*, 116:888–901.

Hayes, S. C., White, D., and Bissett, R. T. (1998). Protocol analysis and the "silent dog" method of analyzing the impact of self-generated rules. *The Analysis of Verbal Behavior*, 15(1):57–63.

Kopcha, T. J. (2012). Teachers' perceptions of the barriers to technology integration and practices with technology under situated professional development. *Computers & Education*, 59(4):1109–1121.

Li, C.-S. and Irby, B. (2008). An overview of online education: Attractiveness, benefits, challenges, concerns and recommendations. *College Student Journal*, 42(2).

Mandinach, E. B. and Jackson, S. S. (2012). *Transforming teaching and learning through data-driven decision making.* Corwin Press.

Mouza, C. (2018). The role of teachers in teaching and learning with technology. *Contemporary Issues in Technology and Teacher Education*, 18(3):491–493.

of Education., U. D. (2014). Learning technology effectiveness. Technical report, Office of Educational Technology. Retrieved from https://tech.ed.gov/wp-content/uploads/2014/11/Learning-Technology-Effectiveness-Brief.pdf.

Paiva, R., Barbosa, A., Batista, E., Pimentel, D., and Bittencourt, I. I. (2015a). Badges and xp: An observational study about learning. In *Frontiers in Education Conference (FIE), 2015. 32614 2015. IEEE*, pages 1–8. IEEE.

Paiva, R. and Bittencourt, I. I. (2017a). The authoring of pedagogical decisions informed by data, on the perspective of a mooc. In *Anais dos Workshops do Congresso Brasileiro de Informática na Educação*, volume 6, page 15.

Paiva, R. and Bittencourt, I. I. (2017b). The authoring of pedagogical decisions informed by data, on the perspective of a mooc. In *Anais dos Workshops do Congresso Brasileiro de Informática na Educação*, volume 6, page 15.

Paiva, R. and Bittencourt, I. I. (2017c). Helping mooc teachers do their job. In *Researcher Links Workshop: Higher Education for All*, pages 52–67. Springer.

Paiva, R. O. A., Bittencourt, I. I., da Silva, A. P., Isotani, S., and Jaques, P. (2015b). Improving pedagogical recommendations by classifying students according to their interactional behavior in a gamified learning environment. In *Proceedings of the 30th Annual ACM Symposium on Applied Computing*, pages 233–238. ACM.

Rossi, D., Beer, C., Janse van Rensburg, H., Harreveld, R., Danaher, P. A., and Singh, M. (2011). Learning interactions: A cross-institutional multi-disciplinary analysis of learner-learner and learner-teacher and learner-content interactions in online learning contexts.

Seaman, J. E., Allen, I. E., and Seaman, J. (2018). Grade increase: Tracking distance education in the united states. *Babson Survey Research Group*.

Sivaldo J. de Santana, Ranilson Paiva, I. I. B. e. a. (2016). Evaluating the impact of mars and venus effect on the use of an adaptive learning technology for portuguese and mathematics. *The 16th IEEE International Conference on Advanced Learning Technologies - ICALT2016*.

Stephenson, J. (2020). *Teaching & learning online: new pedagogies for new technologies.* Routledge.