The Authoring of Pedagogical Decisions Informed by Data, on the Perspective of a MOOC

Ranilson Paiva[1]**, Ig Ibert Bittencourt**[1]

[1]Computing Institute - Federal University of Alagoas (UFAL)
Av. Lourival Melo Mota, s/n - Tabuleiro do Martins, CEP: 57072-970 Maceió/AL - Brasil

`ranilsonpaiva@ic.ufal.br`, `ig.ibert@gmail.com`

Abstract. *Improvements in digital information and communication technologies (DICT) brought a renewed global interest in distance learning. However, there is no equivalent increase in support for the instructors responsible for maintaining such courses, which is evidenced by the large number of dropouts and failures in such courses. In order to provide this support, instructors would have to: (1) to discover situations of pedagogical interest occurring in their courses; (2) understand these situations; (3) make decisions to address them; (4) monitor and evaluate the impact of the decision made. However, instructors do not master these abilities, nor is it practical or appropriate to ask them to do so. We propose a process, and an authoring solution (T-Partner) that implements it, to guide instructors through these 4 steps, promoting cooperation between artificial intelligence and human intelligence. We conducted experiments to evaluate the process and the authoring solution. The results indicate that the process helped instructors make better pedagogical decisions and the authoring solution was positively perceived by the instructors.*

1. Introduction

Improvements in digital information and communication technologies (DICT) brought a renewed global interest in distance learning [Allen and Seaman 2014, de Educação à Distância 2016]. It paves a new path for education, where DICTs are used to mediate learning from anywhere, at any time and for anyone (*AAA Learning*) [Bittencourt et al. 2009].

In this context, MOOCs[1] are gaining attention due to their potential to "propagate and democratize education" [Siemens and d Baker 2012, Baker 2016]. However, MOOCs bring new challenges for the instructors responsible for maintaining them. For instance, researchers[2] [Onah et al. 2014] estimate that, on average, 85% of learners dropout from MOOCs, justifying it as **a lack of support by their instructors**.

In order to provide this support, instructors would need to analyze data from the learners' interactions in search for educational profiles and relevant information for the teaching and learning process [Romero and Ventura 2016], to support pedagogical decision-making [Schildkamp et al. 2012]. However, instructors are not, normally, trained for that and they do not receive appropriate technology aid to deal with these data [Mandinach and Jackson 2012], giving rise to barriers against technology integration and

---

[1]MOOCs: Massive Open Online Courses.
[2]Check: `http://www.katyjordan.com/MOOCproject.html`

instructional practice [Kopcha 2012]. We, then, notice a need for direction and assistance for these instructors [Mandinach and Jackson 2012].

Some authors state that attempts to automatize instructors' tasks are there for about 70 years without huge successes [Baker 2016]. They propose *machine intelligence* should complement *human intelligence* [Chou et al. 2011, Baker 2016, Romero and Ventura 2016], which is aligned with our proposal to solve the *lack of support* mentioned previously. Based on that, we asked the following question: *How can we use digital information and communication technologies to guide instructors' pedagogical decision-making?*

In order to address this question, we defined the following research objectives: (RO1) Design a process that systematically guides pedagogical decision making; (RO2) Create an authoring solution, based on the process, to build informed pedagogical decisions, in course time, for online learning environments; (RO3) Enable the authoring of pedagogical decisions at different levels of granularity without affecting the users' perception about the authoring tool; and (RO4) Define visualizations to allow teachers and tutors to use techniques of educational data processing and analysis, informing them of relevant pedagogical situations occurring during course time.

These objectives align our proposal to the philosophy mentioned[3]. For that, we created a process (the Pedagogical Decision-Making Process - PDMP) to guide pedagogical decision-making for instructors of online learning environments (OLEs), and an authoring solution (T-Partner) that implements the process. Initially, we manually applied the process in an online course [citation removed due to blind review], with the intention of helping instructors use the educational data generated by the studied learning environment to: (1) discover situations of pedagogical interest occurring in their courses; (2) understand these situations; (3) make decisions to address them; (4) monitor and evaluate the impact of the decisions made. These were the 4 steps of the PDMP

The second part of our proposal, the authoring solution (T-Parner, in a proof-of-concept stage), implements the process, allowing instructors to analyse educational data (without having to learn it), in order to detect and address pedagogical situations occurring within a course, such as: detecting performance problems within a group, comparing groups, predicting dropouts and predicting inadequate performance in a group. Our proposal is based on informed decision-making by educational data, promoting cooperation between artificial intelligence and human intelligence.

In order to test our proposal, we conducted two experiments. In the first (a controlled experiment), we manually applied the PDMP to help teachers improve their learners' interactions. In the second (a survey), we invited teachers and tutors (instructors) to interact with the T-Partner's interfaces and evaluate if they allowed instructors to follow the PDMP. The results for the first experiment was an increase (more than 10%) in the learners' interactions. In the second experiment, participants showed positive perceptions about T-Partner's interfaces.

In search of related works, we carried out a systematic literature review and noticed the non-existence of works that brought data-based authorship to support the pedagogical decision making of teachers/tutors, on course time.

---

[3]*Machine intelligence* complementing/augmenting *human intelligence*

In this article, we focused on the essential parts of the work. More details can be found in the thesis. This article is organized as follows: in Section 2, we present the main ideas about our proposal. In Section 3, we describe the two experiments we ran to evaluate our proposal (one for the process and the other for the authoring solution). In Section 4, we summarize the results of the experiments and, finally, in Section 5, we draw our conclusions, based on the results of our experiments.

2. Proposal

In this section, we present the Pedagogical Decision-Making Process (PDMP) and the authoring solution that implements it (T-Partner). Our intention is to allow instructors to use educational data to investigate their students' learning experiences, transforming the online learning environment (OLE) into a lab where the instructors' hypothesis may be tested. These hypothesis appear from the analysis of pedagogical situations and should guide instructors into testing different interventions in order to address them.

2.1. Pedagogical Decision-Making Process (PDMP)

The process is based on *scenario planning* [Chermack et al. 2001] and was created [citation removed due to blind review] to guide online learning environment instructors to: (1) to discover situations of pedagogical interest occurring in their courses; (2) understand these situations; (3) make decisions to address them; (4) monitor and evaluate the impact of the decisions made.

We consider situations with pedagogical value, like those related to: interactions of the (groups of) students, prediction of dropouts and performance problems or good practices. The instructors should define the context in which they want to search for the occurrence of such situations. For example: predict low performance (before a test) of a 9th grade mathematics class, considering the students' interactions in the last 30 days. The initial results are data from the students' interactions during the defined period, which is processed in search of patterns and information that help instructors creating pedagogical hypothesis, guiding their actions (interventions they may test using the system).

These interventions (pedagogical decisions), use the educational resources available in the OLE (preferably) or suggest links to external resources, which are recommended to the target (group of) students, considering the instructors' definitions. Finished the time to follow these recommendations, is it time to evaluate the adherence[4] and the expected results[5].

The PDMP is a cyclical process and consists of two phases: the construction phase and the implementation phase. In the construction phase, there is the collaborative interaction of human and artificial intelligences, to create computational artefacts called *Pedagogical Decision Capsule*. These artefacts are the encapsulation of the instructors' decisions about: (1) which pedagogical situation they want check in the learning environment; (2) what decisions to take to address such situation; and (3) how they wish to assess whether the decisions have been effective or not.

The PDMP was used in some researches to evaluate the effectiveness of gamification elements in educational environments [citation removed due to blind review], to

---

[4]The percentage of students who followed the recommendation.

[5]The results the instructors wish to see in order to consider the recommendations successful.

measure the differences of male and female students' interactions in an OLE [citation removed due to blind review], to improve students' interactions in an OLE [citation removed due to blind review] among other uses[6].

2.2. T-Partner

Following the process (PDMP) manually is not an easy task, being subject to human failures. We created a technological solution to support instructors authoring *Pedagogical Decision Capsules*, which we called T-Partner[7].

The T-Partner needs to be integrated to the learning environment in order to access data about: learners, educational resources available, pedagogical approaches offered, interactional data (user-user, user-content and user-environment), and other data to support pedagogical decision making. Succinctly, this is done as follows: (1) learners interact with the 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 instructors 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 learners; (9) the T-Partner measures the effectiveness of the decisions.

The T-Partner was, initially, idealized in two versions (granularities):

1. *Light Weight*: an easy to use version, but more limited. This version is for users with little experience with computers and for the creation of pedagogical decision capsules (PDC) with simple content or in a faster way. It can also be used as an entry version (beginners).
2. *Heavy Weight*: a version with more features, but higher complexity. It is intended to allow greater detailing and control in creating the PDCs.

2.3. Implementing the PDMP

As mentioned in subsection 2.1, the PDMP contains two phases. In the *construction phase*, instructors interact with T-Partner to define the "hypothesis" they want to check in the educational data, investigate evidences about this "hypothesis", decide what to do based on the evidences and check whether their decisions were effective or not, thus following the four steps in this phase. In the *execution phase*, the successful decisions (those that were effective) are encapsulated and executed, automatically, by the T-Partner. Bellow, we better describe the four steps from the construction phase:

1. **Step 1: Define the Pedagogical Situation**: in this step, instructors choose what kind of pedagogical situation (among those available) they want to investigate in the OLE. Instructors also label the range of results as inappropriate, insufficient and appropriate (level of learning). The system searches the data requested by the teacher, regarding the selected situation, and classifies it.

---

[6]Although not published, yet, the process was successfully used to recommend areas learners should focus to improve their writing performance. Some other works in progress are: detecting and recommending actions to disengaged learners, recommending appropriate educational resources to practice an specific math topic, etc.

[7]Teachers' Partner

2. **Step 2: Investigate Pedagogical Situation**: in this step, instructors select filters they want to apply in the data, as well as how they wish to visualize the data processing result. The system processes the data using an algorithm associated with the chosen pedagogical action and generates the visualization of the result, providing instructors information to decide what to do.

3. **Step 3: Define Pedagogic Decision**: in this step, instructors decide what tasks (among those available in the OLE) learners should do, according to their levels of learning (inadequate, insufficient or appropriate).

4. **Step 4: Define Assessment**: In this step, instructors set the desired percentage for adherence and the expected outcome from those who followed the recommendations, for each level of learning.

3. Experimentation

In this section, we present the experiments to evaluate the PDMP and the T-Partner.

3.1. Experiment 1

In this experiment, the PDMP was manually applied on a MOOC (live). Its objective was to evaluate whether the process was effective, or not, in helping instructors make better decisions using educational data, meeting our first research objective (RO1) (check Section 1). Below, we describe the steps followed in this experiment:

- **Step 1 - Define a Pedagogical Situation:** we collected data from 36 learners who were active for at least 6 months. The pedagogical situation of interest was to identify differences in their interactions within the MOOC.
- **Step 2 - Investigate a Pedagogical Scenario:** we processed the educational data and classified the learners' interaction in: individual learning (those that direct influences learning, like: watching videos, solving problems, etc.), collaboration (those that helps improving the MOOC or helps other learners, like: reporting errors, rating videos, etc.), gamification (those focused on the game elements available, like: earning badges, completing missions, etc.) and social (those focused on interacting with other learners without the need of pedagogical gains, like: chatting, making friends, etc.). We calculated a score for each type of interaction, for each learner, plotted it on a spider web graph, and presented it to the instructors.
- **Step 3 - Define Recommendations:** after studying the visualizations (spider web graph), instructors were informed about the educational resources[8] available in the MOOC and were told to use appropriate resources to help learners. The decision was to recommend learners to interact more with the types of resources they had fewer interactions, and gave learners 30 days to do it.
- **Step 4 - Define Success Criteria:** as a success criteria, instructors defined an adherence of 73%[9] and, as the outcome, an increase, of at least 10%, in the learners' interactions with the recommended types of resources.

---

[8]Educational resources are resources for teaching and learning, including complete courses, teaching material, modules, textbooks, videos, quizzes, educational software and any other tools or techniques used to support access to knowledge[Foundation 2013].

[9]At least 26 learners should follow the recommendation

3.2. Experiment 2

In this experiment, we invited instructors (professors, teachers and tutors) to evaluate the T-Partner, regarding the following metrics: (1) utility perception - UP; (2) ease of use perception - EUP; (3) attitude towards use - ATU; (4) intention to use - IU; (5) correctness on the tasks requested - CTR; (6) perception about the Traffic Light Metaphor - TLM; and (7) perception about the terms that categorize student learning - TSL [Teo 2011].

Participants were randomly assigned to one of three possible versions of T-Partner: (1) lightweight; (2) heavyweight; (3) manual (spreadsheet). The lightweight (LW) and heavyweight (HW) are versions with different granularities (complexity) of the T-Partner, that is, these versions explore different combinations between artificial intelligence and human intelligence. The *spreadsheet* (SS) is a control version that simulates authoring, following the PDMP, in an ultra simplified way, using a familiar resource (spreadsheets). The objective was to evaluate the perception of the versions of the T-Partner (unknown) compared to the control version (ultra simple and familiar).

Each version covered the 4 steps of the PDMP (see subsection 2.1). At each step, participants performed tasks that altogether create a pedagogical decision capsule (PDC) and then respond a questionnaire, expressing their perceptions about (1) utility; (2) ease of use; (3) attitude towards use; (4) intention to use; (5) colours used; and (6) terms used.

After a month, we had 175 valid records (obtained after removing incomplete and testing records). Finally, we performed the appropriate data analysis [Jain 2010].

4. Results and Discussion

In this section, we summarize the results based on our research objectives.

4.1. RO1: Design a process that systematically guides pedagogical decision making

During a six-month period, learners freely interacted with the MOOC, we noticed that the 36 selected learners, on average: made 0.25 friends per month; Rated 0.65 videos per month; Correctly answered 198.20 problems (33.03 problems per learner, per month), and had considerably more problems solved incorrectly than problems solved correctly. These results suggest that these learners were experiencing difficulties.

The learners received personalized recommendations (based on their interactions), and had 30 days to finish them. After this period, the results were: 0.30 friends; 0.70 videos rated; 30.96 correctly answered problems and 35.02 problems answered incorrectly.

We used the Wilcoxon test (for two measures, of the same sample, from a non-parametrical population) to check if the learners' interactions before and after the personalized missions, were different (see Table 1).

Table 1. P-values for the (Wilcoxon signed-rank test).

	Friends	Videos	Correct Problems	Incorrect Problems
Wilcoxon Signed-rank	0.001985	0.0006552	0.0002137	0.001656

As a result, we had an increase in the number of the desired interactions after the instructors' interventions using the PDMP. These facts suggest that the process helped instructors make better decisions and it is a step towards human intelligence augmented by the artificial intelligence approach. We need more experiments, but in the context of this experiment, the PDMP guided pedagogical decision-making, meeting our first research objective.

4.2. RO2: Create an authoring solution, based on the process, to build informed pedagogical decisions, in course time, for on-line learning environments

This is part of experiment 2. We asked participants to accomplish tasks that, in the end, are combined to create a *Pedagogical Decision Capsule*. We measured the participants' precision in accomplishing the tasks. The comparison between authoring versions HW and LW are in Table 2.

Table 2. Wilcoxon-Mann-Whitney test - Precision in accomplishing tasks.

STEPS	P-VALUE	GREATER PRECISION
1	0.0001984*	*HW*
2	0.08513	No Difference
3	0.01698*	*HW*
4	0.0004147*	*LW*
All Steps	0,02922*	*HW*

The conclusion is that the HW version promoted greater precision in the accomplishment of the tasks. The reasons are not clear, but we believe that the greater detailing and control that instructors had using the HW version, helped them accomplish the proposed tasks. In future works, we will study the reasons for such outcome.

4.3. RO3: Enable the authoring of pedagogical decisions at different levels of granularity without affecting the users' perception about the authoring tool

We evaluated if the authoring solution allows pedagogical decisions to be made at different levels of granularity (different complexity levels). For that, we considered the participants' perceptions about: (1) utility perception - UP; (2) ease of use perception - EUP; (3) attitude towards use - ATU; (4) intention to use - IU. Participants scored each criteria (Likert scale from 0 to 6). The results (mean) are presented in Table 3.

Table 3. Average scores of the participants' perceptions.

AUTHORING VERSION	UP	FU	IU	ARU
Heavy Weight (HW)	3.9	3.8	4.1	3.8
Light Weight (LW)	4.1	3.9	4.2	4.0
Spreadsheet (SS)	3.7	4.1	3.9	3.7

Regarding **utility perception**, **ease of use perception**, **intention to use**, and **attitude towards use**, we can conclude that there is a statistical similarity between the three

versions used in the experiment [10]. The mean of the responses are higher or very close to 4 ("I somewhat agree") for all metrics, at all stages and for all versions. These results indicate that the participants' perceptions were positive and similar, concluding that the authoring solution allows pedagogical decisions to be made at different levels of granularity without affecting the participants' perceptions about the authoring.

4.4. RO4: Define visual metaphors to allow teachers and tutors, transparently, to use techniques of educational data processing and analysis, to inform them of relevant pedagogical situations occurring during course time

We evaluated if the visualizations used helped instructors understand the result of the data analysis. For that, we analysed the *precision* of the participants' answers in step 2 (where the visualizations were presented to help participants). We also analyzed the participants' perceptions regarding the *use of RAG colours*[11] and the *terms used*[12].

We created 4 visualizations to help instructors understand the results of the educational data analysis (in step 2). One was used in the *heavy weight* version (Viz1) and the other 3 were used in the *light weight* version (Viz2, Viz3 and Viz4). In Table 4, we compared the visualizations among themselves, considering how well they helped instructors. The results show that Viz1 and Viz2 promoted greater precision in the participants' answers to the tasks (i.e.: helped instructors understand the result of the data analysis). The order in terms of assistance is as follows: Viz1 > Viz2 > Viz3 = Viz4.

Table 4. Comparison between visualizations

VIZ	Wilcoxon Test	BONFERRONI	GREATER PRECISION
Viz2 vs. Viz3	0.0001185024	0.0007110142*	Viz2
Viz2 vs. Viz4	7.216746e-06	4.330048e-05*	Viz2
Viz2 vs. Viz1	1.00	-	No difference
Viz3 vs. Viz4	0.01171951	0.07031707	No difference
Viz3 vs. Viz1	3.330479e-05	0.0001998287*	Viz1
Viz4 vs. Viz 1	1.439644e-06	5.758577e-06*	Viz1

Considering the *use of RAG colours*, we compared all three versions (Table 5) and the participants' perceptions was not statistically different. This is the desired result, once that the use of RAG Colours is a resource to support instructors and should, therefore, be consistently provided in all authoring versions.

Finally, considering the *terms used*, the results are shown in Table 6 and suggest that the way the terms were used in the *light weight* version, caused a significantly greater perception among participants'. This result is not aligned with our expectations, demanding future improvements regarding this matter.

---

[10]There was a small difference in step 4, regarding the utility perception and the intention to use.

[11]We used the colours Red, Amber and Green to categorize learners according to their performance/pedagogical situation (inadequate: red; insufficient: amber/yellow and adequate: green).

[12]We tried to use terms that are familiar to instructors, rather than technical terms, to favour interactions.

Table 5. Participants' perceptions on the use of RAG Colors

Wilcoxon-Mann-Whitney		
STEPS	**P-VALUE**	**GREATER PERCEPTION**
1	0.5728	No difference
2	0.07699	No difference
3	0.1998	No difference
4	0.07501	No difference

Table 6. Participants' perceptions on the terms for the learners' performance.

Wilcoxon-Mann-Whitney		
STEP	**P-VALUE**	**GREATER PERCEPTION**
1	0.4632	No difference
2	0.1593	No difference
3	0.09437	No difference
4	0.0196	*LW*
All	0.02744	*LW*

5. Conclusion

In this paper, we performed 2 experiments, using educational data from a running MOOC (MeuTutor-ENEM). The first was to evaluate the process and the second experiment was to evaluate if the authoring solution could help instructors, supporting pedagogical decision-making (at different levels of granularity and with the use of special colours scheme, RAG colours; data visualization and appropriate terms/expressions).

For the first experiment, the results showed that, following the PDMP, instructors made better decisions that improved students' interactions. For the second experiment, the results demonstrate that the participants were able to perform the demanded tasks and that their responses indicate a positive perception regarding, roughly, all the metrics with statistical significance.

Overall, the results allow us to conclude the PDMP guided teachers through better pedagogical decisions (meeting our first research objective - RO1); the authoring tool conducted teachers/tutors through the PDMP (RO2) at different levels (RO3) and the visualizations and interface resources (RO4) were positively perceived by the users (survey participants).

As future works, we believe it is necessary to research, test and implement more pedagogical situations. Also, it is important to better define the authoring levels, evaluating the trade-off between human intelligence and the artificial intelligence. Finally, we intend to research what is necessary to allow teachers/tutors to use other educational data-mining, learning analytics, data visualization and recommender systems techniques as well as create multiple strategies to address a single pedagogical situation.

References

Allen, I. E. and Seaman, J. (2014). Grade change: Tracking online education in the united states, 2013. *Babson Survey Research Group and Quahog Research Group, LLC. Retrieved on*, 3(5):2014.

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

Bittencourt, I. I., Costa, E., Silva, M., and Soares, E. (2009). A computational model for developing semantic web-based educational systems. *Knowledge-Based Systems*, 22(4):302–315.

Chermack, T., Lynham, S., and Ruona, W. (2001). A review of scenario planning literature. *Futures Research Quarterly*, 17(2).

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.

de Educação à Distância, A. B. (2016). Censo ead br (2015). *Relatório Analítico da Aprendizagem a Distância no Brasil*.

Foundation, T. W. F. H. (2013). White paper: Open educational resources. breaking the lockbox on education. Technical report, The William Flora Hewlett Foundation.

Jain, A. K. (2010). Data clustering: 50 years beyond k-means. *Pattern recognition letters*, 31(8):651–666.

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.

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

Onah, D. F., Sinclair, J., and Boyatt, R. (2014). Dropout rates of massive open online courses: behavioural patterns. *EDULEARN14 Proceedings*, pages 5825–5834.

Romero, C. and Ventura, S. (2016). Educational data science in massive open online courses. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*.

Schildkamp, K., Lai, M. K., and Earl, L. (2012). *Data-based decision making in education: Challenges and opportunities*, volume 17. Springer Science & Business Media.

Siemens, G. and d Baker, R. S. (2012). Learning analytics and educational data mining: towards communication and collaboration. In *Proceedings of the 2nd international conference on learning analytics and knowledge*, pages 252–254. ACM.

Teo, T. (2011). Factors influencing teachers' intention to use technology: Model development and test. *Computers & Education*, 57(4):2432–2440.