Usability Perspective of an Authoring Solution to Assist Pedagogical Decision-Making

**Ranilson Paiva[1], Ig Ibert Bittencourt[1], André Vinicius[1],
Sérgio Amorim[1], Wansel Lemos[2], Diego Dermeval[3], Seiji Isotani[4]**

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

[2]Federal University of Alagoas (UFAL) Campus Arapiraca
Arapiraca/AL - Brasil

[3]Federal University of Alagoas (UFAL) Educational Unity of Penedo
Penedo/AL - Brasil

[4]Institute of Mathematical Sciences and Computing - University of São Paulo (USP)
São Carlos/SP - Brasil

ranilsonpaiva@ic.ufal.br, {ig.ibert, andreducap}@gmail.com

sata@ic.ufal.br, wansel.lemos@arapiraca.ufal.br

diegodermeval@gmail.com, sisotani@icmc.usp.br

Abstract. *There is a quest to provide education from anywhere, at any time and for anyone, using digital information and communication technologies, but there is no equivalent increase in support for the instructors responsible for maintaining such courses, evidenced by the large number of dropouts and failures. We propose an authoring solution to guide pedagogical decision-making in on-line learning environments to help instructors (1) to discover pedagogical situations occurring in their courses; (2) understand these situation; (3) make decisions to address them; (4) monitor and evaluate the impact from decisions made. However, instructors do not master these abilities, nor is it practical/appropriate to ask them to do so. We developed a proof-of-concept version of an authoring solution, named T-Partner, that guides instructors through these 4 steps. We conducted an experiment to evaluate whether it is perceived as useful and easy to use. The results show that its usefulness and ease of use were positively perceived by the instructors.*

Keywords: *pedagogical decision-making process, data-based decisions, authoring tools, on-line learning environments.*

1. Introduction

We are experiencing changes in the educational paradigm. There is a quest to provide education from anywhere, at any time and for anyone (*AAA Learning*) [Bittencourt et al. 2009], using digital information and communication technologies, evidenced by a global interest in distance learning [Chrysafiadi and Virvou 2013,

de Educação à Distância 2016]. Among the types of distance courses, MOOCs[1] are gaining attention due to their potential to "propagate and democratize education" [Siemens and d Baker 2012, Baker 2016].

However, the dynamics of a MOOC brings challenges to the instructors responsible for maintaining it. For instance, in Figure 1 we present the statistical data of a course titled *Bioelectricity: A Quantitative Approach*[Belanger and Thornton 2013]. We notice, among other facts, that only 2,46% of the learners finished the course (earn a certificate). That was not an isolated case. Researchers[2] [Onah et al. 2014, Liyanagunawardena et al. 2014] state that, on average, 85% of learners dropout from MOOCs, justifying it as **a lack of support by their instructors** (teachers and/or tutors).
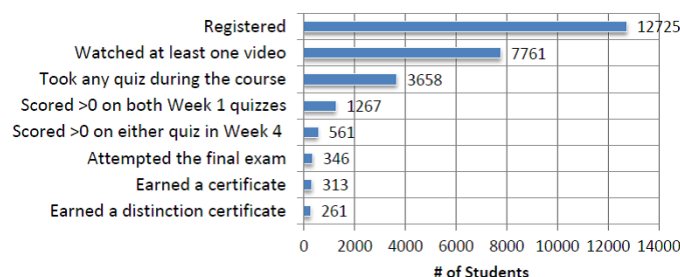
Figure 1. Number of learners per milestone of the course.

Source: Bioelectricity: A Quantitative Approach ([Belanger and Thornton 2013])

In order to provide this support, instructors would need to analyze data from the learner' interactions in search for educational profiles and relevant information for the teaching and learning process [Romero and Ventura 2016], supporting pedagogical decision-making [Bienkowski et al. 2012, Schildkamp et al. 2012]. Instructors do not receive appropriate technology aid to do that [Mandinach and Jackson 2012, Schildkamp et al. 2012], giving rise to barriers against technology integration and instructional practice [Kopcha 2012].

We see the need to direct and aid these instructors [Mandinach and Jackson 2012, Schildkamp et al. 2012], in a way where machine intelligence complements human intelligence [Chou et al. 2011, Baker 2016, Romero and Ventura 2016], raising the following research question: How can we use digital information and communication technologies to guide instructors' pedagogical decision-making?

To address this question, we created a process (the Pedagogical Decision-Making Process - PDMP) to guide pedagogical decision-making for instructors of on-line learning environments, and an authoring solution (T-Partner) that implements this process. Initially, we manually applied the process in an on-line course [Paiva et al. 2016], with the intention of helping instructors use educational data to (1) discover pedagogical situations occurring in their courses; (2) understand these situations; (3) make decisions to address them; (4) monitor and evaluate the impact of the decisions made.

The authoring solution (T-Parner, in a proof-of-concept version) implements the process, allowing instructors to use data analysis techniques (without the requirement of

---

[1]MOOCs: Massive Open On-line Courses.

[2]Check: `http://www.katyjordan.com/MOOCproject.html`

learning them), in order to detect and address pedagogical situations occurring within a course, such as: (1) detecting performance problems within a group, (2) comparing groups, (3) predicting dropouts and (4) predicting inadequate performance. Our proposal uses artificial intelligence to complement/augment human intelligence.

In this work, our objective was to evaluate if the interfaces allowed instructors to follow the PDMP to make the appropriate pedagogical decisions to address a particular problem (detecting performance problems within a group), i.e., evaluate if the interfaces were **useful** and **easy to use**. For that, we conducted an experiment where we invited teachers and tutors to interact with the T-Partner's interface and evaluate its usability, grading each interface according to their perception on its **Utility** and **Ease of Use**. The results show that the T-Partner guided instructors' decision-making, who positively perceived its usefulness and ease of use.

This article is organized as follows: in Section 2, we present some related works. Our proposal is described in Section 3. In Section 4, we describe the experiment design and in Section 5, we present the results. Finally, in Section 6, we draw our conclusion.

2. Related Work

In this section, we briefly discuss works related to our proposal, focusing on the authoring solution they present.

ProTracer 2.0 is an educational system that promotes programming teaching by tracking students' actions to identify problems and provide specific help to learners. The system is also able to help teachers evaluate the progress of their students, as a Virtual Teacher Assistant (VTA), sharing the teacher mentoring tasks in helping students practice programming. The idea of the system is that responsibility and confidence in the performance of the tasks is not solely with human teachers, nor does it depend entirely on ITS [Chou et al. 2011]. However, the system is used in a very restricted domain and the teachers' function is to aggregate information to the system to the point where they are no longer necessary.

ASSISTments is a platform that allows teachers to create artifacts called *ASSISTments* (composed of questions with answers and associated suggestions, solutions, web videos, etc.) or to (re)use pre-built *ASSISTments*, group them into a set of problems and assign them to their students. The system gives immediate feedback to students while they are studying and provides student-level data to teachers on any task. The platform allows any user, primarily researchers, to create controlled and randomized experiments on the content [Heffernan and Heffernan 2014]. However, the instructor can only create questions/tips, monitor the students' answers to the questions and, if necessary, create new questions/tips or improve the environment, limiting pedagogical decision-making.

3. Proposal

In this section, we present the main characteristics of the Pedagogical Decision-Making Process (PDMP) and the authoring solution that implements it (T-Partner).

3.1. Pedagogical Decision-Making Process (PDMP)

The PDMP (Figure 2) was created [Paiva et al. 2016] to instructors to: (1) discover pedagogical situations occurring in their courses; (2) understand these situations; (3) make

decisions to address them; (4) monitor and evaluate the impact of the decisions made.

The PDMP is a cyclic process and consists of two phases: the construction phase, where the collaborative interaction of human and computational intelligences takes place, to create the computational artifacts called *Pedagogical Decision Capsules*. These artifacts are created based on the instructors' decisions about (1) the pedagogical situation they wish to research in a learning environment; (2) how they want the related data to be processed for patterns/information to support decisions; (3) the decisions/recommendations/strategies they make to address such situation; And (4) how they wish to assess if the decisions were effective. In the next phase (construction), the definitions made by instructors (Pedagogic Decision Capsule), if valid, are executed automatically: (1) search for the defined situation, evaluating whether the patterns are the same as those identified in the Construction [3]; (2) recommend the resources and strategies chosen by the instructor; And (3) evaluate the effectiveness of the decisions [4].

Figure 2. The Pedagogical Decision-Making Process.

3.2. T-Partner

Following the process (PDMP), manually, is not an easy task [Paiva et al. 2016], being subject to (human) failures. We need a technology solution to support instructors, the T-Partner[5] authoring the *Pedagogic Decision Capsules*. To that, it will need to be integrated into the learning environment in order to guarantee access to data about: students, educational resources available, pedagogical approaches offered, other educational data resulting from interactions or other data that can be used to support pedagogical decision-making .

In summary, the T-Partner works as follows: students (1) interact with the learning environment. (2) Such interactions generate educational data that are stored. (3) These data are obtained and processed by T-Partner to (4) inform instructors about situations of educational interest. (5) Based on these information, instructors make pedagogical decisions that (6) will use educational resources and (7) the learning environment interface to (8) be sent to target students.

The T-Partner was designed in two versions, described below:

---

[3]Pedagogical situation + patterns = pedagogical scenario.

[4]Further information on the PDMP can be found in [Paiva et al. 2016]

[5]Teachers' Partner

1. *Light Weight* **Version**: a version with less features and less complex to use, but more limited. This version is intended for users with little experience with computers and for the quick creation of pedagogical decision capsules with simple content. It can also be used as an entry version for beginners.

2. *Heavy Weight* **Version**: a version with greater number of functionalities, but with higher complexity. However, this version allows greater detailing of the pedagogical decision capsules, which is appropriate if the user wishes to have more control or more detailed and specific needs for the use of the system.

3.3. Implementing the PDMP

In this subsection, we will describe how T-Partner implements the PDMP.

1. **Step 1: Define Pedagogical Situation**: in this step, the instructor chooses and personalizes the pedagogical action of interest. They then define which values characterize the results as inadequate, insufficient, and appropriate. The system searches the data requested by the instructor, based on the pedagogical action chosen, and classifies them according to the definitions.

2. **Step 2: Investigate Pedagogical Situation**: in this step, the instructor selects how he wants to pre-process, process and post-process educational data related to the chosen situation. The system uses the algorithm associated with the chosen action and generates the defined visualization, to support pedagogic decision making.

3. **Step 3: Define Pedagogic Decision**: in this step, the instructor defines the tasks and strategies that the students should carry out, considering the learning levels[6] for the chosen situation. The system recommends these decisions to students.

4. **Step 4: Define Assessment**: in this step, the instructor sets the desired percentage of adherence[7] and the expected outcome of those who followed them, for each learning level. The system checks if the criteria were met.

3.3.1. Components Diagram

In this subsection, we present the structural organization of the authoring solution. We adopted the components paradigm once it provides the general structure of the system and describes its behavior and its provided and required interfaces[Cheesman and Daniels 2000].

The T-Partner architecture utilizes kernel and satelite-centered development. According to Figure 3, we perceive the following layers: (1) Application, contains the interaction interface with the services; (2) Services, contains the services provided by the system; (3) System, contains the main components of the application, including the kernel; (4) Business, controls and mediates between the system layer and persistence; and (5) Persistence, contains the bundle of data.

4. Design of the Experiment

In this section, we present the experimental design conducted to evaluate the T-Partner.

---

[6]Inadequate, Insufficient or Adequate

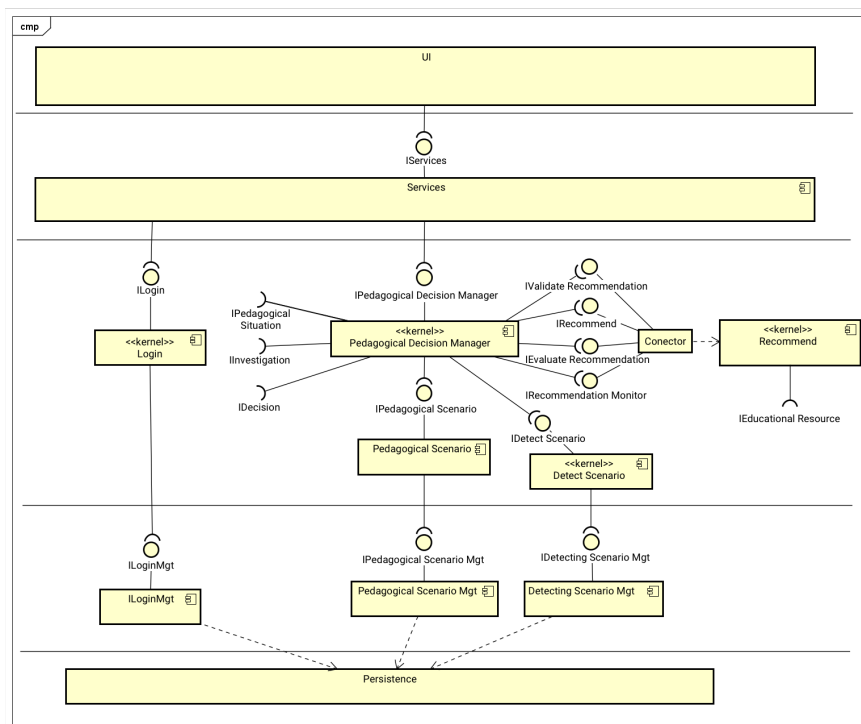[7]Students who followed his recommendation.

Figure 3. Components Diagram - Kernel Architecture.

It consisted of interfaces from the T-Partner (see Subsection: 3.2) and a control version that implemented the PDMP (see Subsection: 2), called spreadsheet (SS), was designed to mime the authoring capabilities of the T-Partner adapted to a known technology (spreadsheets). It was designed to be the simplest and easiest way to author a pedagogical decision capsule. The objective was to identify possible difficulties with the authoring solution (technology) or with the authoring steps (process).

Participants (teachers and tutors), upon accessing the experiment link, were automatically directed to one of the versions of the experiment. After interacting with each authoring step they should complete a questionnaire evaluating[8] each interface with respect to the following metrics: (1) Utility Perception; (2)Ease of Use Perception; (3) Attitude Towards Use; (4) Intention to Use; (5) Perceptions for the Traffic-Light Metaphor and (6) Perception of terms to categorize learning. In this work, we will cover metrics 1 and 2 that are related to the usability of the interfaces [Teo 2011]. The purpose was to evaluate the following hypotheses:

$H_{1.0}$: There is no difference between the authoring solutions regarding *Utility Perception*.

$H_{1.1}$: At least one of the solutions differs regarding *Utility Perception*.

$H_{2.0}$: There is no difference between the authoring solutions regarding *Ease of Use Perception*.

$H_{2.1}$: At least one of the solutions differs regarding *Ease of Use Perception*.

---

[8]Following a Likert scale from 0 to 6, where 0 = I completely disagree; 1 = I disagree; 2 = I somewhat disagree; 3 = I do not agree nor disagree; I agree; 5 = I agree; 6 = I completely agree.

Our goal is for the T-Partner to achieve at least the same perception as the control solution, which was designed to be simple, easy and uses a known technology.

5. Results and Discussion

In this section, we report the results our experiment.

5.1. Participants and Data

From 12/12/2016 to 04/01/2017, we obtained 175 valid interactions with T-Partner. The heavy weight (HW), light weight (LW) and spreadsheet (SS) versions participation count were, respectively, 63, 53 and 59 participants.

The data analyzed came from independent samples, since each participant only used one of the versions of the experiment. However, for each version there are 4 steps, whose interactions complement each other to form the decision capsule. Thus, we realize that the data of the different versions are independent among each other, but dependent on the steps.

Initially, we checked the normality of the sample data in order to choose the most appropriate test given the nature of the data (Shapiro-Wilk test). Then, for data from a non-normal population, we used the Wilcoxon Mann Whitney test to compare two independent samples and the Kruskal-Wallis test for the analysis of samples with respect to 1 factor. For normal data, we used the One-way ANOVA test for 1-factor analyzes, both with a confidence level of 95%, and the Bonferroni adjustment [9].

5.2. Analysis

Figure 4 displays the notes (scale *Likert*) assigned by participants to *Utility Perception*. We observed that the medians are close to note 4. The exceptions are steps 3 and 4 of the SS version, which presented medians less than 4; And stage 4 of the LW version, which presented median near 5.
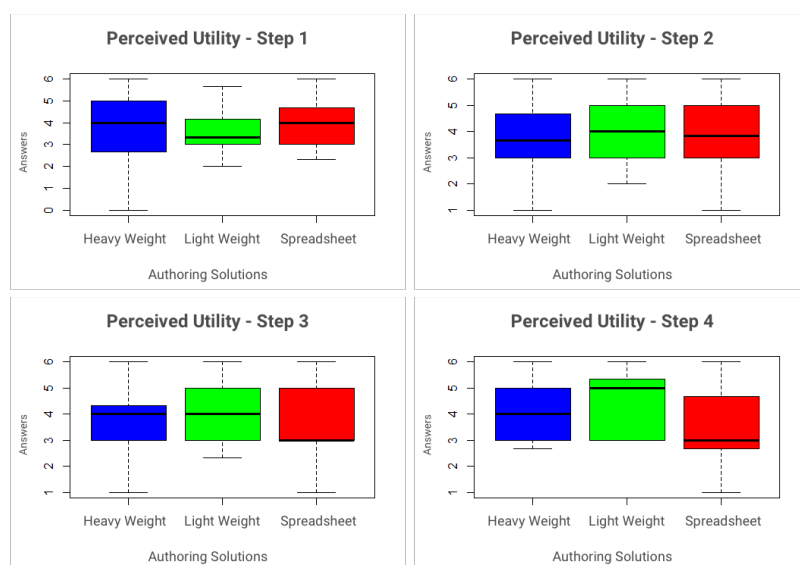
Performing the normality test for the metric *Utility Perception*, we note that only Step 4 of the SS version obtained a significant p-value when compared to the LW version, even after applying the Bonferroni adjustment.

We conclude that *Utility Perception* was statistically similar for the three versions of the experiment, with a slight difference in step 4. Thus, it is not possible to reject the null hypothesis. The implication of this result is that the perception of utility was favorable and similar between the versions, showing that the T-Partner was as useful as the control solution.

Figure 5 displays the notes assigned by the participants to *Ease of Use Perception*. We observed that the medians are close to 4. The exceptions are steps 3 and 4 of the LW version, which obtained medians close to 5.

---

[9]The Bonferroni correction is an adjustment made to the p- Values when several dependent or independent statistical tests are being performed simultaneously on a single data set. To perform a Bonferroni correction, divide the critical p-value by the number of comparisons made.

Figure 4. Utility - All steps of the authoring phase for each version.
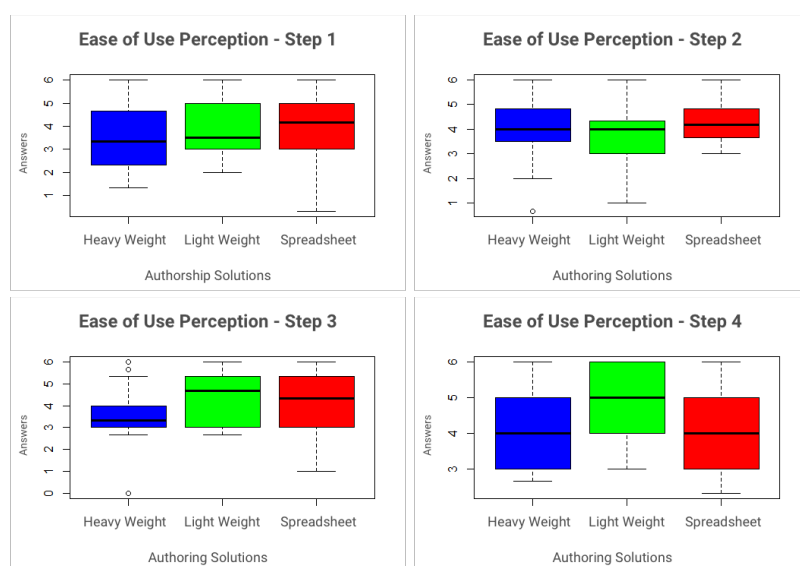
Figure 5. Ease of Use - All steps of the authoring phase for each version.

Performing the normality test for the metric, we noticed that the data belongs to a non-normal distribution and the Kruskal-Wallis test, for factorial analysis, confirms that the differences are not significant.

Table 2. Ease of Use Perception - Statistical Tests

Test	Step 1	Step 2	Step 3	Step 4
Shapiro-Wilk	0.03367	0.02267	0.001013	1.356e-05
Kruskal-Wallis	0.2741	0.2754	0.3305	0.1636

Table 1. Utility Perception - Statistical Tests

Test	Step 1	Step 2	Step 3	Step 4
Shapiro-Wilk	0.03417	0.02051	0.0002181	0.001545
Kruskal-Wallis	0.5943	0.6035	0.6383	0.02322*
Step 4	*SS*	*LW*	*HW*	*P-value*
Wilcoxon-Mann-Whitney Bonferroni Adjustment	x	x		0.01244 0.03731718*
Wilcoxon-Mann-Whitney Bonferroni Adjustment	x		x	0.02627 0.07879614
Wilcoxon-Mann-Whitney		x	x	0.8879

We conclude that the *Ease of Use Perception* was statistically similar for the three versions of the experiment. Therefore, it is not possible to reject the null hypothesis. The implication of this result is that the perception of ease of use was favorable and similar between versions, showing that the T-Partner was as easy to use as the control solution.

6. Conclusion

We reported the problem on-line learning instructors face when trying to handle educational data, making it impossible for them to manually obtain information to help them manage their courses. The consequences are failure and/or dropouts. To address it, we proposed a process (Pedagogical Decision-Making Process (PDMP)), and an authoring solution (T-Partner) to help them follow the PDMP and make informed pedagogical decisions. In order to evaluate the proposal, we evaluated teachers'/tutors' perceptions regarding the usefulness and ease of use of T-Partner interfaces, comparing it with a control version (a version adapted to a spreadsheet).

The results showed positive perceptions for both metrics, with no differences between the tested versions (with subtle exceptions). This demonstrates the authoring solution was perceived as useful and easy to use, supporting the work of teachers/tutors in on-line learning environments. However, there is a need to improve the interfaces, in order to make the T-Partner interfaces better than the (fictitious and ultra-simplified) control solution.

References

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

Belanger, Y. and Thornton, J. (2013). Bioelectricity: A quantitative approach duke university's first mooc.

Bienkowski, M., Feng, M., and Means, B. (2012). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. *US Department of Education, Office of Educational Technology*, pages 1–57.

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.

Cheesman, J. and Daniels, J. (2000). *UML components: a simple process for specifying component-based software*. Addison-Wesley Longman Publishing Co., Inc.

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.

Chrysafiadi, K. and Virvou, M. (2013). Student modeling approaches: A literature review for the last decade. *Expert Systems with Applications*, 40(11):4715–4729.

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

Heffernan, N. T. and Heffernan, C. L. (2014). The assistments ecosystem: building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education*, 24(4):470–497.

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.

Liyanagunawardena, T. R., Parslow, P., and Williams, S. (2014). Dropout: Mooc participants' perspective.

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.

Paiva, R., Bittencourt, I. I., Tenório, T., Jaques, P., and Isotani, S. (2016). What do students do on-line? modeling students' interactions to improve their learning experience. *Computers in Human Behavior*, 64:769 – 781.

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.