

At the Intersection of Technology and Teaching: The Critical Role of Educators in Implementing Technology Solutions

Rachel Van Campenhout, Acrobatiq by VitalSource, United States
Murray Kimball, Acrobatiq by VitalSource, United States

The IAFOR International Conference on Education – Hawaii 2021
Official Conference Proceedings

Abstract

Educators are critical for the successful implementation of any technology. Acrobatiq by VitalSource can use data to demonstrate the dramatic impact instructors—and their course policies—can have on courseware engagement. Acrobatiq courseware incorporates learning content, formative practice, homework assignments, adaptive practice, and summative assessments into a single learning environment for students, with additional data dashboards for instructors. Previous research has shown that the “learn by doing” approach, central to the courseware, has a six-times effect size on learning than reading alone, so engaging with the formative practice is critical to student success. A statewide system of colleges and universities used Acrobatiq’s Probability and Statistics courseware in a grant-funded initiative. The instructors were all provided extensive training on the courseware features, instructor dashboards, and pacing suggestions before the term began, however, each instructor was able to dictate how they incorporated the courseware into their teaching practice and course grades. We analyzed the courseware data using a visualization called engagement graphs and found a surprising level of variability between instructors. These findings demonstrate the impact that instructors and their policies have on the successful implementation of the courseware. Because engagement is a vital component for the learning benefit of the courseware learning environment, research is needed to better identify implementation practices which affect student engagement. At this intersection of learning science-based technology and teaching practice is immense potential to increase student success.

Keywords: Courseware, Student Engagement, Fidelity of Implementation, Teaching Practice

iafor

The International Academic Forum
www.iafor.org

Introduction

When technology is developed on a foundation of learning science and is rigorously researched to iteratively improve and optimize its performance, students benefit from increasingly effective learning environments. The Acrobatiq platform utilizes a proven “learn by doing” method to help students master content efficiently (Lovett et al., 2008). The courseware integrates frequent formative practice with the explanatory text and media, allowing students to practice at the point of learning. This formative practice provides immediate targeted feedback and gives students a low-stakes environment to check their learning. This learning by doing method produces the doer effect: engaging in practice has six times the effect size on learning than reading alone (Koedinger et al., 2015). The doer effect has also been shown to be causal in multiple research studies, including Acrobatiq courseware used at scale, allowing us to recommend this method with confidence (Koedinger et al., 2016; Koedinger et al., 2018; Olsen and Johnson., 2019).

The courseware provides an effective method for learning and practicing new content as well as delivering adaptive activities and graded summative assessments. In previous research done on the course analyzed in this paper, we found that the adaptive activities were beneficial for students, especially low and intermediate performing students (Van Campenhout et al., 2020). After students complete a module of lessons (which includes content and formative practice tied to learning objectives), students completed an adaptive activity before the summative module quiz. The adaptive activity personalized a set of questions based on the students’ needs; their performance on the formative practice informed what level of scaffolding to provide to each student for each learning objective they encountered. Results showed that a significant portion of students who completed the adaptive activities were able to increase their learning estimate (a learning measure generated by Acrobatiq’s predictive model). Students who increased their learning estimates through the adaptive activities scored higher on the summative assessment than their peers who did not (Van Campenhout et al., 2020).

While the benefits of a research-based learning environment are clear, the classroom instructional model has been shown to have a large effect on student learning. Instructional content can be delivered in class or outside of class, through the instructor or through technology. In a meta-analysis of studies on the effectiveness of mixed methods course design, the flipped blended model was the only type that outperformed other models of delivery (Margulieux et al., 2015). The flipped blended model delivers content via technology and provides feedback via the instructor. When courseware is used as the instructional material outside of class in a flipped-blended model, students have the added benefit of receiving feedback from the technology for formative practice as they learn the material, which enhances their mastery of content before working through activities with the instructor in class.

Given the optimization of both the courseware as a technology-based learning environment and the instructional model best fit to utilize this learning resource as the out-of-class instruction, what additional variables could impact the effectiveness of this method? Individual instructor variation in implementation can greatly impact the outcomes expected from a technology or instructional model. As Kessler et al. (2019) noted, “research consistently indicates that instructional innovations are only as

effective as their implementation.” The role of the instructor in computer-directed learning environments is often minimized, as these environments are required to be designed for a wide audience and various complex and divergent learning situations (Kessler et al., 2019). The Acrobatiq courseware was designed to fit a variety of learning models, with research showing effective outcomes in student self-directed asynchronous models as well as faculty-led flipped blended models (Olsen and Johnson, 2019; Van Campenhout et al., 2020). While the student interface is designed as a complete environment for them, the instructor dashboards are a significant feature of the platform. For contexts where instructors are involved in the courseware delivery, the dashboards organize data around actionable questions to facilitate instructor involvement in the interactions between students and their course. The delivery of actionable data to instructors for use at their discretion is a type of Course Signal, which has been shown to help increase course and university retention (Arnold and Pistilli, 2012; Baker, 2016). As other researchers have recently proposed, the proper utilization of both the educational environment and intelligent tutoring systems should produce a better learning experience than either could produce on their own (Ritter et al., 2016).

The importance of implementation is not a new concept; Fullan and Pomfret (1977) reviewed research literature on implementation to define the construct, address its importance, and identify how researchers measured it. O’Donnell (2008) completed a review of the literature to define and measure the relationship of implementation and outcomes in intervention research. O’Donnell (2008) defines fidelity of implementation as a “determination of how well an intervention is implemented in comparison to the original program design during an efficacy and/or effectiveness study.” There are several key ideas in this definition to unpack for their relevance to this paper. First is the concept that the intervention implementation should be compared to the original program design. Courseware is designed using specific learning science principles to elicit specific benefits for students. While there are many different mixed methods teaching models being used, the implementation of courseware into a model should also be compared to the design intentions and the literature to understand what the expected outcomes might be. Meaning, if the efficacy results were measured using a flipped blended teaching model, but an implementation uses a lecture hybrid model, it should not be expected to find the same results as the original design. Second, fidelity of implementation is critical when doing an efficacy study, but not all uses will have this as a goal. With a variety of educational settings for courseware, it is reasonable that not all will be designed to optimize effectiveness for various reasons. However, for uses in which efficacy is a goal or measurement of success, fidelity of implementation is critical. Finally, fidelity of implementation requires a determination of how well an intervention is implemented, which indicates the need to evaluate based on criteria. While a review of public health literature identified five criteria for measuring fidelity of implementation (adherence, duration, quality of delivery, participant responsiveness, and program differentiation), it is also clear that establishing criteria for fidelity of implementation requires a close evaluation of the treatment and its acceptable uses (O’Donnell, 2008).

Implementation

Through a research grant, a state-wide system of universities and community colleges were able to use the same courseware across all introductory probability and statistics courses. There were 8 individual institutions and 20 course sections in the fall pilot. Instructors were required to attend two trainings to onboard them with the Acrobatiq courseware. The first session included an overview of the course and basic navigation of the platform that was held prior to the semester start. After 5 weeks, a subsequent training focusing on utilizing the data in the Learning Dashboard was delivered that focused on how instructors could identify engagement risks and learning objectives that students were struggling to master in their own courses.

Best practices were established for course setup and grading to help increase student engagement. Instructors were encouraged to set due dates on all quizzes and assignments to clearly establish these elements as required course components for their students. It was recommended that instructors give a participation score (5% or greater) to students for completing all the formative practice in the course. Instructors were encouraged to use the courseware in a flipped blended teaching model, so students could complete the foundational work via the courseware and instructors could evaluate their progress via the dashboards before class. Instructors still had the ultimate control over their teaching model and how they implemented the courseware as a part of their syllabus and gradebook.

Data

The Engagement Graph

After the semester had concluded for all institutions, the Acrobatiq Research and Development team used a data visualization called an engagement graph to compare aggregated institutions as well as individual instructor sections. The engagement graph was developed as a way to visualize how students were engaging with the courseware over time. The pages of a course are ordered along the x-axis, and the number of students along the y-axis. This creates a view of a class over time in the courseware. Dots are added to each page to show the number of students who read content on a given page, the number of students who engaged with the formative practice on that page, and the number of students who completed adaptive or summative assessments.

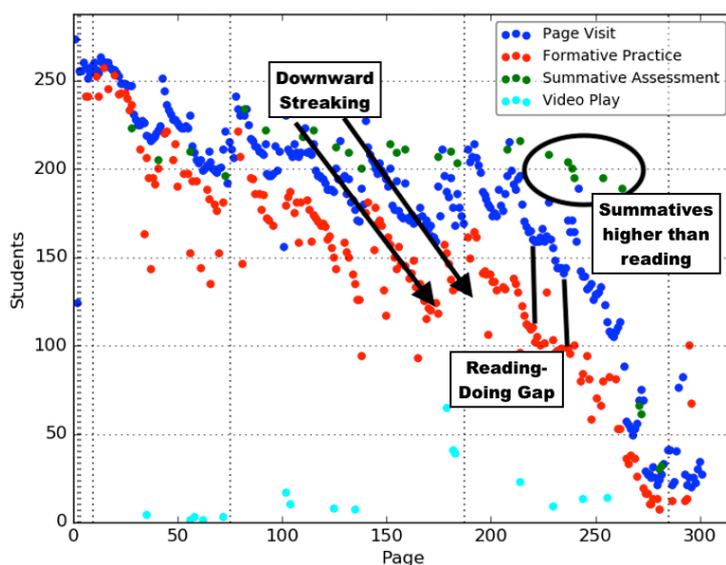


Figure 1. An engagement graph marked to show examples of within-unit streaking, the reading-doing gap, and higher summative engagement than reading or doing engagement.

This engagement graph example shows a relatively typical course. As we move along the x-axis from the beginning of the course to the end, there is a steady decrease in engagement, with a steeper drop-off toward the end. This tells us what is generally known—that some students stop doing their work toward the end of the semester. We also see downward streaking from left to right in a downward repetitive pattern. This notes a pattern that within modules, some students drop out partway through the module only to return at the start of the next. The blue dots indicate the number of students who read each page, while the red dots indicate the number of students who did the formative practice questions. The red dots are below the blue, meaning some students read the page but do not do practice. We call this the reading-doing gap. As seen in this graph, that gap between reading and doing widens over time, meaning fewer students engage in the practice as the course nears the end.

In an ideal world, all students would read every page and do all the practice opportunities, so the engagement graph would be a horizontal line at the number of students in the course. It is unrealistic to set this as the goal, but it is reasonable to aim to reduce the reading-doing gap and increase engagement across the course.

Engagement Graphs by Institution

The first level of inspection took place at the institution level. It was expected that we might see differences between institutions due to variables such as differing student characteristics between institutions of different types. The engagement graphs which had combined data for all sections at the institution confirmed there were drastic differences in how students engaged with the courseware between institutions. Figure 2 shows three institutions as a side-by-side comparison.

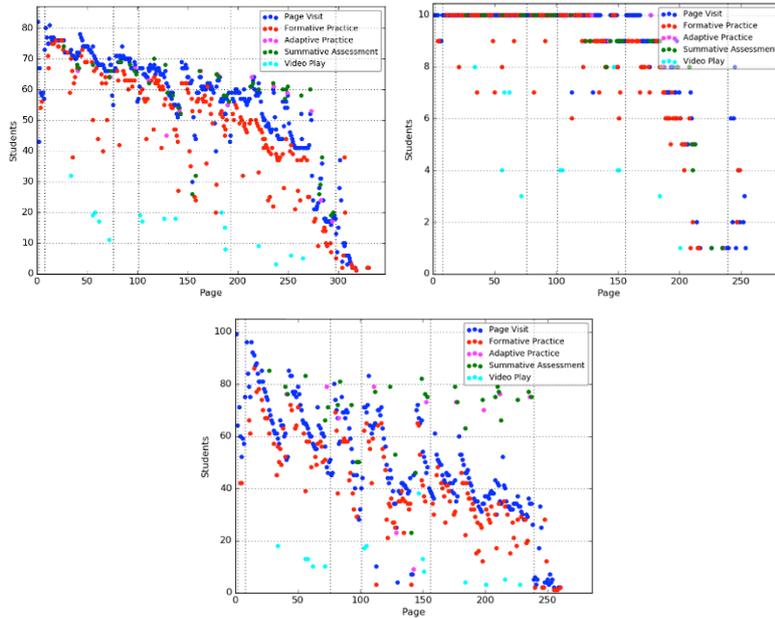
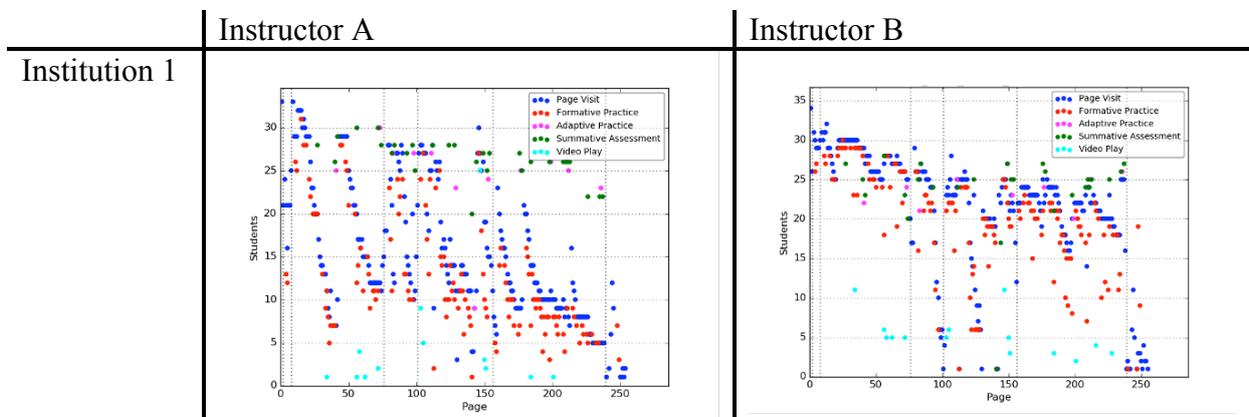


Figure 2. A series of three aggregated institutional engagement graphs which show strongly divergent patterns of engagement.

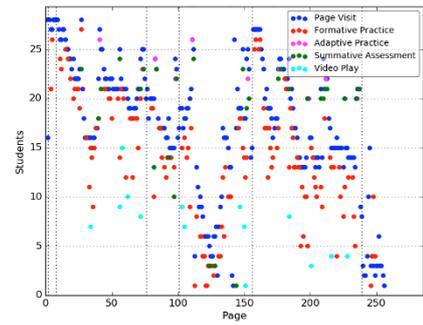
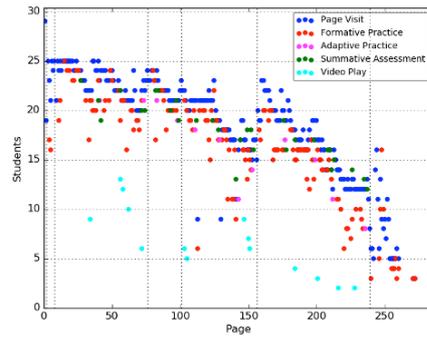
Each engagement graph looks drastically different at a glance. The number of total students varies from 10 to 100. The engagement graph on the left shows a slow decline in use over time but a fairly steady decrease with minimal vertical streaking. The engagement graph in the middle shows dramatic vertical streaking and poor engagement through the majority of the course. The engagement graph on the right has a nearly horizontal line of engagement for reading and doing, which is close to ideal usage. If these were the only data views available we might conclude that the influencing factor could be institutional policies or differing student characteristics.

Engagement Graphs by Instructor

Most institutions had multiple sections of the courseware being used by different instructors. When we look at a selection of engagement graphs separated by instructor, as in Table 1, we see unique differences.



Institution 2



Institution 3

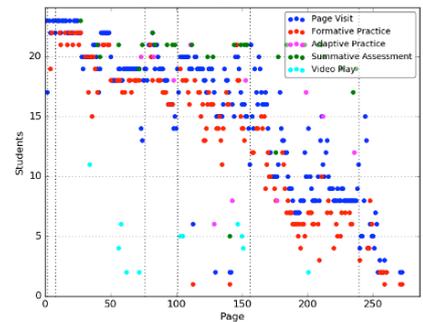
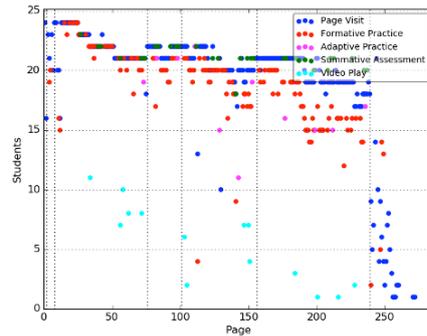


Table 1. A comparison of engagement graphs between two instructors at the same institution.

Visual inspection of these different engagement graphs shows very divergent student engagement patterns with the courseware between instructors. Instructor 1A's section shows extreme vertical streaking and low overall usage. Students in this section took the summative assessments, but most students quickly stopped looking at pages or doing practice within modules, with only about a third of all students working through the courseware consistently. Comparatively, instructor 1B's section shows that the majority of students consistently used the courseware, with vertical streaking limited to a range of roughly five students.

At the second institution, instructor 2A's section shows a fairly tight band of reading-doing, with only a few students fluctuating vertically, and almost all students doing the practice as well as reading. Instructor 2B's section shows a large variation of engagement, with nearly half the class reading or not reading, doing or not doing. While instructor 2A's section shows a slight dip in the middle of the course and a steady decline in usage in the last unit, instructor 2B's section shows a dramatic dip in the middle of the course and wide fluctuations in usage throughout.

At the third institution, instructor 3A's course shows a fairly horizontal line with variation in reading and doing of only two or three students, with a slightly larger decrease in doing at the end of the course. Comparatively, instructor 3B's section shows a consistent vertical variation in reading and doing of five or six students, with a dramatic decrease at the end of the course.

Engagement and Final Exam Scores

While each institution created their own final exam for the course, there was a portion of questions that were the same across all institutions. A comparison of the engagement graph patterns for the institution as a whole with the mean score of the common questions for students at each institution shows a relationship between the overall level of engagement and the mean assessment score for common questions. The courses at the top with more student retention to the end of the course had higher scores while the engagement graphs at the bottom with low engagement had lower scores.

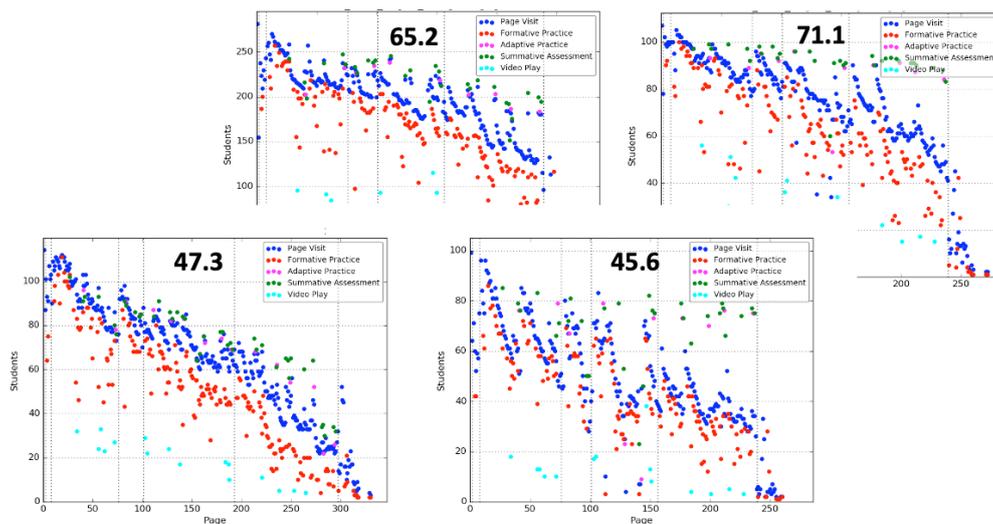


Figure 3. Four institutional engagement graphs of varying patterns with the mean score of the common final exam questions.

Results and Discussion

Inspection of these data visualizations revealed valuable insight into the variability of engagement within individual course sections. While not a randomized experiment with control over all variables, this initiative provided many controls across a large number of courses run in a natural setting. With a single initiative organizing the mission of the project, the same courseware being used, and the same training and instruction provided to instructors, we had some expectation of similarity of usage and outcomes. Our initial assumptions were to see variation according to differing student populations between institutions. Instead, we see significant variation in engagement patterns between sections at the same institution. While it's possible that additional variables could contribute to differing engagement (different course times, different student groups, etc.), it is unlikely that those could account for the entirety of such drastic differences. The instructor and their choices regarding implementation greatly impacted how students chose to engage with the courseware.

Instructors who had the highest student engagement shared several important commonalities: they used due dates for assessments, included completion of formative practice as a part of the student's participation grade, attended all trainings and attempted a flipped classroom model to some degree. Though training was a required element of this pilot, the several instructors who did not attend all trainings

were also the instructors that had some of the lowest student engagement in their course sections. These findings informed how training, instructor resources, and best practices were created and used in subsequent pilots. Using a flipped classroom model was recommended for instructors but this was implemented at varying degrees. Better defining what a flipped classroom model was and providing additional tools for instructors to better leverage this teaching modality was one of the lessons learned from this pilot.

Conclusion

This data validates our belief that the instructor is critical to the success of technology in the classroom. While the courseware itself is proven to be effective in helping students learn, it can only do so if students engage with it. Instructors hold enormous sway over how students engage with the courseware and therefore benefit from the technology.

This data analysis suggests that the usage of courseware should also be paired with a framework to evaluate the fidelity of implementation. The validity of efficacy research is diminished if the results cannot be clearly attributed to the courseware or the implementation of the courseware. Further work should be done to establish a theoretical framework and criteria for implementation as well as an evaluation of the level of fidelity to that implementation.

These findings indicate several avenues for future research. First, it is clear that more work needs to be done to investigate how a fidelity of implementation framework could be leveraged in real-world contexts to increase the validity of effectiveness research. Second, given that engagement with the courseware is the only way to benefit from the proven learning science principles inherent to its design, increasing engagement must be the focus of future research. We would be interested in evaluating how instructor policies such as participation scores, late work policies, and gradebook settings are related to student engagement. Additional research should also be done into how qualitative factors such as approaches to introducing students to courseware, expectation setting, and instructor attitudes can influence student engagement.

Acknowledgements

We would like to thank Benny Johnson and Bill Jerome for their thoughtful review of this manuscript. We also would like to thank the dedicated educators working to better learning experiences for students.

References

- Arnold, K. E., Pistilli, M. D. (2012) Course Signals at Purdue: Using learning analytics to increase student success. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*. New York, NY: ACM, pp. 267–270.
- Baker, R. S. (2016). Stupid Tutoring Systems, Intelligent Humans. *International Journal of Artificial Intelligence in Education*, 26(2), 600–614. <https://doi.org/10.1007/s40593-016-0105-0>
- Fullan, M., & Pomfret, A. (1977). Research on curriculum instruction implementation. *Review of Educational Research*, 47, 335–397.
- Kessler, A., Boston, M., & Stein, M. K. (2019). Exploring how teachers support students' mathematical learning in computer-directed learning environments. *Information and Learning Science*, 121(1–2), 52–78. <https://doi.org/10.1108/ILS-07-2019-0075>
- Koedinger, K., Kim, J., Jia, J., McLaughlin, E., & Bier, N. (2015). Learning is not a spectator sport: doing is better than watching for learning from a MOOC. In: *Learning at Scale*, pp. 111–120. Vancouver, Canada. <http://dx.doi.org/10.1145/2724660.2724681>
- Koedinger, K., McLaughlin, E., Jia, J., & Bier, N. (2016). Is the doer effect a causal relationship? How can we tell and why it's important. *Learning Analytics and Knowledge*. Edinburgh, United Kingdom. <http://dx.doi.org/10.1145/2883851.2883957>
- Koedinger, K. R., Scheines, R., & Schaldenbrand, P. (2018). Is the doer effect robust across multiple data sets? *Proceedings of the 11th International Conference on Educational Data Mining, EDM 2018*, 369–375.
- Margulieux, L. E., McCracken, W. M., & Catrambone, R. (2015). Mixing in-class and online learning: Content meta-analysis of outcomes for hybrid, blended, and flipped courses. In O. Lindwall, P. Hakkinen, T. Koschmann, P. Tchounikine, & S. Ludvigsen (Eds.) *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference* (pp. 220-227), 2. Gothenburg, Sweden: The International Society of the Learning Sciences.
- Lovett, M., Meyer, O., & Thille, C. (2008). The Open Learning Initiative: Measuring the effectiveness of the OLI statistics course in accelerating student learning. *Journal of Interactive Media in Education*, (1), 1-16. <http://doi.org/10.5334/2008-14>
- O'Donnell, C. L. (2008). Defining, Conceptualizing, and Measuring Fidelity of Implementation and Its Relationship to Outcomes in K-12 Curriculum Intervention. (2008). *Review of Educational Research*. 78(1). Pp. 33–84. <https://doi.org/10.3102/0034654307313793>

Olsen, J., & Johnson, B.G. (2019). Deeper collaborations: a finding that may have gone unnoticed. *Paper Presented at the IMS Global Learning Impact Leadership Institute*. San Diego, CA

Ritter, S., Fancsali, S., Yudelson, M., Rus, V. and Berman, S. (2016). Toward intelligent instructional handoffs between humans and machines. *Workshop on Machine Learning for Education, The Thirtieth Conference on Neural Information Processing Systems (NIPS)*, Barcelona.

Van Campenhout, R., Jerome, B., & Johnson, B. G. (2020). The impact of adaptive activities in Acrobatiq courseware: Investigating the efficacy of formative adaptive activities on learning estimates and summative assessment scores. In: Sottolare R., Schwarz J. (eds) *Adaptive Instructional Systems*. HCII 2020. LNCS, vol 12214. Springer. pp 543–554. https://doi.org/10.1007/978-3-030-50788-6_40

Contact email: rachel@acrobatiq.com