Chapter 16 – Reports To Facilitate Improvements of Adaptive Instructional Systems

Andrew J. Hampton[1] and Ryan S. Baker[2]
[1]University of Memphis, [2]University of Pennsylvania

Introduction

Many technologically groundbreaking methods have been proposed for creating self-improving adaptive instructional systems (AISs). Where these proposals primarily embrace innovative uses of advanced technical capabilities, we emphasize maintaining contact with a base from which systems should improve themselves: consideration of how humans perform similar tasks. In many of these systems, developers have processes in place to receive meaningful and interpretable data on various aspects of learner performance in order to understand system quality and learner outcomes. Providing accurate, relevant system information in a comprehensible, usable fashion constitutes a necessary precondition to any reliably and demonstrably self-improving (or in fact, manually improving) AISs. In this chapter, we consider how instructional systems are currently adapted, and propose some ideas for how analogous improvement approaches can be implemented within self-improving systems.

The concept of a self-improving system presupposes that (a) quality is measurable and (b) measurements of quality can inform changes meant to improve those measurements. AIS practitioners continue to explore creative ways to produce autonomously controlled enhancements based on these two pillars, but we must carefully verify that each of these two preconditions is met. That is, a self-improving system may generate improvements by identifying and adjusting sub-optimal interactions with learners, but without an effective way to measure learner outcomes, the changes may be arbitrary or even counterproductive. This could happen for several reasons. Changes may be ineffective because of poor measurement. Changes may be effective but the system is unable to convince an outside practitioner that this is the case. Systems may become trapped in local maxima, unable to reason out of a condition in which parameters of interest (but not actual performance) appear optimally tuned. The system may recognize a problem but not have the operators to address it. In other words, a self-improving system may not always be able to improve itself. Sometimes the only way to solve a problem may be to present relevant, digestible information to humans and enlist them to fix the flaw.

While simple in principle, selection and presentation of information provide an array of challenges. An AIS, by definition, has complex interaction protocols. Learner input changes the system output, using complex formulations of performance to produce truly unique learner profiles. This expansive array of possible paths impedes a priori categorization of learners into discrete data categories. Further, systems typically contain metrics for performance on multiple facets of a topic rather than a holistic score. Layers of multiplicative complexity stem from unstandardized inclusions of psychological variables, coordination with external instruction, or consideration of learners collectively (e.g., how a class performs, and a single learner in relation to that class). Identifying the critical variables requires careful consideration and often sophisticated data analysis. Representing the resulting data may vary drastically depending on their type (individual versus collective, qualitative versus quantitative, etc.) and rate of change. Exceptionally complicated concepts and relationships need to be conveyed efficiently, with primary metrics easily perceptible and the option to dig deeper readily available.

Dashboards provide the primary method of conveying these types of information to practitioners. Understanding the ways in which humans use the dashboard data fundamentally impacts dashboard design. Learning analytics researchers have established methods for developing dashboards for use by instructors and system designers. A careful review of these research traditions with an eye toward cultivating, organizing, and presenting AIS variables to practitioners will yield actionable recommendations for designers.

Designing for Users

"Know thy user", a common mantra among system designers, reminds us that there exists a typical set of use cases, and that those use cases should be made as easy and efficient as possible. In the context of an AIS, the most critical variables typically consist of metrics related to learner performance and quality of instructional content. These metrics then form the basis of changes made to the system or curricula to improve learner outcomes. "Know thy user" thus suggests that designers present information in such a way as to match the existing mental models of instructors and practitioners. With this match established, common evaluations or procedures proceed as seamlessly as possible.

However, it can in some cases be challenging to develop reporting and dashboards that support teachers in conducting data-driven decision-making (Marsh & Farrell, 2015). As a result, many dashboards are used unevenly across teachers (Hawn, 2019). This relates broadly to the challenge of providing professional development training to teachers that encourages individualized instruction (Al Otaiba, Connor, Folsom, Greulich, Meadows, & Li, 2011). Principled dashboard design is unlikely to fix this problem on its own, but certainly constitutes a step in the direction of broader adoption, both of dashboards and of the effective instructional practices that they can promote.

The conventional human–computer interaction approach of matching the presentation of information to existing mental models assumes appropriate mental models for users (Flach & Voorhorst, 2016) who may have limited experience in pedagogical theory or educational technology. The challenge then becomes to present information in such a way as to *encourage* a mental model in line with the affordances and constraints of the system, with respect to the ultimate goal of improving learning outcomes. In this way, information can directly suggest action by encouraging an intuitive understanding of the relationship between intervention and outcome, between intentional cause and desired effect.

For example, data representations of individual instructional items should immediately suggest to instructors which items fail to challenge learners (ceiling effect) or serve only to frustrate and confuse (floor effect). Areas ripe for improvement should be accentuated, and methods of intervention intuitively suggested. Dashboard design provides a direct opportunity to influence mental models toward appropriate action.

Applications of Dashboards in Improvement of Instruction

Learning analytics dashboards provide potentially useful resources for many stakeholders attempting to improve the performance of AISs. This includes developers working *on* the systems, teachers working *in* the systems, and school personnel working *with* the systems as part of their broader instructional approach. Many AISs have been designed for and used by school and university personnel, including teachers, instructors, academic advisors, school counselors, and higher-level administrators. Though most dashboards are presented through traditional computer screen displays, recent research has considered if other modalities may be better for providing reports to classroom teachers, including wall-based ambient displays (Alphen & Bakker, 2016) or augmented reality headsets (Holstein, McLaren, & Aleven, 2018).

Perhaps the most common application for dashboards in the learning analytics space is in communicating data on learner student performance to instructors. These types of dashboards have become commonplace within K–12 adaptive learning systems and were present from the first large-scale usage of AISs. However, their inclusion was initially poorly documented. For example, a "teacher's toolkit" with reports on student performance has been part of the *Cognitive Tutor* system and used in schools since the mid-1990s but is not discussed in detail in publications regarding that system. Teacher reports were considerably better documented in the later *ASSISTments* system, with an article discussing the range of reports offered by the system as well as their use by teachers and school leaders (Feng & Heffernan, 2006). Both *Cognitive Tutor* and *ASSISTments* are relatively straightforward mathematics problem-solving environments. This type of dashboard has been incorporated in recent years by other types of learning systems such as exploratory learning environments (Mavrikis, Gutierrez-Santos, & Poulovassilis, 2016) and group learning environments (Martinez-Maldonado, Clayphan, Yacef, & Kay, 2015).

Another popular use for learning analytics dashboards is in risk prediction—typically dropout or course failure. Perhaps the most widely used platform for predicting which students are at-risk in higher education is the *Civitas* platform, used by universities worldwide and providing a range of reports about students, leveraging data on course-taking, grades, admissions data, and questionnaires administered to students (Milliron, Malcolm, & Kil, 2014). The *Course Signals* platform, perhaps the first widely used platform in this space, integrates dashboard reports with recommendations and scaffolding for email-based interventions by instructors (Arnold & Pistilli, 2012); when an instructor clicks on an at-risk student within the dashboard, the platform goes straight to a recommended action and scaffolded intervention tailored to that specific learner, which the instructor has the option to modify before sending to the learner. At the K–12 level, *BrightBytes*, used by dozens of school districts nationwide, offers reports on the at-risk status of students (Singh, 2018). These are only a few examples of the increasing number of platforms and vendors now offering at-risk prediction dashboards to universities and school districts. There have also been efforts to create dashboards for risk prediction within the context of MOOCS (e.g. Chen et al., 2016), though adoption lags the use of dashboards in for-credit online university programs. Many other uses exist for dashboards in learning, from presenting course recommendations to students (Denley, 2013), to providing visualizations to students of their own activity and progress (Kim, Jo, & Park, 2016), to visualizing group work (Kay et al., 2006).

However, there has not yet been sufficient attention in the published literature to dashboards for the enhancement of learning content. Existing dashboards used primarily for teachers to monitor student progress can effectively serve a secondary function of evaluating content (e.g. Feng & Heffernan, 2006) or the effectiveness of the learning platform. While the design and enhancement of dashboards is often seen as a secondary goal by the developers of adaptive learning systems, we would argue that dashboards are all but indispensable to the creation of high-quality computer learning environments. Though there are efforts to create dashboards for course designers, this work has often not been published, despite its pivotal role in enhancing the instructional quality of adaptive learning systems.

How Dashboards are Used

Dashboards are used in a variety of ways that potentially impact the practice of teachers (Miller et al., 2015; Xhakaj et al., 2017) and academic advisors (Arnold & Pistilli, 2012; Lonn, Aguilar, & Teasley, 2015). Teachers frequently access dashboards during class to drive pedagogical decisions, consulting dashboards as frequently as eight times per class session (Molenaar & Knoop-van Campen, 2017), and using this information to inform pedagogical strategies. For example, many teachers use proactive remediation, where teachers speak to a student struggling with specific material (Miller et al., 2015). Other teachers use dashboards to identify students who have recently succeeded and provide them with encouragement (Molenaar

& Knoop-van Campen, 2017). Molenaar and Knoop-van Campen (2017) note that teachers using dashboards shift from providing general learning support to the most struggling students to providing targeted support to a broader range of students.

Not all dashboards are used in real-time. Indeed, within the *ASSISTments* system, one of the most common uses of dashboards is by teachers to identify content that students struggled with during homework and redesign the next class session's lesson to target specific errors common across students (Kelly, Heffernan, D'Mello, Namias, & Strain, 2013). Teachers also use *ASSISTments* dashboards to split students into groups, putting students in groups together if they make similar patterns of errors, and giving advanced content to groups of students who made no errors (Heffernan & Heffernan, 2014).

The use of dashboards connected to adaptive (or non-adaptive) instructional systems is part of a broader move towards using data in instruction, what is referred to as data-based decision making or data-driven decision making (Halverson, 2014). Teachers now commonly receive a broad range of types of data within dashboards, including data from student information systems, formative assessment systems, disciplinary data, attendance data, and teacher observation data (Mandinach & Jackson, 2012). While there has been skepticism about the effectiveness of data-based decision-making involving coarser-grained data (Mandinach & Jackson, 2012), teachers' use of dashboards involving finer-grained data (such as is generated by AISs) is associated with better student outcomes, both in K–12 (Xhakaj et al., 2017) and in higher education (Arnold & Pistilli, 2012). It is not yet entirely clear which of these uses produces the benefits associated with dashboards—but this is exactly the sort of question that could be answered by a self-improving AIS. By trying the pedagogical strategies associated with the use of data in dashboards, one-by-one, an AIS could determine which approach to data-driven instruction works for which students in which contexts.

Recommendations and Future Research

These past projects suggest some steps for taking AISs forward. First, the types of decisions captured in instructor responses to dashboards may be directly applicable to training AISs to replicate instructor behaviors. Take, for example, the *Course Signals* system discussed above. This system presents recommended interventions to instructors, but instructors do not have to follow that recommendation—they can ignore it or modify the intervention. By tracking the situations in which instructors choose to intervene and how they choose to modify interventions, an AIS may be able to more closely replicate instructor behavior.

One could envision that this could be accomplished via linking between the Generalized Intelligent Framework for Tutoring (GIFT) architecture and an external system that provides recommendations to an instructor, via the Learning Management System Module. The recommendations could be generated from the Pedagogical Module, using information from the Learner Module. In turn, the Learning Management System could be designed to track whether and how instructors follow its recommendations, and this information could be passed back to the Pedagogical Module to update its algorithm for when and how to recommend intervention.

More directly, it may be possible to create a repertoire of AIS behaviors based on the behaviors in which teachers engage with dashboards, such as the proactive remediation behaviors identified by Miller and his colleagues (2015) and the encouragement behaviors identified by Molenaar and Knoop-van Campen (2017). Again, these behaviors could be embedded into the Pedagogical Module and triggered by the Learner Module. In this case, the behaviors would actually occur within the Tutor Module rather than the Learning Management System Module. An AIS could then use reinforcement learning (running within the Pedagogical Module) to determine which situations are the most beneficial for applying these strategies.

Further study of teachers' behaviors with dashboards would likely be beneficial to research and development along these lines. For example, it may be possible to create activities similar to those used by Kelly et al. (2013), where performance within *ASSISTments* is used to drive the selection of problems to work through, providing a practice experience to one student and a worked-out erroneous example (Adams et al., 2014) to other students. Instruction modeling analysis (Khachatryan, in press) can be conducted on the practices of teachers such as Dr. Kelly to provide a foundation for creating an AIS that can replicate observed pedagogical strategies, in the case of GIFT within the Pedagogical Module.

The future of AISs is strong. However, in developing AISs that can improve themselves, it is worth considering the example of how humans already enhance existing instructional systems. By doing so, we may be able to speed the enhancement of these systems and understand which enhancements work to the ultimate benefit of the learners.

References

Adams, D. M., McLaren, B. M., Durkin, K., Mayer, R. E., Rittle-Johnson, B., Isotani, S., & Van Velsen, M. (2014). Using erroneous examples to improve mathematics learning with a web-based tutoring system. *Computers in Human Behavior*, *36*, 401–411.

Al Otaiba, S., Connor, C. M., Folsom, J. S., Greulich, L., Meadows, J., & Li, Z. (2011). Assessment data-informed guidance to individualize kindergarten reading instruction: Findings from a cluster-randomized control field trial. *The Elementary School Journal, 111*(4), 535–560.

Alphen, E. V., & Bakker, S. (2016, May). Lernanto: Using an ambient display during differentiated instruction. In J. Kaye, & A. Druin (Eds.), *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 2334–2340). San Jose, CA: ACM.

Arnold, K. E., & Pistilli, M. D. (2012). Course signals at Purdue: Using learning analytics to increase student success. In S. B. Shum, D. Gasevic, & R. Ferguson (Eds.), *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 267–270). Vancouver, Canada: ACM.

Chen, Y., Chen, Q., Zhao, M., Boyer, S., Veeramachaneni, K., & Qu, H. (2016, October). DropoutSeer: Visualizing learning patterns in Massive Open Online Courses for dropout reasoning and prediction. In G. Andrienko, S. Liu, & J. Stask0 (Eds.), *IEEE Conference on Visual Analytics Science and Technology* (pp. 111–120). New Brunswick: IEEE.

Denley, T. (2013). Degree compass: A course recommendation system. *Educause Review Online*.

Feng, M., & Heffernan, N. T. (2006). Informing teachers live about student learning: Reporting in the ASSISTment system. *Technology Instruction Cognition and Learning*, *3*(1/2), 63.

Flach, J. M., & Voorhorst, F. (2016). *What matters*. Dayton, OH: Wright State University Library.

Halverson, R. (2014). Data-driven leadership for learning in the age of accountability. In A. J. Bowers, A. R. Shoho, & B. G. Barnett (Eds.), *Using Data in Schools to Inform Leadership and Decision Making* (pp. 255–266). Charlotte, NC: Information Age Publishing Incorporated.

Hawn, M.A. (2019) *Data-wary, value-driven: Teacher attitudes, efficacy, and online access for data-based decision making* (Unpublished doctoral dissertation). Teachers College, Columbia University, New York.

Heffernan, N. T., & 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.

Holstein, K., McLaren, B. M., & Aleven, V. (2018). Student learning benefits of a mixed-reality teacher awareness tool in AI-enhanced classrooms. In C. P. Rosé, R. Martínez-Maldonado, H. U. Hoppe, R. Luckin, M. Mavrikis, K, Porayska-Pomsta, B. McLaren, & B. du Boulay (Eds.), *Proceedings of the International Conference on Artificial Intelligence in Education* (pp. 154–168). London, England, UK.

Kay, J., Maisonneuve, N., Yacef, K., & Reimann, P. (2006). The big five and visualisations of team work activity. In M. Ikeda, K. D. Ashley, & T. W. Chan (Eds.), *International Conference on Intelligent Tutoring Systems* (pp. 197–206). Berlin, Heidelberg: Springer.

Kelly, K., Heffernan, N., D'Mello, S., Namias, J., & Strain, A. (2013). Adding teacher-created motivational video to an ITS. In C. Boonthum-Denecke, & G. M. Youngblood (Eds.), *Proceedings of the Twenty-Sixth Florida Artificial Intelligence Research Society Conference*. pp. 503–508. St. Pete Beach, FL: AAAI Press.

Khachatryan, G. (in press) *Instruction modeling and blended learning.* Cambridge, MA: Oxford University Press.

Kim, J., Jo, I. H., & Park, Y. (2016). Effects of learning analytics dashboard: analyzing the relations among dashboard utilization, satisfaction, and learning achievement. *Asia Pacific Education Review*, *17*(1), 13–24.

Lonn, S., Aguilar, S. J., & Teasley, S. D. (2015). Investigating student motivation in the context of a learning analytics intervention during a summer bridge program. *Computers in Human Behavior*, *47*, 90–97.

Mandinach, E. B., & Jackson, S. S. (2012). *Transforming teaching and learning through datadriven decision making*. Thousand Oaks, CA: Corwin Press.

Marsh, J. A., & Farrell, C. C. (2015). How leaders can support teachers with data-driven decision making: A framework for understanding capacity building. *Educational Management Administration & Leadership*, *43*(2), 269–289.

Martinez-Maldonado, R., Clayphan, A., Yacef, K., & Kay, J. (2015). MTFeedback: Providing notifications to enhance teacher awareness of small group work in the classroom. *IEEE Transactions on Learning Technology, 8*, 2. pp. 187–200.

Mavrikis, M., Gutierrez-Santos, S., & Poulovassilis, A. (2016). Design and evaluation of teacher assistance tools for exploratory learning environments. In D. Gašević, & G. Lynch (Eds.), *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 168–172). Edinburgh, Scotland: ACM.

Miller, W.L., Baker, R., Labrum, M., Petsche, K., Liu, Y-H., Wagner, A. (2015) Automated detection of proactive remediation by teachers in reasoning mind classrooms. In J. Baron, G. Lynch, & N. Maziarz (Eds.), *Proceedings of the 5th International Learning Analytics and Knowledge Conference*, 290–294. Poughkeepsie, New York: ACM.

Milliron, M. D., Malcolm, L., & Kil, D. (2014). Insight and action analytics: Three case studies to consider. *Research & Practice in Assessment*, *9*, 70–89.

Molenaar, I., & Knoop-van Campen, C. (2017). Teacher dashboards in practice: Usage and impact. In *European Conference on Technology-Enhanced Learning (EC-TEL)* (pp. 125–138). Tallinn, Estonia: Springer.

Singh, R. P. (2018). Learning Analytics: Potential, Protection, and Privacy in the Educational System. In M. K. Singh, Z. Zerihun, & N. Singh (Eds.), *Impact of Learning Analytics on Curriculum Design and Student Performance* (pp. 1–18). Hershey, PA: IGI Global.

Xhakaj, F., Aleven, V., & McLaren, B. M. (2017). Effects of a teacher dashboard for an intelligent tutoring system on teacher knowledge, lesson planning, lessons and student learning. In E. Lavoué, H. Drachsler, K. Verbert, J. Broisin, & M. Pérez-Sanagustín (Eds.), *European Conference on Technology-Enhanced Learning* (pp. 315–329). Tallinn, Estonia: Springer.