Computational Infrastructures for School Improvement: How to Move Forward

R. Benjamin Shapiro, Hisham Petry, and Louis M. Gomez

{rbs, h-petry, l-gomez}@northwestern.edu

Learning Sciences, Northwestern University

Abstract. The instructional practices common in today's schools reveal a disconnect between instruction and evidence of the effects of that instruction on student learning. In this paper, we propose the creation of computational infrastructures that will help teachers make more informed decisions in their practice. These infrastructures formalize student and teacher routines to facilitate data collection and mining, in order to create actionable information. We then show an instance of such a computational infrastructure and describe its potential for improving instruction.

1 Introduction

Studies of today’s (American, and many European) schools reveal a disconnect between instruction and evidence of the effects of that instruction on student learning. New constructivist understandings of learning, knowledge, and effective instruction [11] challenge school reformers to instantiate formative assessment and instruction by reconstituting the daily routines of teaching into ones "where a teacher's day-to-day decision-making is instrumentally constructed based on the interaction of detailed observations about students' work in the classroom (and the personal background students bring to their work) and the aims in view for subsequent instruction" [2]. Transitioning the existing American system of education into a system of practice where this occurs could be made more likely by the development of new Computational Infrastructures for School Improvement (CISIs) that help teachers to use, and reflect on their use of, evidence to make decisions. This paper describes a rationale and framework for such architectures, gives a concrete illustration of the implementation of such an infrastructure, and shows how it provides an opportunity to apply Educational Data Mining (EDM) techniques to some of the most challenging problems that teachers face.

This is not fundamentally an empirical paper. Rather, it is an argument for how we could organize the EDM field and the technologies it produces to have higher cumulatively and greater impact on educational practice and research. It is also not a claim about a specific technology; some of the ideas here about feedback and interactive support are manifested in some form in existing work, such as Interactive Tutoring Systems. Instead, we are trying to make a broader claim about how attending to the daily practices of teachers, learners, and school leaders could guide the design of interconnected technologies and practices, allowing EDM to have a much broader impact than it otherwise might.

2 Problem of Practice

Teaching is a complex, difficult, and uncertainty-ridden job. As Higgens notes:

> One starting point for inquiry into the moral phenomenology of teaching is Philip Jackson’s famous observation that teachers make over 200 decisions per hour. If Donald Schon is right, all practices require ‘reflection in action’, but teaching takes the demand for improvisation to new levels. It is in large part this radical unpredictability of teaching which shapes its phenomenology. It does not seem to matter how many times one has taught. Each time one begins a class, there is that unique blend of excitement and dread

> caused by the unpredictability of what will unfold. Perhaps this explains the fetishisation of the lesson plan: we want to deny just how much is unplanned. [5]

In large part it is this uncertainty that makes teaching difficult and inhibits school improvement [9,12], for even the most principled and motivated teacher trying to implement Ambitious Instruction (i.e., trying to make instruction build upon learners' prior knowledge and needs while engaging learners in consequential work; [2]) faces the currently insurmountable challenge of making sense of the volumes of information that are created in the course of daily instructional practice. At the same time, today's information infrastructure through which evidence of past performance flows from class to class and year to year requires that a massive amount of data be discarded on a continuous basis. The rich detail of students' daily work is distilled into grades that provide extremely low resolution; one cannot deduce from a past grade which concepts a student mastered and which he or she did not. Consequently, teachers face at once a glut of temporally local information, too much to completely comprehend, and a dearth of specific information about past performance that could inform present decision making.

Real-time and retrospective EDM, coupled to new architectures and practices for collecting, displaying, and using data, may offer a way past this seeming paradox, leap-frogging the current technology and practice of schooling by helping practitioners to make sense of the volumes of information generated through daily student and teacher practices. A growing number of scholars are discussing data and data use to quantify problems of achievement deficiencies, but few have investigated which data teachers should use and how they should use it to solve problems. Where does the teacher get the most decision-making bang for the data-analytic buck? What problems are students having? Do the problems that students are facing indicate a need to change teaching practice? Are they indicative of gaps in prior understanding? What kind of instruction has helped students with similar misunderstandings solve problems? What are the data that can reliably yield that information in order to support improved teaching? Thus far, most of the tools that support student thinking are either within single domains (such as cognitive tutors e.g., [3,6]) or systems that don't provide information in a way that lets teacher compare across classes and domains to examine their practices (Blackboard, for example, could do this, but does not). We are proposing building integrative architectures that allow information from such systems, used across disciplines/subject areas, over time and across space, to deeply understand students' intellectual development and the instructional activity that is necessary to support it.

Fullan, Hill, and Crevola also perceive instructional technologies to be on the cusp of dramatic improvement [4]. They believe that we now know enough about effective teaching to make large strides through increases in personalization, precision, and professional learning. At the same time that teachers must manage an entire classroom of students, students still have individual skills, proficiencies, and deficits that require more personalized instruction. To respond to those needs, teachers must be precise in their actions, which they suggest requires more precise information. This information could come from data mining, analyzing student work (data) to transform it into action by the teacher, reducing the uncertainty and increasing the precision of practice.

3 Moving From Data to Action

As described above, supporting students' learning is difficult due to the complexity of extracting actionable information from the raw data of student activity. Actionable information is information that directly supports decision making by entailing specific actions in regard to some problem at hand.

In order to situate technology development (i.e., the design of technologies to create this actionable information) within the broader scope of school improvement support, we conceptualize the work of teaching and learning as sets of routines enacted by intentional agents, facilitated by tools which provide a mechanism for action (agents work through the tools) and are a source of actionable information. Agents do their work with particular goals and knowledges, which are by-products of agents' social/historical contexts and their individual experiences.

Given the right circumstances (e.g., time, need, context...), an expert agent will interpret particular observed information as indicative of specific facts about the nature of the world (including the knowledge of other actors), which have specific implications for action. For example, a particular student response *X* to some task likely indicates some student conceptualization *Y* about the subject matter which has instructional implications *Z*. It is in the sense that from *X* follows *Y* and *Z* that *X* can be said, through the teacher's knowledgeable perception, to be actionable information. And it is because the observing teacher can see *X* as indicative of *Y* and that understanding *Y* warrants *Z* that we can say the teacher is an expert. Another, less expert, teacher might not notice *X* at all or, noticing *X*, might come to some other conclusion *Y'* and action *Z'*. Even the expert teacher, pressed for time or otherwise distracted, might not notice *X*'s significance (its actionability). We believe that tools can encapsulate knowledge (expertise) and, through so doing, can help practitioners to notice more and act more expertly, by guiding action toward Z. Tools that encapsulate this expertise, and help teachers to notice the instructional implications of student work, may help teachers to implement cycles of formative assessment and instruction.

Tools providing actionable information articulate/encapsulate an assembly model by suggesting what the meaning of some pieces of information (raw data) are and what the pedagogical responses to that meaning are (ala *X*, *Y*, and *Z*, above). In order to generate actionable information to guide learning and teaching, our design and evaluation work should include helping teachers to build routines and tools that engage learners in tasks that evoke evidence about their literate and subject-area development and provide actionable renderings of this evidence to guide pedagogical responses to it (i.e., instructional routines informed by the evidence, mediated by the knowledge encapsulated in the tools). Data mining and information visualization techniques form the core of such tools; a crucial empirical challenge, yet to be solved, is exploring the space of such tools to know which can most easily support more informed teacher practice.

4 SPACE, A Prototype CISI

Working with teachers using participatory design methods [7], we have created the beginnings of a CISI, called SPACE. SPACE was originally created to support many of the most information intensive aspects of project-based Inquiry teaching and learning, including teachers' planning, assignment specification, artifact management, formative feedback and assessment, reflection, as well as students' planning, performance, revision, peer critique and assessment. SPACE is a database-backed web application that so far incorporates simple data aggregation and visualization techniques. The case of its design and implementation to support practice at a middle school in science, social studies, and literacy makes clear two things. First, there is evidence that CISIs support a paradigm shift in teachers' use of evidence for improving instruction and second EDM and data visualization techniques are at the core of what makes CISIs useful. SPACE's underlying information architecture is roughly captured by Figure 1.

In SPACE, project assignments are comprised of ordered sets of task assignments. Each task is subject to a set of assessment criteria, each of which is related to one or more standards of goals (the standards may be State or Federal Standards or local constructs). As students (working alone or in groups) do their work in SPACE, the electronic artifacts they create (the student tasks) are related to the appropriate task template as well as to the authors themselves. Teachers' and students' assessments of work, along with information about (co-)authorship, form a kind of incidental social network, the implications of which we discuss below.

The SPACE user interface makes it easy for students to find out what work they need to do and what criteria will be used to assess it. Teachers and students may easily browse the database to see the work of others and to assess that work. Aggregate representations make it simple for teachers to discover which skills need the most instructional focus. At present, these aggregations are of the most simplistic kind, showing, for each standard or assessment criterion, on a by-student, by-assignment, or by-cohort basis, the mean, standard deviation, and trend over time. Presently SPACE employs relatively simple EDM techniques, but as we discuss next, combining the structured data of a system like SPACE with advancements in EDM could dramatically increase the quality and quantity of actionable information available to teachers.

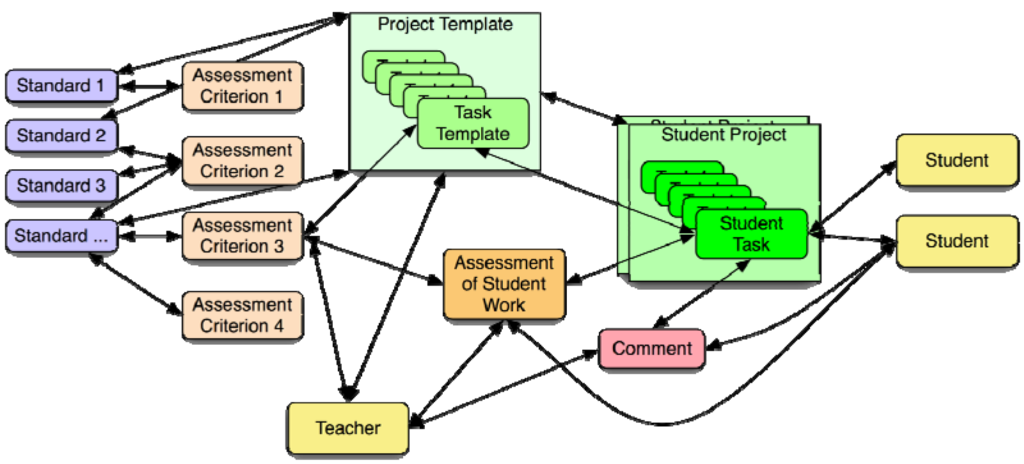

Figure 1: SPACE Information Architecture

5 SPACE in Practice

SPACE was intended to help teachers enact cycles of formative instruction throughout the course of project-based Inquiry. The following example arose during a recent classroom implementation of SPACE supporting a science fair project: Students are working on different projects and they are at different stages of progress. Students submit work to be assessed by their peers and the teacher. The work can then be revised and resubmitted as needed to complete the task. Managing the student schedules is critical because the science fair represents a hard deadline that must be met. This is a lot of data for a teacher to keep track of, and students (and the teacher) can easily fall behind.

Figure 2 shows a visualization of an eighth grade class midway through a project. The stages represent progress towards the science fair presentation. As an example of what these stages represent, stage one is to read some current events and find something technological that excites the student and take some notes on it (e.g. Nike's new tennis shoe). Stage two is finding a related area of interest (e.g. materials science), and stage three is a project proposal (e.g. comparing springiness of different materials). From there stage ten and eleven are building a data table and collecting data for analysis. Finally, stage 14 and 15 is the write up and presentation of the work. At the time that this figure was generated several students were starting on stage 10.

A tremendous amount of data is being generated: the student work (the pieces of work individually, but also chains of revisions), the critiques, rubric-based assessments, and logs of who's looking at what. This data must be analyzed and presented in a way that is useful to the teacher to become actionable knowledge. Analysis and visualization techniques appraise teachers of students that are waiting for teacher feedback in order to go forward. Other visualizations indicate whether the student is ahead or behind schedule for the science fair.

Looking at this visualization the teacher can know immediately which students are falling behind or need attention. The 'w' indicates tasks that are waiting for teacher feedback. The large ratio of 'w's to non-'w's indicates that this specific teacher is significantly behind in offering feedback, and may be overwhelmed. Moreover, the number of pink and red boxes illustrates that students are struggling to turn work in on time; the number of lagging revision required symbols illustrates that many students have been asked to revise but have not, perhaps because they are confused about what the teacher wants them to do.

The representation in Figure 2 is a high-level view of the underlying data; it foregrounds punctuality and task status, while backgrounding the actual content of the work and assessments thereof. Teachers can click the column headings, author names, or project titles to see all instances of work on a task (across projects), to see information about the authors (including links to other work), or all work on the project, respectively. This representation is only one of many possible ways of aggregating project work. We have already implemented a skill-based aggregator that shows summary statistics about how students' work (at the individual student, project, or cohort level) has been assessed, on a skill-by-skill basis, showing teachers what skills students have mastery of and which they need additional support in. Because all interfaces are massively hyperlinked, it is easy for teachers to find concrete examples of relevant work. We imagine the application of a number of information retrieval and data-mining techniques to provide additional depictions of the data. The following section describes several possibilities.

6 A Way Forward

Singley and Lam describe a number of interesting heuristics that could be helpful to practitioners trying to decide where to direct their attention [14]. Their Classroom Sentinel alerts teachers to conditions, identified through data mining, such as a student's grade dropping significantly from past assessments, a struggling student performing above average on assessments, the student being within range to increase their letter grade if they do well on the next assignment, or ESL students performing below average on a mathematics assessment. [8] describes a framework for data driven decision-making in schools that "highlights three key components: the process by which raw data becomes useable information, the role of prior knowledge of the decision-maker, and the effect of the data-reporting tool in shaping that process”. We try here to describe, in a more concrete fashion, what infrastructure for such decision-making might be.

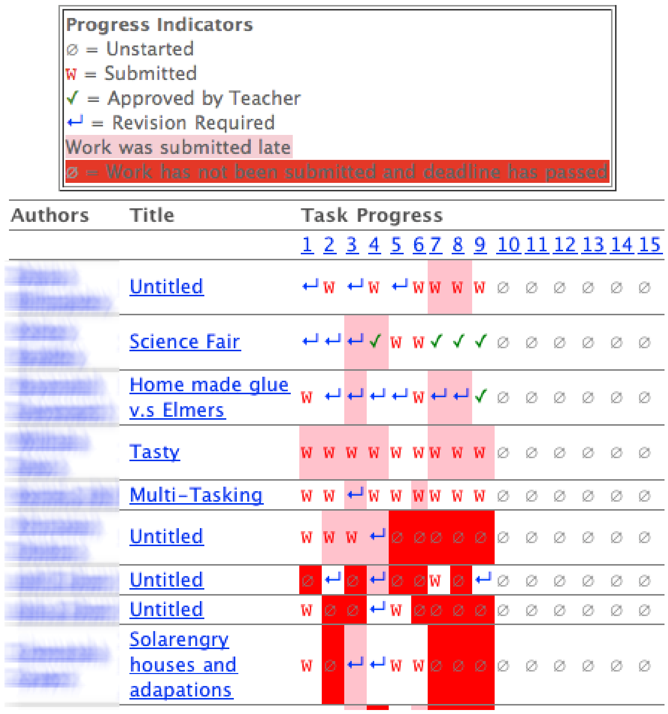

Figure 2: SPACE Project Progress View (for Cohort)

The data structure behind SPACE affords a number of powerful analyses each of which promises to deliver actionable information to teachers and leaders. These analyses range from discovering students' mastery of necessary prior knowledge or related activities, to the leveraging students' social networks, to recommendations of pedagogical options, each of which we now describe.

The SPACE database makes it simple for teachers to connect summary statistics to concrete examples of student work. For example, the Illinois state standards for literacy state that middle school students should be able to "identify appropriate resources to solve problems or answer questions through research"; this standard is relevant to any activity where students must collect sources "from the wild" (e.g., the Internet) in order to do their work. Consequently, in a full implementation of SPACE across a school or district, there would be multiple assignments, in a variety of subject areas, that are related to the standard. Through SPACE, a teacher can see not only summary statistics about students' performances related to the standard, but also concrete examples of work. Teachers could search for all examples (in his/her own class or in others') of poor performance in selecting appropriate articles, or find out which students tend to be poor selectors.

Having this rich data offers interesting opportunities for increasing the (measurement) validity of teachers' assessment practices. Ideally, there would be a high correlation between teachers' assessment of students' work with respect to some set of standards and other measures of those same students' skill with respect to those same standards. For example, if a recent assessment says that a student has trouble summarizing texts, then we'd expect that teacher assessments of that student's summaries of texts should be similarly low. Divergence between the two measures may indicate a disconnect between teacher understanding of standards and how those standards are measured. Tools like SPACE would allow teachers to take problem areas identified by standardized tests, and then comb through a student's work to verify similar pre-existing issues. This verification process helps teachers identify and recognize problem areas for present and future students. Consequently, tools like SPACE offer the potential to allow teachers to reflect on their practice as well as make standardized tests formative rather than summative assessments.

Education research has found patterns in peer groups and achievement using social network analysis [10,13]. The SPACE database also affords a number of social network analyses. For example, students' critiques of each others' work might represent an edges between students. We can use the amount of subsequent improvement in students' work as the strength of the edges. We can then analyze a network constructed in such a manner over time to understand which student's feedback influences another student (i.e., students' whose critiques are useful to a broad spectrum of peers). This information could be highly actionable for teachers, supporting decision-making about instruction (e.g., knowing who needs help giving good critiques) and classroom management (e.g., which students to pair up so that they'll be maximally mutually supportive).

CISIs have potential beyond what we have described here. Not only could they better support the daily cycles of instruction, but they can also be used to aggregate information at the school level to make decisions. For example they could assist school leaders making decisions about how to allocate discretionary support resources (e.g., teachers' aids, additional training, supplemental money for extra teacher hours). Similar to how teachers have an increased awareness of students' learning, leaders could spot school level problems early on (such as the teacher falling behind in the above example) and intervene to be proactive instead of reactive.

7 Conclusion

Personalizing instruction for students is demanding of teachers. Presently teachers do not have tools that support their daily instructional practice. Computational infrastructures that merge data storage, mining, and presentation can help teachers manage classroom data to make more informed and responsive decisions. Through the use of these sorts of tools, levels of instruction that previously required vast pedagogical content knowledge and heroic effort could now be much more reasonably achieved, with benefits for every student.

8 References

[1] Anjewierden, A., Kolloffel, B., & Hulshof, C. (2007). *Towards educational data mining: Using data mining methods for automated chat analysis to understand and support inquiry learning processes.* Paper presented at the International Workshop on Applying Data Mining in e-Learning, Crete, Greece.

[2] Bryk, A., Gomez, L., Joseph, D., Pinkard, N., Rosen, L., Walker, L. (2006). *Activity Theory Framework for the Information Infrastructure System*. Unpublished Manuscript, Accessed April 7, 2008. http://www.iisrd.org/documents/IIS_Activity_Framework_2006.pdf

[3] Constantino-Gonzalez, M. and Suthers, D.D. (2003). Automated Coaching of Collaboration Based on Workspace Analysis: Evaluation and Implications for Future Learning Environments. *Proceedings of the 36th Hawaii International Conference on System Sciences (HICSS'03)*.

[4] Fullan, M. G., Crévola, C., & Hill, P. (2006). *Breakthrough*: Corwin Press.

[5] Higgins, C. (2003). MacIntyre's Moral Theory and the Possibility of an Aretaic Ethics of Teaching. *Journal of Philosophy of Education, 37*(2), 279-292.

[6] Kay, J. and Maisonneuve, N. and Yacef, K. and Zaïane, O. (2006). Mining patterns of events in students' teamwork data. *Proceedings of the Workshop on Educational Data Mining at the 8th International Conference on Intelligent Tutoring Systems (ITS 2006),* Jhongli, Taiwan

[7] Kensing, F., & Blomberg, J. (1998). Participatory Design: Issues and Concerns. *Computer Supported Cooperative Work (CSCW), 7*(3), 167-185.

[8] Light, D., Wexler, D. H., & Heinze, J. (2004). Keeping Teachers in the Center: A Framework of Data-Driven Decision-Making. *Education Development Center (EDC)*, 6.

[9] Lortie, D. C. (2002). Schoolteacher: A Sociological Study.

[10] Petry, H. J., & Gomez, L. (2008). *On par with peers: Social dynamics of players and workers.* Paper presented at the AERA, New York, NY.

[11] Resnick, L. B., Lesgold, A., & Hall, M. W. (2005). Technology and the New Culture of Learning: Tools for Education Professionals. In P. Gärdenfors & P. Johansson (Eds.), *Cognition, Education, And Communication Technology*.

[12] Rosenholtz, S. J. (1991). *Teacher's workplace: the social organization of schools*: Teachers College Press.

[13] Ryan, A. M. (2001). The peer group as a context for the development of young adolescent motivation and achievement. *Child Development, 72*(4), 1135-1150.

[14] Singley, M. K., & Lam, R. B. (2005). The classroom sentinel: supporting data-driven decision-making in the classroom. *Proceedings of the 14th international conference on World Wide Web*, 315-321.

[15] Stiggins, R. J. (2002). Assessment Crisis: The Absence of Assessment for Learning. *Phi Delta Kappan, 83*(10), 758-765.