Scripting and orchestrating learning communities:  
A role for learning analytics

James D. Slotta, Boston College, slotta@bc.edu
Alisa Acosta, University of Toronto, alisa.acosta@utoronto.ca

Abstract: This paper describes our efforts to add structure and formalism to the design of a CSCL curriculum for high school science—integrating individual, collaborative and whole-class inquiry activities into a coherent “learning community.” A pedagogical model called Knowledge Community and Inquiry (KCI) guided our design of a curricular sequence in which one activity feeds into the next, responding differentially to students, and scaffolding new forms of interaction. We include real-time analysis of student interaction data as a source of input into the orchestration of complex scripts, which can influence the assignment of students to groups, the distribution of materials or sequencing of activities. It can also be used to determine which groups may need help, to provide groups with formative feedback, and to provide the instructor with information concerning student groups. The primary outcome of this paper is the design itself, which is evaluated in terms of its theoretical coherence.

Introduction

By now, most educators have heard about the need to foster “21st century knowledge skills,” such as critical thinking, collaborative problem solving, and evidence-based reasoning (Hargreaves, 2003; Pellegrino, Hilton, & others, 2013). The world of science, in particular, has become infused with new technologies and information practices, data-intensive methods and large, multidisciplinary collaborations distributed across space and time (e.g., the Human Genome Project, astronomical mapping, climate tracking). In response, many have argued that traditional modes of science instruction are inconsistent with the task demands of science and the wider STEM workplace (Collins & Halverson, 2010; diSessa, 2001). Thus, science education should help students develop relevant skills and literacies, in addition to basic skills and factual knowledge (NGSS, 2013).

Science educators have responded to this challenge, exploring new modes of learning and instruction. “Active Learning” (Bishop & Verleger, 2013; Charles et al., 2011; DeLozier & Rhodes, 2016), is one such approach that has now engaged many STEM educators, resulting in professional societies (e.g., SALTISE.ca) and university-based centers to support the design of active learning courses. Ruiz-Primo et al. (2011) summarize active learning as comprising four dimensions: (1) conceptually oriented tasks, (2) collaboration, (3) technology, and (4) inquiry based projects. Several studies have now measured the benefits of active learning (Code, Piccolo, Kohler, & Maclean, 2014; Dori & Belcher, 2005; Linton, Pangle, Wyatt, Powell, & Sherwood, 2014). Freeman et al. (2014) performed a meta-analysis of active learning in STEM, finding that exams scores improved by 6% and students were 1.5 times less likely to fail compared with traditional lecture approaches.

Despite this evidence of success, however, active learning remains largely ill-specified and difficult to study with any control (Brownell, Kloeser, Fukami, & Shavelson, 2013; Ruiz-Primo et al., 2011). For example, while specific group strategies are often invoked (e.g., cooperative learning, groups, gallery walks, collaborative projects, problem solving or case based learning) very little definition is provided about the learning processes, materials or assessments, nor about the instructor’s role during the activities (Henderson & Dancy, 2007). Simply naming those collaborative approaches fails to provide sufficient detail about the content, structure or sequencing of activities. What makes a hands-on lab activity effective? When should it be used within the active learning sequence? How will students collaborate, and to what end? How should design projects be structured, and how should they be assessed? Practitioners and researchers require more detail about the curricular designs, in order to develop a deeper understanding of active learning.

This paper describes our recent efforts to add structure and formalism to the study of active learning, as we co-designed (i.e., with the teacher) a new high school biology curriculum that integrated individual, collaborative and whole-class inquiry activities into a coherent “learning community” design. While our curricular design is currently in the process of being enacted and studied as part of a broader research program, this paper focuses on the role of a pedagogical model in guiding a CSCL design. In that regard, the paper has a theoretical focus, although the specific design (i.e., of our active learning biology curriculum) can be considered an empirical outcome. We focus on the important questions of how a curriculum design can be constructed in a principled way that weaves together the different forms of activities into a coherent sequence, in which one activity feeds into the next, responding differentially to students, and scaffolding new forms of interaction for instructors.
Thus, the primary outcome of this paper is the design itself, which can be evaluated in terms of its theoretical coherence.

We begin with a discussion of the theoretical perspective of learning communities, including our own theoretical framework, called Knowledge Community and Inquiry (KCI). We then introduce important notions of scripting and orchestration, and the role of a formal model in guiding the design of CSCL curricular scripts. We include an emphasis on the real-time analysis or processing of student interaction data, as a source of input into the orchestration of complex scripts (e.g., where students’ contributions on one activity may determine the condition or materials they are assigned to in a subsequent activity). We also focus on the important notion of group process analytics, for both scripting and orchestration processes. By introducing such real-time analytics of group process (e.g., whether the group is progressing according to the designed activities, whether all members are contributing, etc) we can add an important theoretical capacity to our scripting and orchestration of CSCL curriculum. This can influence the assignment of students to groups, the distribution of materials or sequencing of activities. It can also be used to evaluate, in real-time, the process of a scripted group interaction, in order to determine which groups may be on task, which may need help, to provide groups with formative inputs or feedback, and to provide the instructor with important information concerning the state of student groups. Our design-based research addresses the following questions:

1. What sequences of small and large group activities, including social media and technology-mediated learning, support a community of learners in our courses?
2. How can a learning community approach reinforce the lectures and other course activities, adding structure, coherence, and connections across topics?
3. What is the role of the instructor within these designs? Beyond simply acting as “guide on the side”, what forms of classroom discourse must the instructor emphasize? What conditions or markers of progress should be monitored to determine needed discussions or activity transitions?

Theoretical Background

Active learning has become a movement amongst secondary and post-secondary educators (Freeman et al., 2014), founded on constructivist and social constructivist learning principles (Bransford, Brown, & Cocking, 1999), and informed by deep pedagogical expertise within the specific disciplines. In the life sciences, a surge of interest has driven a growing community of scholars, as evidenced by online communities like LifeSciEd.org and the Society for Advancement of Biology Education Research (SABER). While many undergraduate biology educators have advocated flipped classrooms (e.g., Gross, Pietri, Anderson, Moyano-Camihort, & Graham, 2015; van Vliet, Winnips, & Brouwer, 2015), others have cautioned that flipping alone will not improve student outcomes, unless accompanied by effective learning designs in the classroom (Jensen, Kummer, & Godoy, 2015).

One prominent form of active learning in biology education is concerned with the enhancement of whole-class discussion and lectures. The most effective lectures engage students in responding to questions, where the instructor “re-voices” their ideas, blending multiple responses, and bridging to new topics. The nature of instructor-led discourse, sometimes referred to as “accountable talk” (Michaels, O’Connor, & Resnick, 2008), has been a topic of growing interest for educational researchers. In biology as in other disciplines, the use of audience response systems (“clickers”) has greatly increased the opportunities for instructor-led discussions that connect to student ideas (Smith et al., 2009). Following the wealth of work from the physics education community (i.e., the use of clickers and peer instruction methods), biology educators are also studying these methods. Giuliodori et al. (2006) incorporated peer instruction discussions four times during each 90-minute physiology lecture, resulting in statistically significant positive gains on qualitative questions. In a paper titled “Teaching more by lecturing less,” Knight & Wood (2005) reported improved student outcomes in an upper level developmental biology course from the integration of collaborative problem solving and whole-class discussions. Similarly, Gardner & Belland (2012) observed that these various techniques work best when applied synergistically to create an active learning environment for students.

Theoretical Framework: Knowledge Community and Inquiry

Another promising approach to the design of active learning is to consider the entire classroom as a learning community, in which students draw upon their diverse interests and expertise with a common goal. They share the understanding that their learning activities will align to advance the community’s cause, while at the same time helping individuals learn, and allowing everyone to benefit from the community’s resources (Bielaczyc & Collins, 2005). In a review of learning community models, Slotta & Najafi (2013) articulated three common characteristics: (1) An epistemic commitment to collective advancement, (2) a shared community knowledge base, and (3) common modes of discourse. Several scholars have observed that it is challenging for teachers or researchers to coordinate a learning community approach (Slotta & Najafi, 2013; van Aalst & Chan, 2007). As
observed by Kling and Courtright (2003, p. 221) “developing a group into a community is a major accomplishment that requires special processes and practices, and the experience is often both frustrating and satisfying for the participants.” The limited success or uptake of this approach has been due to the pragmatic and epistemic challenges of shifting from a didactic mode of “knowledge transmission” into one of collective inquiry. But it is also due to the lack of explicit models to guide the design of curriculum where students are interconnected in a progression of individual, small group and whole class activities, creating and consuming materials from a community knowledge base (Slotta & Peters, 2008).

The Knowledge Community and Inquiry (KCI) model was developed to guide the design of such curricula, in which the whole class (or even multiple class sections) work together, with all students held accountable for content learning gains (Slotta, 2014; Slotta & Najafi, 2013; Slotta & Peters, 2008). The model includes principled requirements for (1) a knowledge base that is indexed to the targeted science domain (2) collective, collaborative and individual inquiry activities in which students co-construct the knowledge base and then use it as a resource for further inquiry, and (3) assessable learning outcomes that allow teachers to evaluate student progress. KCI curricula typically span multiple weeks or months, and are developed through a sustained process of co-design (Roshelle, Penuel, & Shechtman, 2006) amongst researchers, teachers and designers. Within KCI curriculum, inquiry activities are designed to engage students individually and in small groups where they make use of their community knowledge base as a resource. The designed curriculum constitutes a “script” that includes student-contributed content, social media, and small-group activities such as design, debate, critique, argumentation and reflection. The script is “orchestrated” by the instructor, who is enabled, in turn, by features within the physical environment (e.g., large screen projections of students’ pooled votes, resources or other products) as well as the technology environment, which can help track student progress, distribute instructions and prompts, pause students for planned or spontaneous discussions, etc. The orchestration of the script often depends upon in-the-moment decisions by the instructor, whose role is one of collaborator and mentor, responding to student ideas as they emerge, and orchestrating the flow of activities. Teachers are not just a “guide on the side” but rather have an explicitly scripted role at all times, as well as responsibility for overall coordination.

Prior KCI studies have investigated various forms of learning content, activities and environments, including mobile technology applications for student-contributed observations (e.g., forms, photographs, notes, votes, tags), large, projected “emergent representations” of the collective knowledge, and various forms of classroom instrumentation (Cober, McCann, Moher, & Slotta, 2013; Fong et al., 2013; Moher et al., 2015). Students are typically engaged in computer-supported inquiry activities, including note taking, observations, brainstorms, problem solving, modeling and simulation, design and argumentation (Slotta, Tissenbaum, & Lui, 2013). Large projected displays help teachers identify pedagogically meaningful signals from amidst the noise of student contributions, and track the community’s learning progress.

KCI research has produced a technology environment called Common Knowledge (CK) that includes server software that captures student contributions (i.e., the knowledge base), and a wide range of Web applications for students and teachers that support the collection, distribution, curation and application of that content. CK is a “bespoke technology,” meaning that it was developed in close alignment with the epistemological commitments of the model, for purposes of the research, and so provides a good fit for the complex activity sequences and dependencies on student interactions that are required by KCI designs. In recent versions of CK, the technology architecture has been improved to allow interoperability with many other platforms, including shared authentication (i.e., using the LTI standard). This allows our designs to include a variety of tools or other platforms, as well as the existing functionality offered by CK or new features that can be readily developed. CK provides a flexible foundation for technology-mediated collective inquiry, which has been extended and applied in the current work, supporting a rich array of biology learning materials, activities and interactions.

**Scripting and Orchestration**

One area of research from the learning sciences that is central to our designs are the concepts of scripting and orchestration (Dillenbourg & Jermann, 2007; Kollar, Fischer, & Slotta, 2007). Similar to a theatrical script, which specifies all aspects of a play (i.e., stage, props, lines, actions, and attitudes), a pedagogical script explicates a learning design in terms of the participants, roles, goals, groups, activities, materials, and logical conditions or determinants of activity boundaries (Fischer et al., 2013). Like its theatrical counterpart, a pedagogical script is only an abstract or idealized description…until it is performed. Orchestration refers to the enactment or coordination of the script, binding it to the local context of learners, classrooms, curriculum and instructor, and giving it concrete form in terms of materials, activities and interactions amongst participants (Tchounikine, 2013). Pedagogical scripts are orchestrated in the classroom, online or across contexts (i.e., home, school, or mobile), with the “orchestral load” shared or distributed across several agents: (1) the instructor, who can tell students what to do, pause activities to hold short discussions, or advance the lesson from one point in the script to another;
(2) the materials, including text or other media, instructions, or interactive Web sites; (3) the technology environment, including online portals, discussion forums, note sharing or wiki environments, Google Docs, etc.; and (4) the physical learning environment (i.e., classroom configuration, furniture, walls, lighting). The notions of scripting and orchestration can inform our design of active learning, encouraging specificity about the materials, activities and sequencing, as well as deep understanding about the role of the instructor, and any scaffolding environments (Slotta, 2014).

**Learning Analytics**

While technology is often invoked as an important ingredient of active learning, the specific role of technology is seldom explicated (i.e., in terms of how it scaffolds individual or group work, the role of technology-enhanced media in student learning, or the best practices for instructors in working with any given technology). Indeed, many current models of Active Learning de-emphasize commitments to specific technologies, focusing on flexible classroom configuration, table-group collaborations with whiteboard surfaces and paper-based problem solving. Students may have their own laptops, and engage with any number of tools and materials, but the technology itself is not intrinsic to the design. While these approaches may be practicable and engaging, they fail to capitalize on promising new media like social networks, learning analytics, user-contributed content, tangible and embodied interactions, and “gameful design” (Fishman & Deterding, 2013). In such approaches, technology can play a central mediating role, supporting functions or features that would not otherwise be possible, connecting students and enabling real-time processing of student interactions (e.g. to inform new groupings or distribution of materials). In the past few years, the field of learning analytics has grown as a specialized discipline focused on “the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs” (Ferguson, 2012, p. 305). One application of learning analytics relevant to scripting and orchestration is in the support of adaptive learning designs (i.e., scripts) in which students’ interactions with technology environments (i.e., click logs, Web form data, uploaded content, tags, votes, etc) are processed in real time to inform their assignment of materials, groups or activities (Lockyer, Heathcote, & Dawson, 2013).

**Methodology**

KCI informs the design of inquiry curriculum that engages a community of learners at three levels of granularity: (1) the individual level, (2) the small group level, and (3) the whole class (i.e. knowledge community). As described above, materials and activities at each level are carefully designed to promote the development and re-use of a community knowledge base. Individual activities may engage students in adding content to the knowledge base. Small group activities may divide students according to the levels of an important organizational variable, and ask them to sort and tag the elements in a knowledge base, or to apply the contents in some design project (e.g., designing a solution to some environmental problem). Whole class activities could entail brainstorming, sorting and tagging, or whole class discussions. The activities are all indexed to a common domain model that also provides the structure or indexing of the knowledge base itself. In this way, all activities and assessable outcomes are assured of promoting progress on the targeted science learning goals (Slotta et al., 2013).

We employed a design-based research methodology (Brown, 1992; Collins, 1992; Edelson, 2002), wherein we worked closely with a high school biology teacher and team of technology developers to co-design an innovative active learning curriculum and corresponding technology environment called CKBiology. The structure of our designs was guided by KCI, and is comprised of three distinct elements: (1) the content model – specific forms of user-contributed content, Web form elements, votes and tags, photos or other media, emergent learning objects, and connections to course elements like lectures, homework, quizzes or exams; (2) a process model – how groups will be formed, roles for students and instructor, content logic, feedback and materials, generation of emergent learning objects, and specific bindings to the content model; (3) a discourse narrative – a detailed description of the expected forms of interaction between students, peers and instructors, relating to any materials or activities (i.e., expected discourse patterns and amongst students, peers and instructor, and orchestrational roles for instructor and technology environment).

The articulation of the content model began by defining and parametrizing the content domain of the course, including pertinent aspects of scientific inquiry (e.g., for molecular genetics, identifying the impacts of a mutation), as well as inquiry skills like collaboration and problem solving. Next, we defined a knowledge base, indexed to the domain parameters, to ensure that all student contributions are directly connected to targeted content areas. Finally, we designed the inquiry script, including materials, activities and tools that linked explicitly to the knowledge base, and a community of learners, including students, groups, roles, and any relevant metadata.

A substantial head start on the technology environment was gained from the existing CK technology, including the capacity for collecting, aggregating and re-distributing any form data (i.e., text entry fields, image
uploads, radio buttons, check boxes, etc.), and fixed keyword tagging. To support real-time evaluation and feedback, we have built upon the current capacity of CK for learning analytics of individual and group activities. The next section outlines our designed curriculum, in terms of the three underlying models (content, process and discourse), including how we implemented learning process analytics to support scripting and orchestration for a learning community in high school biology.

Results

Content Model
The content model included a major index to five primary units of the course (i.e. biochemistry, metabolic processes, molecular genetics, homeostasis, and population dynamics), each of which comprised a set of lessons (see Figure 1). For example, the molecular genetics unit included lessons on DNA replication, protein synthesis, gene expression and regulation, and biotechnology. Each of these topics was further indexed in terms of core concepts, as shown in Figure 2. All concepts were defined by students and connected in a semantic Web, then systematically incorporated into inquiry activities in which students relied on the definitions and benefited from the semantic web (see Process Model section below). We adapted the Common Knowledge environment to create CKBiology, which supported students in working across contexts (home and school; small group and whole class), ensuring that all student contributions were added and indexed to the knowledge base, and that activities that could benefit from the knowledge base were able to do so.

![Figure 1: CKBiology home screen, depicting a series of lessons within one curricular unit. Progress bars corresponding to individual and community-level progress are shown in purple and blue, respectively.](image)

![Figure 2: Students’ contributions aggregated to a shared community knowledge base, serving as a resource for subsequent inquiry activities.](image)

Process Model
Students log into CKBiology at home to complete a series of tasks following each day’s regular classroom lesson. The home screen for each unit consists of a series of lessons, each displaying two progress bars; one depicting the student’s own individual progress, and the other depicting the progress of the whole knowledge community (see Figure 1). Upon selecting a lesson, students are assigned three different kinds of tasks: (1) Providing an explanation for a particular term or concept, (2) identifying the relationship between two terms or concepts, and (3) vetting explanations that have been contributed by other members of the knowledge community. The vetting task ensures that all students’ ideas are read, discussed and improved upon by others in the knowledge community. As students progress through their assigned tasks, their contributions are aggregated to the shared community knowledge base (see Figure 2). Students and the teacher can access this knowledge base at any time using the navigation toolbar at the top of the screen. In cases where vetting has led to a disagreement around a particular explanation, a yellow dot is added to the term or concept within the knowledge base screen, serving as a cue to the teacher that this may warrant a follow-up discussion in class the following day. A teacher dashboard has also been created, which provides an overview of each student’s progress as well as the state of the knowledge base.

A second important element of the process model is a series of in-class inquiry activities in which students individually read one of several current “real world” research articles, tagging terms and concepts from within the knowledge base (i.e., providing an explicit link to the domain content). Students then form small groups to negotiate their choice of tags and provide explanations as to how each term or concept applies within the context of the article (see Figure 3). Next, they form teams and complete a review “challenge” activity in
which they consolidate knowledge, applying the concepts they have learned within a new context of inquiry, and synthesizing their knowledge in response to a broad socio-scientific issue (e.g. climate change). The progress of each review challenge team is represented by a group-level progress bar, which is also available on the teacher’s dashboard (see Figure 4). Five distinct activity sequences were designed, each indexing to the core concepts, and engaging students in small group applications of that knowledge base. For each, we developed a group process model that tracked groups in terms of their overall process (completeness), successful coordination of the task (fidelity) and equity of participation.

**Figure 3**: Working in small groups, students negotiate how concepts from the knowledge base apply to their chosen article. Shades of blue represent levels of agreement among group members, with dark blue representing strongest agreement. Green tags reflect the end-product of the group’s negotiation efforts.

**Figure 4**: Teacher dashboard showing group-level progress bars for the in-class review activity. Clicking on an individual team’s icon will display the work that team has contributed so far, giving the teacher a deeper understanding of each group’s progress and when/where to intervene.

**Discourse Model**
The teacher portal to support at-a-glance information about the state of all groups in the various activities. We delineated specific determinants of teacher-led discourse, such as when no group had made any contributions for a specified amount of time, or when there were a given number of contested relationships in the knowledge base. We also expected extemporaneous discourse, as the teacher noticed opportunistic moments for intervention, based on information she received during small group visits, or by examining the teacher dashboard or student knowledge base. At present, our model for classroom discourse includes three primary dimensions: Small group discussion (students coordinating, with occasional teacher visits), whole class discussions (teacher initiated), and targeted mini-lectures, which emerge in response to revealed student misconceptions or lack of understanding.

**Implications and Next Steps**
The formal specification of learning designs has been elusive. Yet we know from nearly all other science disciplines that formalisms lead to greater progress in research, allowing reliable communication and opening the doors to a wide range of applications. For learning scientists, formal descriptions could allow for comparison of learning designs, or they could inform the creation of taxonomies of pedagogical structure. Without them, we are reduced to deciphering the descriptions offered by course designers (which vary in detail and granularity), to infer the structure of the underlying script (including material design, activity sequencing, dependencies or conditions, etc.). For science educators, it is important that our designs are appropriated widely by colleagues, in part to ensure fidelity of adoption, but also to encourage experimentation and adaptation. This is how innovations spread and evolve. This project hopes to make a contribution to the growing community of biology educators, offering one complete course design that is equipped with an underlying formal structure, adheres to a central pedagogical perspective (learning communities and inquiry), and advances particular forms of collective and small group engagement. Ultimately, the goal would be to support the exchange, uptake, adaptation and critical evaluation of such design, nurturing a learning community of biology educators who build their own knowledge base of innovative designs, validated assessments and shared understandings about learning and instruction.

**Acknowledgments**
This work was supported by the Social Sciences and Humanities Research Council of Canada.
References


