Envisioning a Learning Analytics for the Learning Sciences

Alyssa Friend Wise, New York University, alyssa.wise@nyu.edu
Yi Cui, New York University, yc65@nyu.edu

Abstract: There is both great excitement and substantial concern within the learning sciences about what educational data science and learning analytics (EDS/LA) have to offer our understanding of learning and ability to support it. This paper lays out three concerns often raised about the use of EDS/LA approaches in learning sciences work: reliance on algorithmic processing over human insight, attention to generalized structures over contextualized processes, and emphasis on empirical findings over theory building. Through an overview of work conducted on the MOOCeology project it then shows specific ways that such concerns can be meaningfully addressed. The paper concludes by elucidating a set of seven initial principles for "learning sciences aware" EDS/LA work and opening the question of to what extent these principles might be appropriate to guide EDS/LA work more broadly.

Introduction

There is both great excitement and substantial concern within the learning sciences (LS) about what educational data science and learning analytics (EDS/LA) methods have to offer our understanding of learning and ability to support it (Wise & Schwarz, 2017). On the enthusiastic front, proponents see these new techniques as offering powerful ways to find patterns in and made sense of the large amounts of data that technological learning environments (and technological data captured in physical learning environments) can now produce. From a more cautious perspective, doubts have been raised about the appropriateness and ultimate usefulness of these methods for generating deep insight into the complex cognitive and interpersonal processes of learning. While there are important issues raised on both sides of this debate, the simplistic discourse of "should we or shouldn't we (embrace EDS/LA as part of the methodological repertoire of LS)" misses the point and distracts us from the real conversation we should be having. Certainly there is research within EDS/LA that may be at odds with some of the core tenets of LS (for example learning theory plays a central role in LS research while it may or may not be central to an EDS/LA study; there are also questions of how theory is mobilized see Wise & Shaffer, 2015). But to declare EDS/LA approaches as fundamentally incompatible with the learning sciences unwisely ignores the exciting new avenues to understanding learning it offers (for example extending manual content analysis to examine learning processes across many more contexts than is currently possible, generating in-the-moment feedback tailored to the learning activities of individuals and small groups), and closes any conversation of how it could thoughtfully be put to use. In other words, the discussion that needs to occur is not one of *if* EDS/LA approaches are appropriate for LS, but *how* we can develop such approaches, as well as norms and practices surrounding their use, in ways that can make valuable contributions to the field.

This paper takes a first step in that direction by laying out three concerns often raised about the use of EDS/LA approaches in LS work. In doing so, it is important to note that EDS/LA is still very much an emerging area. Thus, to the extent that these concerns represent critiques of the actual body of current work (as opposed to worries of a more abstract nature), it does not imply that such characteristics are unchangeable. To the contrary, EDS/LA are undergoing rapid processes of development and maturation which LS can help inform. With this in mind, the paper then shows, through an overview of one particular research project, ways that such concerns can be meaningfully addressed through specific decisions about how the work is conducted. The paper concludes by elucidating a set of seven initial principles for "learning sciences aware" EDS/LA work and opening the question of to what extent these principles might be appropriate to guide EDS/LA work more broadly.

Concerns for the learning sciences in embracing learning analytics

Concern I: Reliance on algorithmic processing over human insight

A core concern for LS in the adoption of EDS/LA approaches relates to the relative balance of the roles of human and machine in the generation of knowledge; specifically that computational processing will supersede the role of human intellect (Wise & Schwarz, 2017). This concern can be elaborated at two levels. From a basic perspective, the application of EDS/LA relies greatly on the computer's ability to apply sophisticated algorithms to relatively large quantities of data; thus the bulk of the responsibility for knowledge generation falls to the computer. The counterargument to this points to the many important decisions made by humans about which overall class of computational methods to apply, which specific algorithm(s) to use, what features to include, and how to set the hyperparameters of the model. While advances in "deep learning" techniques can shift responsibility for some of these decision to computers (e.g. the multiple layers used in neural networks can build up features based on patterns that occur in the data rather than having them determined a priori), there are still numerous other judgments to be made. For example, there are important choices about the architecture of the network with respect to the size and number of layers included, the way the initial data inputs are represented, and even the appropriateness of adopting a neural network in the first place. It is important to be transparent about such decisions both to support algorithmic accountability (Buckingham Shum, 2016) and to acknowledge the role of human input in the modeling process. A key point here is that there are many different ways that humans and machines can "collaborate" in EDS/LA. For example, human judgement may be more efficiently used to verify or correct machine-generated coding than to manually apply codes to data from scratch (Cui, Jin, & Wise, 2017). A similar notion underlies work on open learner models that allows students to not only inspect the model but also edit or negotiate it with the system (Bull & Kay, 2016). It is also important to remember that computational outputs do not themselves represent a contribution to knowledge; human intellect is required to make sense of analytic results in light of the existing knowledge base. Thus, computational approaches to studying learning processes can be thought of as providing new kinds of support for generating human understanding rather than attempting to replace or automate it. Similarly, humans and computers offer different capabilities to provide support for learning based on EDS/LA; thus we should replace the false question of which support is better overall with exciting questions about how responsive scaffolding should be distributed across humans and computers over time. The second level of elaboration of the computational critique of EDS/LA acknowledges the role of humans in setting up analyses, but suggests that such high-level methodological decisions play a fundamentally different role in knowledge-generation than when the researchers themselves are the instruments through which meaning is extracted from data. This latter concern is connected to larger questions about valued knowledge claims and how they are best generated.

Concern II: Attention to generalized structures over contextualized processes

Another concern for LS in incorporating EDS/LA approaches emanates from existing tensions between quantitative and qualitative methodologies and their associated sets of epistemological assumptions. Stated (over)simply, quantitative approaches are often concerned with identifying generalizable regularities in learning while qualitative approaches tend to focus on understanding the particularities tied to specific contexts. While there is a place for well-conceived quantitative work as one of many sources of useful knowledge contributions within LS, the concern is that computational approaches exacerbate existing concerns about work that focuses on structural relations over nuanced processes, and on generalizability over contextualization. However, there are important differences to note between EDS/LA and traditional quantitative methodologies. First, as a field, EDS/LA is explicitly devoted to the study of learning processes. While these processes may be indexed and studied in a quantitative way, the importance of probing mechanisms for learning is built in to the core of the field. The unwillingness to accept black-box relations between treatments and outcomes as a satisfactory account of learning thus represents an area of commonality between LS and EDS/LA that is not always present in all quantitatively oriented work. Second, EDS/LA offers tools for pattern recognition that can be applied to look across large quantities of data collected from many different contexts. This offers the potential for insight into regularities with much greater actual support for generalizability than statistical inference based on data from just a handful of classrooms. Third, EDS/LA is also concerned with building models based on fine-grained patterns of data within individuals to develop personalized insights for students or subpopulations of students. This is quite a different enterprise than aiming to develop generalized knowledge, and one which presents better potential for alignment with LS, especially when such models are used to help students self-regulate their learning (for an elaborated critique of the problems with RCTs and how EDS/LA can help see Winne, 2017). Finally, computational approaches are often less pre-constrained in examining variables and the relationships between them compared with traditional statistical methods in which these must be hypothesized in advance.

Concern III: Emphasis on empirical findings over theory building

A final concern that has been raised about the use of EDS/LA approaches in LS is an outstated focus solely on documenting empirical patterns without sufficient attention to simultaneously developing and testing theories that can explain and cohere patterns seen in the data (Wise & Schwarz, 2017). This is a valid critique of some early work in the field; however other efforts, particularly those focused on developing intelligent agents, have worked closely with theory, developing conceptual models of learning and interactional processes which are then made computational (e.g. Zhao, Papangelis, & Cassell, 2014). Such use of computational models to instantiate, test, and refine theoretical models offers a powerful tool for the process of theory building. Currently, the importance of and need for increased attention to theory is well-acknowledged among many EDS/LA researchers (Wise & Shaffer, 2015), and it has also been noted that the introduction of new analytic methods can create new needs for theorization (for example, with regards to temporality, see Knight, Wise, & Chen, 2017). However, efforts are still required to develop the expertise and processes to do so. This actually represents an opportunity for LS to bring its long history of theorizing learning processes and multidisciplinary scholarly collaboration to bear productively in conversation with EDS/LA researchers.

Addressing the concerns: Examples from the MOOCeology project

The prior section laid out three of the key concerns for the learning sciences related to the adoption of EDS/LA approaches and demonstrated that they relate to characteristics that are neither inherent nor intractable. This section concretizes the argument by detailing the specific ways these concerns were addressed in the course of our work over the last three years on the MOOCeology project. This serves as an intermediary step towards generating a set of principles for conducting "learning sciences aware" EDS/LA work. The project originated from the premise that while *all* interactions in large-scale open online learning environments do not necessarily represent rich collaborative learning, that does not mean that *no* such learning takes place. Furthermore, we had a hypothesis that a lack of attention to differentiating among such interactions might explain some of the divergence found in the literature as to the roles and importance of discussion forums to MOOC learning.

Part 1: Seeking to differentiate "content-related" and "non-content" interactions

Research need: Identify learning-related interactions in large-scale discussion forums

Massive Open Online Courses (MOOCs) provide accessible learning opportunities, attract learners with diverse backgrounds and perspectives, and have the potential to support learning across the globe (Dillahunt, Wang, & Teasley, 2014). Discussion forums in MOOCs are important venues for interpersonal interactions, generally populated with numerous potential learning partners, but often flooded with a variety of topics not directly related to learning (Stump, DeBoer, Whittinghill, & Breslow, 2013). This creates a problem of practice (it is hard for forum participants to find learning opportunities and support) and a problem of research (analyses of very different kinds of interactions that play different roles in learning are considered together).

Prior to computation: Conceptualizing the role of MOOC discussions in learning

Discussion forums can play a variety of roles in the learning process, being designed as sites for collaborative knowledge building, community development, or a place to ask questions and seek answers. In the majority of MOOCs, discussion forums are offered as a supplemental form of support, with use for learning not prescribed. Based on this aspect of MOOC pedagogy, we chose to define "learning-related" very broadly as any discussions that relate to the course content. This included seeking/providing help, information, or resources on the course topic, whether specifically about content mentioned in the course materials or not. The remainder of discussions were considered "non-content" and tended to include technical / logistical issues or socializing (though they might serve other purposes). A second important decision was to categorize discussions at the level of the thread, rather than individual post. As MOOC discussions happen in threaded conversations, categorizing them at this level allowed us to understand interactions in the intact conversation context.

Computation: Building a supervised model using NLP to extend human content-analysis

Prior research has shown that even when MOOC discussions are designed to be segregated by topic, substantial cross-posting occurs, making this a poor indicator of discussion content (Rossi & Gnawali, 2014). A more accurate approach is to assess the content of discussion posts directly; however manual content analysis is extremely time-consuming given the quantity of posting in MOOCs. This creates an exciting research agenda and interesting questions around the extent to which computational models can be created to scale-up the work of human judgement, based on linguistic features of the posts. This study (Wise, Cui, Jin, & Vytasek, 2017) investigated these issues by first looking at whether starting posts of content-related threads in a statistics MOOC discussion (manually identified by humans with $\alpha >= .75$ reliability) had linguistic features that distinguished them from those starting non-content threads. (Starting posts set the frame for a discussion and thus are a reasonable first approximation of thread topic; the DIPTiC approach including reply posts was later developed to make a more refined characterization of thread topic [see Part 1 coda]). Once linguistic features distinguishing content and non-content starters were identified, we then asked if they could be used to create a model that reliably classified the content-related ones. Results showed that a binary L2 regularized logistic regression classification model based on unigrams and bigrams from the discussion posts showed good results on the original statistics course as evaluated by ten-fold cross-validation (accuracy = .80, kappa = .61, recall and precision both = .79). Generalizability was good to another offering of the same course and a different MOOC on the same topic (statistics), with negligible differences in accuracy. Exploring generalizability to two courses on progressively more distal topics (psychology and physiology) showed expected decreases in accuracy.

After computation: Unpacking the model to understand its relevance to different contexts

Following model creation and testing, we unpacked key linguistic features used to make the predictions to better understand the ways in which learners were starting content-related threads across the different courses. The top 30 features (ranked by kappa) from content and non-content starting posts in the five courses were extracted. Researchers then went back to the actual discussion text to examine how these features (words and pair of words) were used in context. This was important because the same word could have different meanings in different contexts of use that would not be detected by the model (e.g. "I have a *question* about transforming data" versus "Is the answer to *question* 2 choice B?"). Based on this examination, the features were organized into categories of word types. Top features of content-related starting posts were primarily terms related to the process of learning (e.g. understand), question words (e.g. why or what), and terms that connected ideas (e.g. but). In contrast, top features of non-content starting posts were terms related to the course tasks and platform (e.g. videos), effort / action (e.g. do), appreciation (e.g. great), and first person singular pronouns (e.g. my). Words related to the course domain (e.g. probability) were also present, but notably (and unexpectedly) these represented less than 20% of the top features across both classes. While somewhat counterintuitive (c.f. Rossi & Gnawali, 2014), given the diversity of specific course topics, this suggests that it is the language that surrounds the varied particular domain-specific words (e.g. interrogatives, learning process words, connectors) used again and again that are most important for identification. This finding can contribute to theories of learning through discussion by directing attention to how questions are asked as an important element leading to the discussion that results (i.e. a question on the same topic asked slightly differently might elicit vastly different replies, e.g. "Is the answer I should put 'scale the data'?" versus "How will scaling the data fix the distribution"). This focus on question form also helps make sense of the model's generalizability, not in terms of topic specific vocabulary, but in the discourse practices of a discipline and the pedagogical approach of the course (the physiology course differed from the others in connecting its topics to learners' daily lives, leading to more personal conversations).

Part 1 coda: Better differentiation via hybrid human-machine categorization

As a follow-up to the first study, we sought to examine whether we could combine the contributions of nuanced human insight and algorithmic processing power to increase the accuracy of categorization. In this work (Cui et al., 2017) we labeled discussion threads twice: once based on the content/non-content classification of their starting post via the model described in Part 1 and once based on a threshold proportion of content/non-content replies (application of the model to reply posts was first validated with accuracy = .85, kappa = .68). We could then automatically compare the two results, accepting the categorizations when they converged and using human judgment to resolve discrepancies. Compared to starter-only thread categorization, this method improved classification performance (estimated accuracy = .88 [vs .81 for starter-only]; estimated kappa = .76 [vs .62 for starter-only]) with the addition of 16 person-hours (using two humans to examine discrepancies). We refer to this process as Dynamic Interrelated Post and Thread Categorization (DIPTiC) and see it as one powerful way to position human and computer contributions to analysis in support of one and other, rather than in competition.

Part 2: Examining learning interactional processes in MOOCS

Research need: Deepen understanding of interaction in MOOC discussion forums

Understanding interactions in MOOC discussion forums can be challenging given the scale of activities and diversity of participation patterns. SNA can be useful for this purpose due to its strengths in identifying interaction patterns from complex activities (Scott & Carrington, 2011), but it alone may not be sufficient when understanding learning in discussion is the primary goal. As MOOC forum interactions consist of both content and non-content discussions, this raises the question of whether it is necessary to differentiate social relationships formed in different types of discussions. The literature indicates good reasons for doing so. First, academically-related and unrelated social interactions were found to impact college retention differently (Kuh, 2002). It is possible that social relationships develop in distinct patterns when the content and contexts of interactions differ. Second, MOOC learners engage with the courses in distinct patterns associated with different motivations (Kizilcec, Piech, & Schneider, 2013). It is possible that learners with interest in learning the course content and those who participate for social experiences take part in different types of discussions, and thus develop social relationships with distinct groups of people. These considerations substantiate the quest to understand social interactions and relationships in context, and the need to address this quest using SNA and content analysis methods in combination.

Prior to computation: What counts as "interaction" and "relationship"? How to operationalize them?

Constructing social networks involves critical conceptual and operational decisions. For instance, the stance on what activities (posting, reading, or both) reflect / associate with learning determines what tie definition is appropriate for network construction. Specifically, reply-based tie definitions (e.g. Direct Reply) construct networks for posting and are based strictly on the reply relationship between discussion contributors; copresence-based definitions (e.g. Total Copresence) construct networks for both posting and reading, and are based on coparticipation relationship among learners. Different tie definitions (even from the same category) can produce dramatically different social networks and substantially influence how the observed patterns should be interpreted (Wise, Cui, & Jin, 2017). These decisions need to be made by human researchers with careful consideration and in-depth understanding of many factors, such as the nature of learning being examined, the learning context, and the characteristics of the chosen tie definition.

Computation: What can we learn about social relationships through social properties at network, community, and individual levels?

Based on the characterization of content and non-content discussion threads, we can investigate the social relationships in these two types of interactions through examining the structural characteristics of the social networks. The study was conducted on discussion forums in a statistics MOOC. The forums were provided for optional interactions. Two instructors and 565 learners participated in the forums. The discussions were categorized using the methods introduced in Part 1 and social networks were constructed separately for the content and non-content discussions. As reading constitutes a substantial proportion of learning activities in online discussion (Wise, Speer, Marbouti, & Hsiao, 2013), social networks were constructed using the Limited Copresence definition, which constructs ties based on either thread coparticipation (for threads with a smaller number of replies) or subthread coparticipation (for bigger threads, see Wise, Cui, & Jin, 2017). Content and non-content social networks were found to have distinct characteristics at network, community (detected using the Louvain method), and individual levels. For instance, comparison of structural network properties (average node degree and average edge weight) showed learners interacted with more people and had more repeated interactions with the same people in content discussions than in non-content discussions. Moreover, the two networks were participated by substantially distinct people, with only 28% of all forum participants contributed to both kinds of threads; for learners who contributed to both content and non-content threads, those who were highly connected in one network were not necessarily highly connected in the other. Furthermore, examination of the major communities (containing > 5% of the network populations) from the two networks yielded two additional findings that expanded our understanding of forum interactions. First, a learner-only community in the content network had a web structure with a distributed core consisting of multiple central learners connected via strong ties, which was dramatically distinct from the wheel or elongated structures in other communities; the community members interacted with substantially more people and had more repeated interactions with the same people. Second, in the content network, learners in the community around one instructor showed stronger ties with a greater number of peers than those around the other instructor.

After computation: Probe where the computation flags to get in-depth understanding of interaction

Comparing structural properties of content and non-content social networks led to improved understanding of interaction and relationship patterns in the two types of discussions. These results also flagged areas that worth to investigate in ways that humans can do better than machines. We conducted inductive analysis on threads contributing to major communities in the two networks following the constant comparative method (Auerbach & Silverstein, 2003), to identify emergent themes and patterns through probing the characteristics, similarities and differences between interactions in the communities. It was found that interactions in the content and non-content networks involved different communication purposes. Non-content interactions often involved straightforward factual information exchanges and did not evolve into extended conversation; content interactions commonly involved problem-solving or understanding complicated concepts, which required multiple rounds of back-and-forth comments to resolve. In the content interactions, it was common for participants to use diverse interaction techniques such as paraphrasing, giving examples, and asking leading questions and follow-up questions. Moreover, conversation structures in content and non-content communities showed differences. Content conversations often developed into complicated structures, such as multiple subtopics and new topics extended from the original one. In contrast, non-content conversations usually had relatively simple and linear structures. These qualitative findings in return provided insights for the differences in structural properties. For instance, it is possible that learners developed bonds with the same peers through multiple rounds of exchanges in the same content conversations, which encouraged more subsequent interactions and resulted in higher edge weight.

The qualitative examination of threads contributing to the learner-only community and the two instructor communities in the content network also yielded useful findings. The learner-only community showed unique nascent community-like characteristics. For instance, some members in this community called on peers to have group discussions and expressed the desire for study partners; they used many social presence indicators when interacting with each other; they valued the collaborative discussions and felt they learned from them; some members who received help often revisited the conversation to help answer others' questions. Examination of the two instructors' contributions revealed contrasting facilitation patterns and styles. One instructors not only revisited the threads that he/she had participated in, but also commented on other learners' replies to learner-initiated threads. This instructor often tried to encourage and help learners to work out the answer or solution themselves, using hints and leading questions. He / she also used a variety of social presence indicators such as greetings and addressing learners by name in his/her messages. In contrast, the other instructor only replied to the thread starting posts. He / she tended to provide straightforward answers or instructions to address learners' questions and used social presence cues infrequently. These qualitative findings may help explain the social network properties (such as node degree and edge weight) for these communities.

Part 3: Untangling the relationship between interaction and learning in MOOCs

Research need: Unpack the relationship between learning and forum interaction in MOOCs

Understanding the connections between discussion forum engagement and learning outcomes can have multifold implications. Theoretically, it can provide a foundation to better investigate and articulate the mechanism(s) by which interaction in forum discussion contributes to and/or reflects learning. Practically, such understanding can inform the facilitation of forum activities to maximize the intended type of learning. It might additionally provide grounds for the inclusion of MOOC forum activity as an integral (rather than supplemental) element of pedagogical design.

Prior to computation: Conceptualizing and operationalizing learning and interaction

When investigating the relationship between learning and interaction in forum discussions, we need to first define the two. Learning is maybe most straightforwardly and frequently measured by course performance, although the specific performance measures used can vary course by course (e.g. pass / fail, normal / extinction, grades). There are two issues that credit special attention. First, in addition to reflecting difference in the kind of learning outcomes being measured, performance measures may associate with some latent factors (e.g. orientations, commitment) that impact the outcomes and interaction, and should be taken into consideration when interpreting the observed relationships between the two. For instance, a learner who participated actively in the discussions may fail to pass a MOOC due to the lack of interest in obtaining a certificate, rather than having learning difficulties. Second, before entering performance variables into computation models, researchers may want to know how the performance assessment was operationalized in the specific learning context, such as whether multiple submissions for quizzes were allowed, whether more weight was given to formative assignments or summative exams, or whether the assessment was conducted through "quantitative" automated grading or more "qualitative" approaches, such as peer reviewed projects or writing assignments. These variances can have important implications for interpreting the observed relationships between learning and interaction. In this regard, CSCL theories on learning assessment can provide valuable insights. Compared to learning outcomes, defining interaction in MOOC discussions can be even more complicated. Researchers need to make two primary decisions: what interaction to measure and how to measure it. First, what to measure is a non-trivial decision that reflects fundamental assumptions about what contributes to / reflects learning. For instance, variables can be constructed for activities such as posting, editing, reading, voting, and following threads. Second, how a certain form of interaction is measured is a critical decision. For instance, forum contributions can be measured for quantity (Gillani & Eynon, 2014), acceptance by peers (Coetzee, Fox, Hearst, & Hartmann, 2014), cognitive engagement characteristics (Wang, Yang, Wen, Koedinger, & Rosé, 2015), and relatedness to course content (as we did in this project). Thoughtful decisions on these fronts can help yield findings that both explain what contributes to learning and are actionable for supporting learning and teaching.

Computation: Does content/non-content discussion involvement relate differently to performance?

In the MOOCeology project, we approached this topic by examining the connections between forum interaction (measured by quantity of forum contributions and several social centrality properties, both differentiated based on content relatedness) and course performance (measured by final grade and pass /fail) [Wise & Cui, 2018]. The purpose was two-fold. One was to investigate whether or not the quantity of content and non-content interaction differs in explanatory power for course performance; the other was to investigate whether or not social centrality properties add to the explanatory power on top of interaction quantity. The study was conducted on the same statistics MOOC examined in Part 2. It was found that for relatively committed learners (who gained more than 1% for final grade), those who contributed to the discussion forum had a significantly higher rate of successfully passing the course than non-contributors (64% vs 32% passing); learners who made posts to both types of threads had a higher passing rate than those who only contributed to content or non-content threads (77% vs 60% / 57% passing). Among learners who successfully passed the course, there were no differences in course grade when comparing discussion contributors and non-contributors overall; however those who contributed to content-related threads performed slightly better than those who did not (course grade of 87% vs 85%). A regression model based on the number of posts made to content-related threads explained 3% of variance in course grades. Addition of other interaction quantity measures (including number of threads contributed to, total number of posts and non-content posts) did not add to the model's explanatory power significantly; neither did addition of any social centrality measures (including degree, weighted degree, closeness, betweenness, and eigencentrality in the content and non-content networks).

After computation: A need to reconsider the definition of "learning" in MOOC discussions

It is notable that this research did not document a strong relationship between MOOC discussion interaction and course performance. This can be explained in several ways. It is possible that the discussion forums were not helpful for improving course performance (as they were not pedagogically integrated into the course but a supplementary venue for optional participation). It is also possible that the discussions were helpful for course performance, but the variables we used were not well tuned to capture the relationship. This is a plausible explanation because we only measured the quantity of learner's forum contributions differentiated based on content-relatedness (in contrast to more nuanced differences, such as quality and questions / answers); and we did not take reading into consideration. However, on top of these two explanations, a third possibility that looks beyond what was in the model should be noted: certain type of learning did happen, but course performance is not the right proxy for it. For instance, Nelimarkka and Vihavainen (2015) found that some alumni learners participate in the same MOOCs repeatedly; instead of pursuing a high grade, their interest may involve networking and assisting others. Their developing expertise and identity in the domain can be considered as a form of learning, but not one captured by course performance. This interpretation of the lack of strong relationship between forum interaction and course performance presents a need to reconsider the definition of learning in MOOC discussions as well as to differentiate short term performance and long term learning.

Conclusions

The description above illustrates the ways that our work on the MOOCeology project has taken seriously the three concerns outlined at the start of this paper. We addressed the relative balance of responsibility accorded to computational processes and human insight by leveraging both in complementary ways in the DIPTiC method and conducting manual follow-up analysis on computational results. We attended to the process of learning in the context in which it occurs by going back to the data to understand how the top linguistic model features were used by learners and following up on communities identified by SNA methods to probe their interactional processes. Finally, we mobilized theory to both frame and be informed by the empirical analyses performed as we considered the meaning of different tie definitions and recognized the need to reconceptualize learning in MOOCs. Our work is not unique in attending to learning sciences concerns in the context of an EDS/LA project, but we believe by making our conscious consideration of them explicit, we offer an important contribution to a conversation that needs to happen at the intersections of these fields. In conclusion, we offer an initial set of principles for conducting "learning sciences aware" EDS/LA work. We see these as a starting point for dialogue about directions of EDS/LA work in LS, and perhaps EDS/LA work more broadly.

Initial principles for a learning sciences aware learning analytics

1. **Ground Analysis in Theory**: A theory of learning in the area being studied (or an explanation of why no existing theory is applicable) should be used to ground the framing of the study.
2. **Characterize the Context Richly**: Details of the learning context(s) from which the data was collected should be characterized in detail (e.g. in terms of pedagogy, technology, populations etc.).
3. **Justify Choice of Data and/or Features**: Decisions about what data, variables or features to collect, construct or include in an analysis should be clearly justified, ideally with reference to learning theory.
4. **Make Sense of High-Level Patterns using Low-Level Data**: Claims made based on computational analyses should be additionally supported by analysis methods that go back to the data to examine if / how the detailed traces bear out the interpretation of the higher-level patterns.

5. **Present Analytics Results Connected to Learning Processes**: Representative examples from the underlying data should be presented to help draw connections between the learning events as they occurred and their computational representations.
6. **Appraise Scope / Boundaries of Applicability**: The extent to which the models built (or results obtained) are thought to be specific to the context(s) studied or in what kinds of similar learning environments they might apply should be acknowledged.
7. **Consider Theoretical Implications**: Implications of the results for confirming, challenging, or refining existing theories of learning or new avenues for theorization should be addressed.

References

Auerbach, C., & Silverstein, L. B. (2003). *Qualitative data: An introduction to coding and analysis*. NYU press.

Buckingham Shum, S. (2016, March 25). *Algorithmic accountability for learning analytics*. Retrieved from http://simon.buckinghamshum.net/2016/03/algorithmic-accountability-forlearning-analytics

Bull, S., & Kay, J. (2016). SMILI: A framework for interfaces to learning data in open learner models, learning analytics, related fields. *International Journal of Artificial Intelligence in Education, 26*(1), 293-331.

Coetzee, D., Fox, A., Hearst, M. A., & Hartmann, B. (2014). Should your MOOC forum use a reputation system?. *Proceedings of CSCW 2014* (pp. 1176-1187). ACM.

Cui, Y., Jin, W., & Wise, A. (2017). Humans and machines together: Improving characterization of large scale online discussions through Dynamic Interrelated Post and Thread Categorization (DIPTiC). *Proceedings of Learning@ Scale 2017* (pp. 217-219). ACM.

Dillahunt, T. R., Wang, B. Z., & Teasley, S. (2014). Democratizing higher education: Exploring MOOC use among those who cannot afford a formal education. *IRRODL*, *15*(5). doi: 10.19173/irrodl.v15i5.1841

Gillani, N., & Eynon, R. (2014). Communication patterns in massively open online courses. *Internet and Higher Education, 23*, 18-26.

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. *Proceedings of LAK'13* (pp. 170-179). ACM.

Knight, S., Wise, A., & Chen, B. (2017). Time for change: Why learning analytics needs temporal analysis. *Journal of Learning Analytics, 4*(3), 7-17.

Kuh, G. (2002). From promise to progress: How colleges and universities are using student engagement results to improve collegiate quality. *NSSE Annual Report*. Bloomington, IN: Indiana University.

Nelimarkka, M., & Vihavainen, A. (2015). Alumni & tenured participants in MOOCs: Analysis of two years of MOOC discussion channel activity. *Proceedings of Learning@ Scale 2015* (pp. 85-93). ACM.

Rossi, L.A, & Gnawali, O. (2014). Language independent analysis and classification of discussion threads in Coursera MOOC forums. *Proceedings of IEEE IRI 2014* (pp. 654-661). IEEE.

Scott, J., & Carrington, P. J. (2011). *The SAGE handbook of social network analysis*. London: SAGE.

Stump, G. S., DeBoer, J., Whittinghill, J., & Breslow, L. (2013). Development of a framework to classify MOOC discussion forum posts: Methodology and challenges. *Proceedings of NIPS 2013 Workshop on Data Driven Education* (pp. 1-20). NIPS Foundation.

Wang, X., Yang, D., Wen, M., Koedinger, K., & Rosé, C. P. (2015). Investigating how student's cognitive behavior in MOOC discussion forums affect learning gains. *Proceedings of EDM 2015* (pp. 226-233).

Winne, P. H. (2017). Leveraging big data to help each learner upgrade learning and accelerate learning science. *Teachers College Record, 119*(3), 1-24.

Wise, A., & Cui, Y. (2018). Unpacking the relationship between discussion forum participation and learning in MOOCs: Content is key. *Proceedings of LAK'18* (pp. 330-339). ACM.

Wise, A., Cui, Y., & Jin, W. (2017). Honing in on social learning networks in MOOC forums: Examining critical network definition decisions. *Proceedings of LAK'17* (pp. 383-392). ACM.

Wise, A., Cui, Y., Jin, W., & Vytasek, J. (2017). Mining for gold: Identifying content-related MOOC discussion threads across domains through linguistic modeling. *Internet and Higher Education*, *32*, 11-28.

Wise, A., & Schwarz, B. (2017). Visions of CSCL: Eight provocations for the future of the field. *International Journal of Computer-Supported Collaborative Learning 12*(4), 1-45.

Wise, A., & Shaffer, D. (2015). Why theory matters more than ever in the age of big data. *Journal of Learning Analytics*, *2*(2), 5-13.

Wise, A., Speer, J., Marbouti, F., & Hsiao, Y. (2013). Broadening the notion of participation in online discussions: Examining patterns in learners' online listening behaviors. *Instructional Science*, 1-21.

Zhao, R., Papangelis, A., & Cassell, J. (2014). Towards a dyadic computational model of rapport management for human-virtual agent interaction. *International Conference Intelligent Virtual Agents* (pp. 514-527).