PLAY Minecraft! Assessing secondary engineering education using game challenges within a participatory learning environment

Erin Shaw, University of Southern California

Erin Shaw is a Computer Scientist at the University of Southern California's Information Sciences Institute, a research center at the USC Viterbi School of Engineering. Her research focuses on modeling and assessing student knowledge in the areas of science and mathematics, experimenting with new technologies for aiding assessment in distance learning, and studying computer mediated social dialogue and team collaboration in post-secondary engineering education. She received an MA in Online and Distance Education from The Open University, an MS in Computer Graphics from Cornell University and a BS in Mathematics from Massachusetts State University, Fitchburg. Ms. Shaw has directed research as a co-principal investigator on several National Science Foundation sponsored grants. In 2013, she served as a STEM outreach specialist at the USC Viterbi School of Engineering.

Mr. Minh Tuan La, University of Southern California

Minh La is an undergraduate student at University of Southern California majoring in Computer Science. He was a junior when he wrote this paper, and is expected to graduate in May 2015. He has been interested in the field of Computer Science since high school when he attended Center for Advanced Technologies in Florida. His passion leads him to constantly ponder on how evolving technologies can be deployed to find it's applicable usage. After completing his studies in USC, Minh pursues a career in Software Engineering.

Richard Phillips

Richard Phillips, University of Southern California Richard Phillips is an undergraduate student at the University of Southern California majoring in Computer Science and Business Administration. He was a sophomore when he wrote this paper, and is expected to graduate in 2016. He plans on getting his Masters in Computer Science as part of USC Viterbi Engineering School's progressive degree program in 2017. After graduating, Richard wants to pursue a career in the field of software engineering and eventually management.

Erin B. Reilly, University of Southern California Annenberg Innovation Lab

Erin Reilly is Creative Director & Research Fellow for Annenberg Innovation Lab at USC's Annenberg School for Communications & Journalism. In her role, she oversees all aspects of lab programming, product design and mentoring students in developing applications and business ideas using digital media and how it impacts society. Her research focus is children, youth and media and the interdisciplinary, creative learning experiences that occur through social and cultural participation with emergent technologies.

Erin is currently developing PLAY!, an educational collaboration platform helping learners tap into broad interest based peer communities as well as exploring new forms of reading and writing through dynamic book prototypes. She most recently published her first digital book, Flows of Reading, to inspire educators to reflect on what can be considered as reading and what kinds of reading they perform in their everyday lives.

She was Research Director for Project New Media Literacies at MIT and also has conducted classes as a Visiting Lecturer at MIT's Comparative Media Studies Department and Harvard University's Project Zero Summer Institute. Reilly is a graduate of Emerson College and has her Master of Fine Arts degree from Maine Media Workshops + College. She is a member of the Academy of Television Arts & Sciences, the first-Vice President board member of NAMLE (National Association for Media Literacy Educators) and serves on advisory boards, such as PBS Emmy-award winning Sci Girls. Erin consults with private and public companies in the areas of mobile, creative strategy and transmedia projects.

PLAY Minecraft! Assessing secondary engineering education using game challenges within a participatory learning environment (Research to Practice)

Strand: Addressing the NGSS: Supporting K-12 Teachers in Engineering Pedagogy and Engineering-Science Connections

Abstract

This paper describes an initial step towards understanding how computational tools such as natural language processing and machine learning might be used to assess K-12 student learning in engineering education. The study used an online participatory learning environment, *PLAY! (Participatory Learning and YOU!),* as a platform for student work. Minecraft, an online construction game popular with young teens, was chosen as the learning topic to be assessed. Within *PLAY*, students created and shared Minecraft 'challenges' during a focus session consisting of five boys, ages 9 to 16. Machine learning techniques were used to create a classification scheme for engineering standards based on the *Science and Engineering Practices* in the Next Generation Science Standards. Natural language processing and data mining techniques were applied to student challenges to assess and report on students' engineering domain and topic learning. Results show that student application of engineering standards and student discussion of domain topics varied consistently by age. Responses to a corresponding questionnaire showed that the session was a highly positive experience for the children. The potential for use in engineering education is discussed.

Participatory Learning

Participatory cultures are defined as cultures with relatively low barriers to artistic expression and civic engagement, strong support for creating and sharing one's creations, and some type of informal mentorship[1]. Participatory cultures foster new media literacies that build on traditional literacy skills taught in the classroom but focus on social skills developed through collaboration and networking, such as play, performance, simulation, appropriation, distributed cognition, and judgment.

As schools shift from traditional systems of teaching, student engagement, and professional development, the norms for education are changing. How do individuals best learn and communicate through visual media? How is teaching quality improved by pulling in diverse resources and perspectives that are typically overlooked? How do you build enthusiasm and alignment behind these new models of teaching? The *PLAY!* (*Participatory Learning and You!)* platform is a social environment helping students and teachers tap into broad interest-based peer communities – encouraging the four C's of participation in the learning process, creativity, connection, collaboration and circulation[2]. It is a visual platform that enables people to collaborate and problem-solve through creating and sharing media-rich content. The basic unit of collaboration on *PLAY* is called a *canvas;* however, the term *challenge* was used before adopting the more neutral term *canvas* and is the term we used with participating users in the study (hence the title of the paper).

This multimedia canvas encourages users to engage, discuss and share their questions, ideas and knowledge with others. To build a canvas, users simply drag and drop media elements onto the canvas. They can include a wide range of media, from text to photos to videos to RSS streams and more. Canvases can be published publically, or shared privately with a selected group of people. Viewers can take action on a canvas through the *Your Turn* feature, which enables them to respond to the canvas via a photo or a video. Users can also drive discussions with specific questions that they pose in *What Do You Think*, the targeted commenting feature. And if a viewer feels inspired, they can remix the learning canvas and build upon the idea to further expand on the creative exchange.

PLAY is based on five core principles for participatory learning that have been identified over two years of working with elementary and secondary teachers from the Los Angeles Unified School District who were seeking to develop a more participatory environment in their classroom[2].

- Participants have many chances to exercise *creativity* through diverse media, tools, and practices;
- Participants adopt an ethos of *co-learning,* respecting each person's skills and knowledge;
- Participants experience heightened *motivation* and *engagement* through meaningful play;
- Activities feel *relevant* to the learners' identities and interests;
- An integrated learning system – or *learning ecosystem* – honors rich connections between home, school, community and world.

The affordances of Web 2.0, associated with social software such as blogs, folksonomies, and peer-to-peer media sharing, and the ubiquity of networked computers in K-12 schools have made possible new educational practices like *PLAY,* that have the potential to produce "radical and transformational shifts" in learning[3]. However, for new educational practices to become accepted they must fit into the curriculum, be aligned to state standards, and have appropriate assessments[4]. This research described in this paper addresses the latter two challenges, NGSS standards alignment and learning assessment.

Assessment Pipeline for STEM Learning

The goal of the project was to develop a standards-based assessment pipeline for *PLAY* and to test the feasibility of using it as a STEM learning environment. The work was an inter-disciplinary collaboration between researchers at the University of Southern California's Annenberg School of Communication and Journalism and USC's Information Sciences Institute. Engineering was chosen as the STEM domain, and the game of Minecraft (2013) was chosen as the learning topic. Work included development of

- A data pipeline for analyzing the text and context of a *PLAY* canvas;
- A machine learning classification system for identifying the application of engineering standards and domain topics;
- Visual analytics for instructional assessment.

The work was conducted over ten weeks during the summer of 2013, and was designed as a collaborative research experience for two undergraduate engineering students, one entering his sophomore year, and one entering his junior year, both of whom are co-authors of this paper. Breadth of research was emphasized over depth of research, as was appropriate given the students' level of expertise, and which allowed for a wider range of research experiences, including using different computational and analytic tools, and conducting a small focus session to collect data that could be analyzed.

Next Generation Science Standards

The 2013 Next Generation Science Standards (NGSS)[5] are arranged by disciplinary core ideas (e.g. heredity and energy) for each subject (e.g., life science and physical sciences), for each school level (e.g., middle and high school). NGSS sets performance expectations for each disciplinary core idea based on age level appropriate *Science and Engineering Practices*, *Disciplinary Core Ideas* (previous building blocks), and *Crosscutting Concepts* (Figure 1, left).

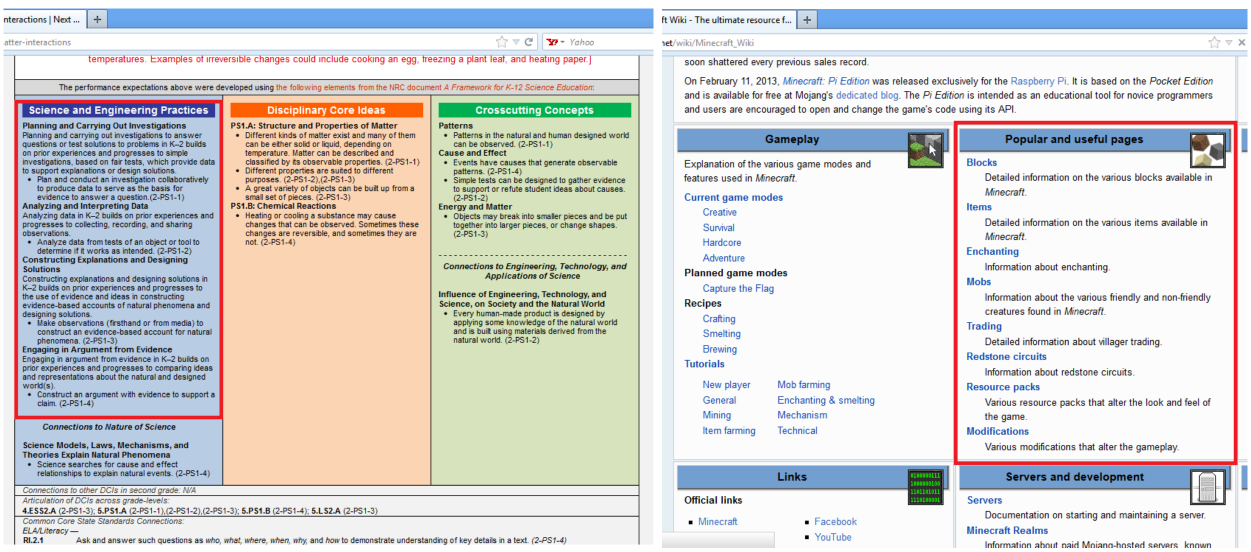

Figure 1. On left, NGSS Science and Engineering Practices associated with a core disciplinary idea. On right, the top-level page of the Minecraft Wiki, showing its main topics.

Minecraft

Minecraft is a first-person free to play indie PC/Mac game with crafting, building and exploration at its center[6,7]. Its dynamic virtual worlds consist of LEGO-like blocks that represent natural resources. Players can combine blocks to "craft" items and to build fantastical structures. There are both single player and networked multiplayer versions of the game. Minecraft relates to engineering practices because it inspires players to think creatively to solve in-game problems using mathematical and spatial analysis[8]. Because Minecraft as a disciplinary topic might span many grades (for example, the focus session participants spanned grades 4 through 11), we used the Science and Engineering Practices (SEP) from all grades to build a set of categories for assessment. The final SEP categories were:

- Analyzing and interpreting data

- Asking questions and defining problems
- Constructing explanations and designing solutions
- Developing and using models
- Engaging in argument from evidence
- Obtaining, evaluating, and communicating information
- Planning and carrying out investigations
- Using mathematics and computational thinking

Text Processing

For the purposes of this experiment, Minecraft was treated as a subject, and information about Minecraft was retrieved from the website www.minecraftwiki.net (Figure 1, right). Content corresponding to one of seven categories was fetched using the Python XML and HTML parsing library, lxml.html[9] and processed using NLTK (Natural Language Toolkit)[10,11]. Processing included filtering out irrelevant links and removing html tags, stop words, and irrelevant characters. This resulted in several pages of text each for the final seven categories: Blocks, Items, Enchanting, Mobs, Trading, Redstone Circuits, and Modifications.

Topic Modeling

Mallet (Machine Learning for Language Toolkit*)*[12] was used to perform topic modeling using LDA (Latent Dirichlet Allocation)[13] for both SEP and Minecraft categories. LDA produces a set of topics, each of which consists of a set of topic words based on their co-occurrences in the document corpus. Topic modeling works best with a large corpus so the lack of text was a concern for the SEP categories; in practice it resulted in fewer topics and concise, but fairly well defined, categories (consisting 50%-75% of only one topic). Additional topic keys (set of words) were provided manually for the Minecraft corpus, to create stronger categories. The SEP and Minecraft categories were pre-processed and these results were measured against text written by *PLAY* users. The final processing can be done in near real time, although in a school setting, real time analytics are not necessary. Instead, teachers might check results after each classroom or homework-based *PLAY* session.

Focus Session Methodology

Assessment of the system included a two-hour *PLAY* session and a corresponding questionnaire in July 2013. The goal was to collect data for exploratory analyses. Five boys participated: two entering grade 4, two entering grade 8 and one entering grade 11. Participants were given two seed challenges to start, “Building a House in Minecraft” and “Building a Sewer in Minecraft”, which were created by two of the authors, ages 19 and 20, both undergraduate engineering students, to allow students to become accustomed to the *PLAY* interface before creating their own challenges. The seed challenges were purposefully created to provide data that would correlate to engineering standards. After thirty minutes of commenting on and debating the seed challenges, students were instructed to create their own challenges, which they did for the next thirty minutes. The final hour was spent participating in peer challenges and also iterating on their own challenges (responding to peer comments). Canvases created by the students are shown in Figure 2., by student grade. The 11th grade student created two canvases.

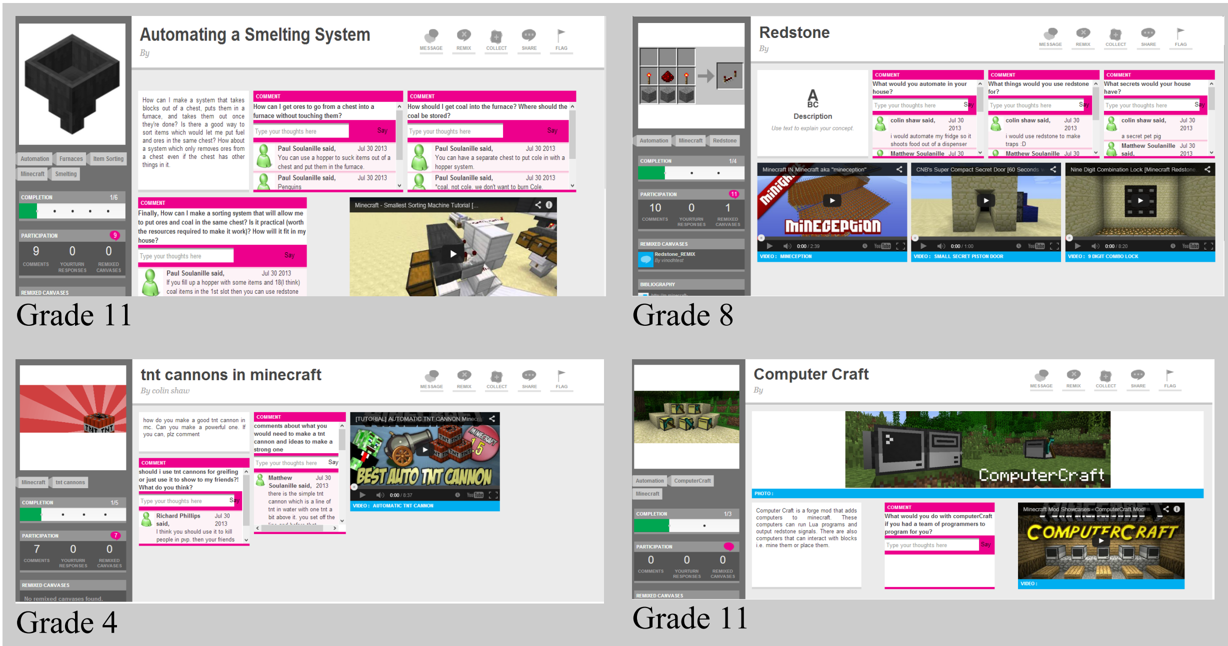

Figure 2. *PLAY* canvases created by students.

Participants responded to a questionnaire after the *PLAY* session, providing feedback on Using PLAY, Responding to Challenges, and Creating Challenges, shown in Table 1. Students had fun using the platform, even found it entertaining, and especially liked sharing ideas, although, as the older student commented, there was not so much building off of other people's ideas. In general, students liked sharing their own ideas and commenting on others, but did not extend the ideas of others, which could be explained by an absence of explicit instructions and also by the two-hour time constraint. The software is in development and the older student noticed and commented appropriately about some issues. All students responded that they played Minecraft often or all the time when they had free time, and three of the students responded that they have played Minecraft for more than two years.

Table 1: Responses to questionnaire.

Using *PLAY*	
Was using PLAY! fun? a. *Why?*	"GUI oriented", "you got to respond", "fun to give people challenges"
b. *Why not?*	"it seemed a bit buggy"
Responding to challenges	
a. *What did you like about responding to challenges?*	"other ppls ideas", "you get to share your thoughts", "it was slightly entertaining", "you could share what you thought about the challenge, your ideas"
b. *What didn't you like about responding to challenges?*	"not much", "liked everything", "it wasn't not, fun", "there didn't seem to be much collaboration between people of building off others' ideas", "hard to choose stuff",
Creating challenges	

a. *What did you like about creating challenges?*	"easy and fun", "you learned how to make a challenge", "fun to think about and test out", "you can share what you like about something." "You can ask other people for help on projects"
b. *What didn't you like about creating challenges?*	"not much", "liked everything", "nuttin'", "mostly the software bugs"
Minecraft experience	

a. *When you have free time, how often do you spend it playing Minecraft?*
Never (0), Sometimes (0), Often (4)(G4,8,11), All the time (1)(G8)

b. *How many years would you estimate that you have been playing Minecraft?*
Less than 1 year (1)(G4), 1-2 years (1)(G4), More than 2 years (3)(G8,8,11)

Data Processing Methodology

The text written by each student during the two-hour focus session was collected. For collaborative responses, the original question or prompt that the student responded to was also included in the student's data set. MALLET LDA was used to analyze student data to determine which topics, determined from the topic modeling of the SEP and Minecraft corpora, the student text most closely matched. For example, 64% of one student's text was consistent with topic 8 and 34% was consistent with topic 4. Since 70% of the text used to describe the SEP category *Analyzing and Interpreting Data* was consistent with topic 8, the student's text most closely matched this category. Topic 4 was common to all SEP categories, so we reported only the 64% match for *Analyzing and Interpreting Data*.

Data from each user were analyzed and plotted. Results for the SEP categories are shown in Figure 3 and results for the Minecraft categories are shown in Figure 4. The user legend is the same for both figures and is based on student grade, "G"; so that user G4 was in grade 4 and users G8a and G8b were both in 8th grade, etc. Users G14 and G15 were undergraduate students entering their sophomore and junior years, respectively, who provided the seed challenges; they did not participate collaboratively with the younger students. The vertical axis represents the percent of student text matching the category of interest. Because we reported only significant topic matches, e.g., the 64% in the example above, most matches fell between 0 and 60 percent.

Results

In looking at the students' application of SEPs (Figure 3), we found that the youngest student, G4, applied only one SEP standard, while the oldest student, G11, applied 7 of 8 standards, including two that were not applied by any of the other three K-12 students, *Planning/Carrying Out* and *Obtaining/Evaluating*. The two G8 students applied 3 and 5 standards, respectively. We know that student G8b is a mathematically gifted and we found evidence of this in his coverage of the *Mathematics/Computation* standard. In going back to the raw data, we found that the older students used more sophisticated words and engineering specific words while the younger users commented less and used more ambiguous terms with respect to engineering.

We also see that many of the students focused on the standards *Asking Questions/Defining Problems*, *Constructing Explanations*, and *Analyzing/Interpreting Data*. Per the raw data, this makes sense because in the challenges they created, they asked others how to build an object,

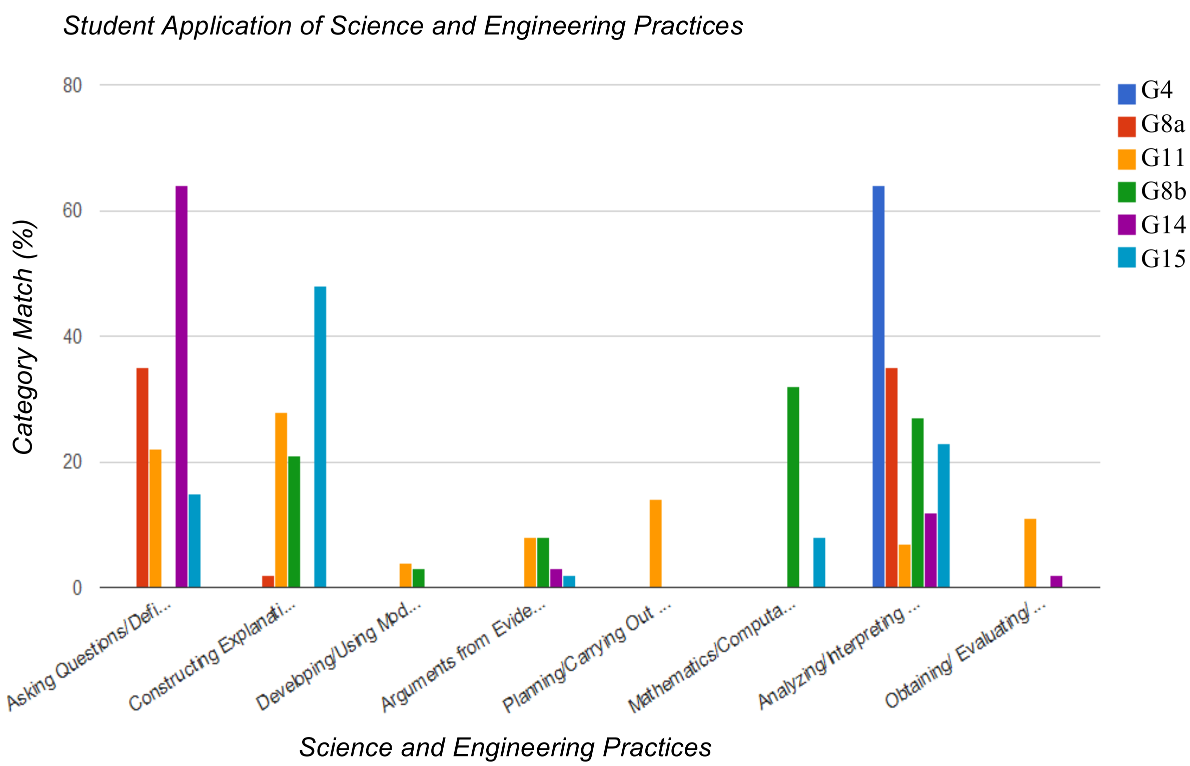

Figure 3. Results showing application of SEP categories for each student.

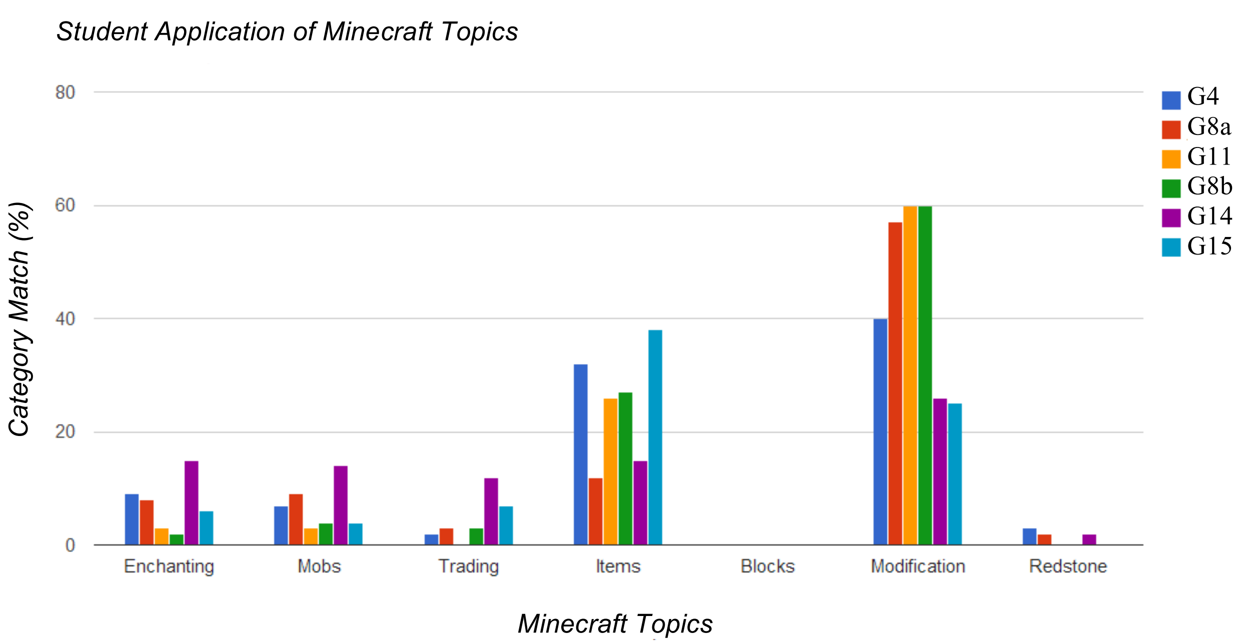

Figure 4. Results showing application of Minecraft categories for each student.

which correlates with *Asking Questions/Defining Problems*. Then they gave opinions on how to build the objects, which correlates with *Constructing Explanations*. Lastly the users debated whether different options would be a viable way to build the object, which correlates with *Analyzing/Interpreting Data*.

The Minecraft topic coverage (Figure 4) also makes sense. We see that students focused mainly on the topics "Modifications" and "Items" which makes sense because the two seed challenges focused on how to build an object and the students usually responded with a "*you could build this out of ...*"

Discussion

Several surprising results came out of this exploratory research. Regarding *PLAY*, which had been piloted only by graduate students and teachers for professional development, we found that children as young as nine could use it easily, and enjoyed sharing their creations and ideas. Regarding Minecraft, we found, again, that children as young as nine understood how to create objects, and owned the objects they created.

The Mallet LDA results were slightly misleading because topic percentages must add up to 100% as opposed to being true coverage percentages for each category. The ratio of coverage across all categories, however, was accurate. Another problem with the method is that keywords may be inaccurately categorized because of the absence of context. For example, Mallet may categorize the word "redstone" in *Mods* when in context the more appropriate topic is *Redstone*. Similarly, the simple language of younger users and their use of generic words like make, bad, good, etc. was also a challenge for analysis. The data appeared to be reliable, however, when we compared the engineering standards with student text. The next step would be to acquire more data and do a deeper analysis.

Because this was a summer research project, the biggest constraint was time. Mallet LDA has several parameters that affect its performance, some of which we experimented with (e.g. number of iterations and topics) and others that we relied on default values for (e.g. optimization bounds and interval). There are other versions of LDA (e.g., TMT, Online LDA, Parallel LDA and we could also have tried a different method or an alternative method (e.g., Compound Topic Model, PCA or Perceptron). Developing metrics for analyzing the validity and reliability of automated assessment tools is an important area of research regardless of the methods chosen. We are currently designing a study to help us evaluate our LDA results that require human annotators to assign SEP categories to student text. If high annotator agreement occurs, the resulting *ground truth*[14] can be used to improve the accuracy of our results. In addition, there is a pedagogical structure to *PLAY* prompts that we did not explore, i.e. the *Your Turn* and *What do you think?* features, but that we think could be incorporated in future studies.

Conclusion

We have described an initial step towards understanding how computational tools such as natural language processing and machine learning might be used to assess K-12 student learning in engineering education by applying new science and engineering standards to student activities within a participatory learning environment. With the advent of the development of educational mods and curricula for Minecraft, such as Quantum Physics for Minecraft (qcraft.org), the game has enormous potential for engineering education, and *PLAY* provides a pedagogically principled environment for participatory learning and assessment.

Regarding the use of LDA to infer the application of *Science and Engineering Practices* in text-based exchanges, we were surprised, especially given the small amount of text and low number of study participants, to see that SEP coverage aligned so well with student age levels and actual student data, although objective methods will be necessary for validating more general results, and results for other disciplines. Taken as a whole, the combination of computational assessments based on NGSS and the *PLAY* participatory learning environment has the potential to positively impact future science learning.

Acknowledgements

This research was funded through an Undergraduate Research Associates Program grant from the University of Southern California. The authors thank *PLAY* administrator Aninoy Mahapatra for providing the database exports and guidelines. We also thank graduate student Sam Shuster for mentoring the undergraduate students on MALLET.

Bibliography

1. Jenkins, H., Clinton K., Purushotma, R., Robinson, A. J., and Weigel, M. (2006). Confronting the challenges of participatory culture: Media education for the 21st century. Chicago, IL: The MacArthur Foundation. Available at http://www.macfound.org/media/article_pdfs/JENKINS_WHITE_PAPER.PDF.
2. Reilly, E., Jenkins, H., Felt, L.J., and Vartabedian, V. (2012). *Shall we PLAY?*, http://www.annenberglab.com/sites/default/files/uploads/Shall_We_PLAY_final_small.pdf.
3. McLoughlin, C., and Lee, M. J. (2007, December). Social software and participatory learning: Pedagogical choices with technology affordances in the Web 2.0 era. In *ICT: Providing choices for learners and learning. Proceedings ASCILTE Singapore 2007* (pp. 664-675).
4. Nobori, M. (2013). A Step-by-Step Guide to Best Projects: Discover a project-based learning model that motivates students to pursue knowledge and drives academic achievement, Edutopia. Online at http://www.edutopia.org/stw-project-based-learning-best-practices-guide.
5. NGSS (2013). Next Generation Science Standards, http://www.nextgenscience.org/next-generation-science-standards.
6. Johnson, S. (2010). *What the hell is Minecraft and why the hell should you care?* Online at http://www.g4tv.com/thefeed/blog/post/707771/what-the-hell-is-minecraft-and-why-the-hell-should-you-care/.
7. Minecraft (2013). Online at https://minecraft.net/.
8. Heick, T. (2013) 5 Lessons To Learn From Minecraft In Education, TeachThought, Feb. 2, 2013. http://www.teachthought.com/trending/5-lessons-to-learn-from-minecraft-in-education/.
9. Richter (2013). XML and HTML with Python. Online at http://lxml.de.
10. Bird, S., Loper, E., and Klein, E. (2009). Natural Language Processing with Python. O'Reilly Media Inc.
11. Loper, E. and Bird, S. (2002). NLTK: the Natural Language Toolkit. In *Proceedings of the ACL-02 Workshop on Effective tools and methodologies for teaching natural language processing and computational linguistics - Volume 1* (ETMTNLP '02), Vol. 1. Association for Computational Linguistics, Stroudsburg, PA, USA, 63-70.
12. McCallum, A.K. (2002). MALLET: A Machine Learning for Language Toolkit. Online at http://mallet.cs.umass.edu.
13. Blei, D.M., Ng, A.Y., and Jordan, M.I. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res.* 3 (March 2003), 993-1022.
14. Wkipedia (2013). Ground Truth. Online at http://en.wikipedia.org/wiki/Ground_truth.