Curricula Analysis and Field Study Design and Publication

Final Report

Spring 2018

Adiza Awwal, Mahzabin Hasnath, Lalitha Madduri

COMS 3998 W 028

Table of Contents

Abstract

Our goal for this project was to implement a means of data collection and analysis to better improve the SAGE's (Social Addictive Gameful Engineering) user-centered design aspect within the "Gameful Affinity Space" and "Field Study Design and Publication Strategy" epics, in particular, Curricula Integration, Learning Metrics, and SAGE Feasibility Study. In this report we will describe the progress made in the Sage Feasibility Study and Publication, Curricula Integration, and Gameful Direct Instruction epics. The features and stories we have completed so far are the Parson's Puzzle Library (26), Curricula Analysis (27), User Centered Design | codeHER (325), Semi-Structured Interview Protocol (347), Survey and Distribution Plan (341), IRB Training and Conflict of Interest Document (356), Pre-distribution Volunteer/Teacher Survey (366), and Curriculum Presentation (30). Feasibility Publication (340) is currently underway.

1. Introduction

Our motivation for this project is to update the way students interact with SAGE's UX/UI as well as a concentrated effort to gather and analyze student and teacher feedback. We want to incorporate user feedback analysis when implementing SAGE's features to better present its mission of infusing computational thinking in 6-8 curricula. We are motivated by the need to bring SAGE out into the real world into diverse classrooms, preparing surveys and setting up parameters to make the best use of experimentation time. Within the Field Study Design and Publication Strategy epic, meanwhile, we will collect and analyze user feedback from volunteers, teachers, and students before and after classroom sessions, ensuring seamless integration along the way.

2. Related Work

2.1) Hour of Code Impact Studies

Code.org, a digital playground with the mission of expanding access to computer science in schools and increasing participation by women and underrepresented minorities with online gameful instruction, releases its impact studies on the research portion of its website. The page assess the impact of Hour of Code programs, a global movement by Computer Science Education Week and Code.org reaching tens of millions of students in 180+ countries through a one-hour introduction to computer science and computer programming. The website details its user count, number of hours students spent on their learning materials, analysis of quantitative and qualitative data about the projects, as well keeping an emphasis on demographic information. On the research portion of their website, this data is presented and areas for

I like computer science

When examining the largest changes from pre- to post- for the statement, "I like computer science", the findings suggest that female students' attitudes towards computer science were the most positively impacted by the Hour of Code activities.

Table: 5 largest increases in perception shift for "I like computer science"

Student Gender	Age Group	Prior Experience	N	Pre Hour of Code likert-scale rating	Post Hour of Code likert-scale rating	Absolute change in % who agree
Female	High school	Never done an Hour of Code	158	2.42	2.78**	20.3%
Female	High school	All	265	2.59	2.88**	17.7%
All	High school	Never done an Hour of Code	513	2.64	2.91**	11.1%
Female	All	Never done an Hour of Code	644	2.76	3.06**	10.1%
Female	Middle school	Never done an Hour of Code	286	2.66	2.98**	9.4%

improvement are identified.

Figure 2.1: Data from “The Hour of Code: Impact on Attitudes Towards and Self-Efficacy with Computer Science”

One of the most interesting observations appears to corroborate the hypothesis that an exposure to informatics since the early age is important to attract students independently from their gender. Findings showed that after completing one Hour of Code activity students report liking computer science more, feel that they are better able to learn computer science, and are better at computer science than their peers. Below is a graph representation created by Stanford researcher Chris Piech evaluated the various computer programs that students submitted to two computer programming puzzles from our popular Hour of Code tutorial.

Figure 2.2: Data from 2 puzzles from Hour of Code 2013

2.2) An Emergent Approach for Studying Serious Games

Files\\guy
1 reference coded, 0.24% coverage

Reference 1: 0.24% coverage

: Six to eight at night Johns Hopkins position.

Files\\lalitha
1 reference coded, 2.57% coverage

Reference 1: 2.57% coverage

I teach through code her and take it all and we teach a different venues taken out how standards in terms of the curriculum that we want to follow so the Baltimore city school curriculum that we want to make sure the students are learning.

Files\\saniya
3 references coded, 2.72% coverage

Reference 1: 0.03% coverage

Yes

Reference 2: 1.27% coverage

Is on that all schools I go start with block based programming and then the transition into type based.
Speaker 2: All schools have to cover web development in one way or another.

Reference 3: 1.42% coverage

The programming languages that they teach to the students entrance and strings so whatever up in the final product or whatever they end up teaching has to be directly linked to what the kids want to do

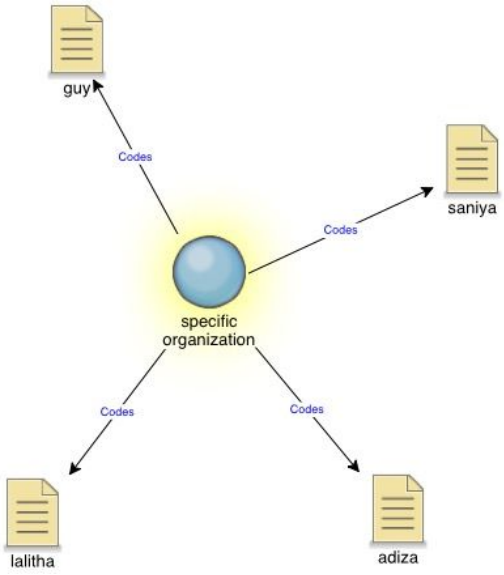


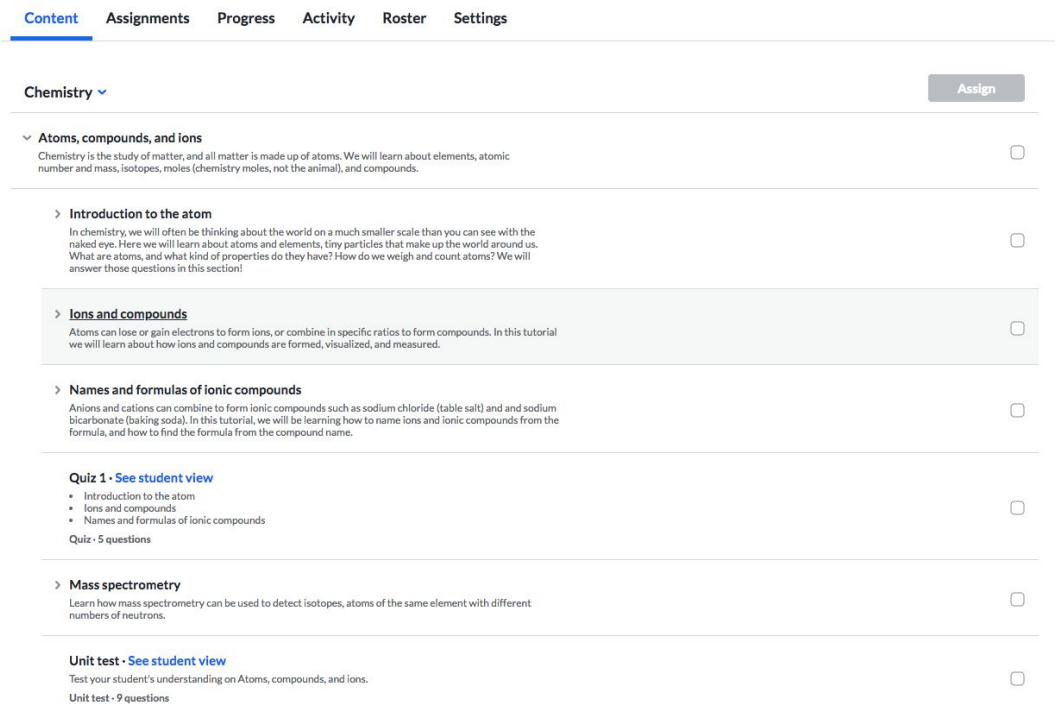
Figure 2.3: Aggregated data and visualizations from NVIVO transcript analysis

Much of the field work format was grasped from the research from Matthew Sharritt's *Serious Educational Game Assessment* (SEGA.) Much of his work chronicles his techniques for qualitative analysis surrounding the user experience for game players. Sharritt emphasized open ended inductive experiments to then begin deductive hypothesis testing. In his methods, the research and the researcher's hypothesis dictate or predetermine what is most important.

He goes on to take from many different fields such as ethnomethodology and, a crowd favorite, grounded theory which is heavily utilized in the Field Studies analysis work. Matthew further describes his work to be something that is a "hybrid" in guiding researchers to find patterns in participant behavior and measure engaging learning experiences.

2.3) Khan Academy

Khan Academy is an online learning platform geared towards democratizing education by making it more accessible via instructional videos. Similar to SAGE, Khan Academy has an instructor dashboard where teachers can assign lectures and delegate assignments to their students. In delegating assignments, instructors can toggle between subjects. Within each subject, instructors have the ability to assign entire sections or certain subsections across different sections. This grants instructors freedom to mix and match different parts of the curriculum to assign to the students. The user interface consists of a drop-down tree menu and checkboxes for the options.

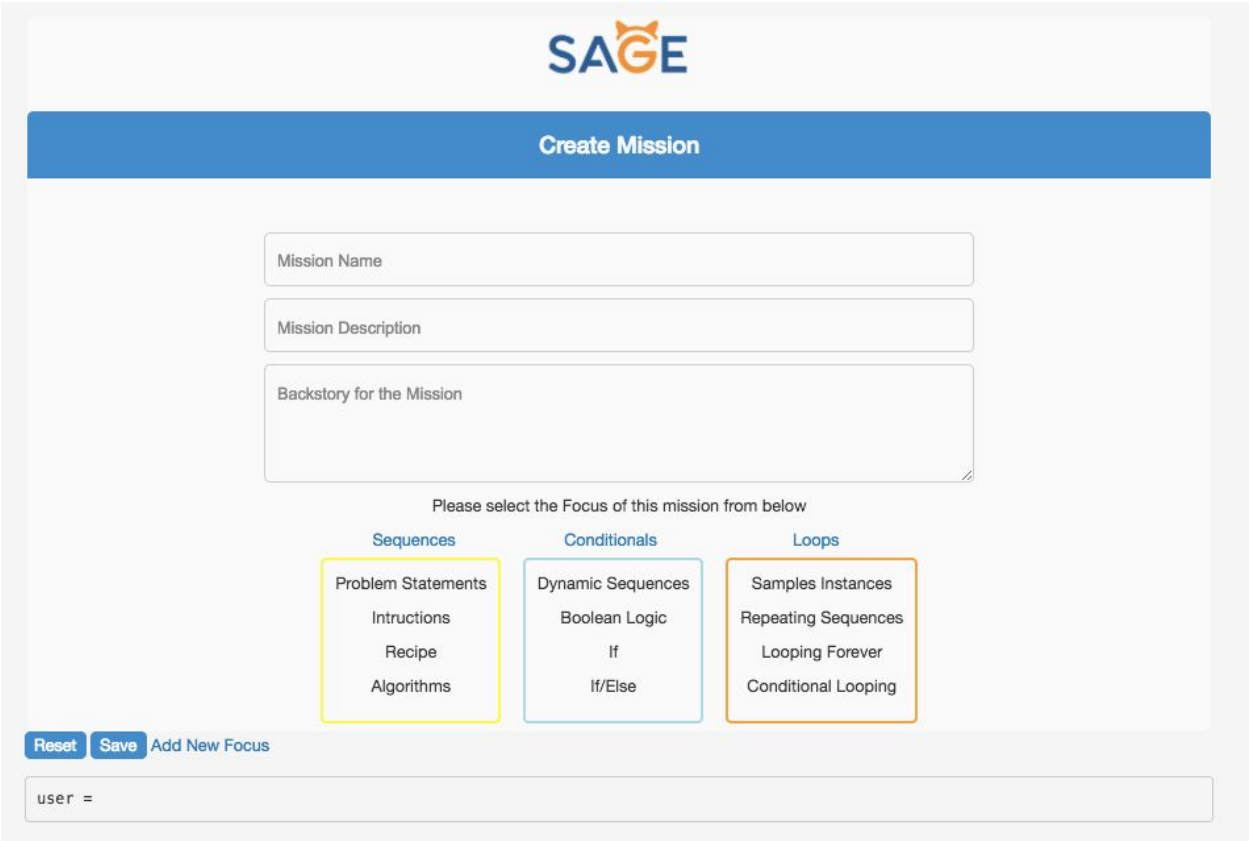


Accomplishments

This semester, we focused on [fill in your parts here], and centering the computational thinking and curriculum within the Gameful Affinity Space.

3.1 Instructor Dashboard - Mission Management

This semester, we aimed to revamp the Instructor Dashboard to better present computational thinking concepts. We modeled the UI off of Gavi Rawson's wireframes [**2**]. Moving away from the traditional tree structure, the presentation of the computational thinking concepts was separated into a list structure. On hover, a description of each concept appears which reinforces the curriculum within the dashboard. From the semi-structured interview sessions instructors provided feedback that they wish to have the ability to add their own focus in creating missions, thus prompting the creation of the Add Focus feature where instructors can add their own computational thinking concept or curricular item.

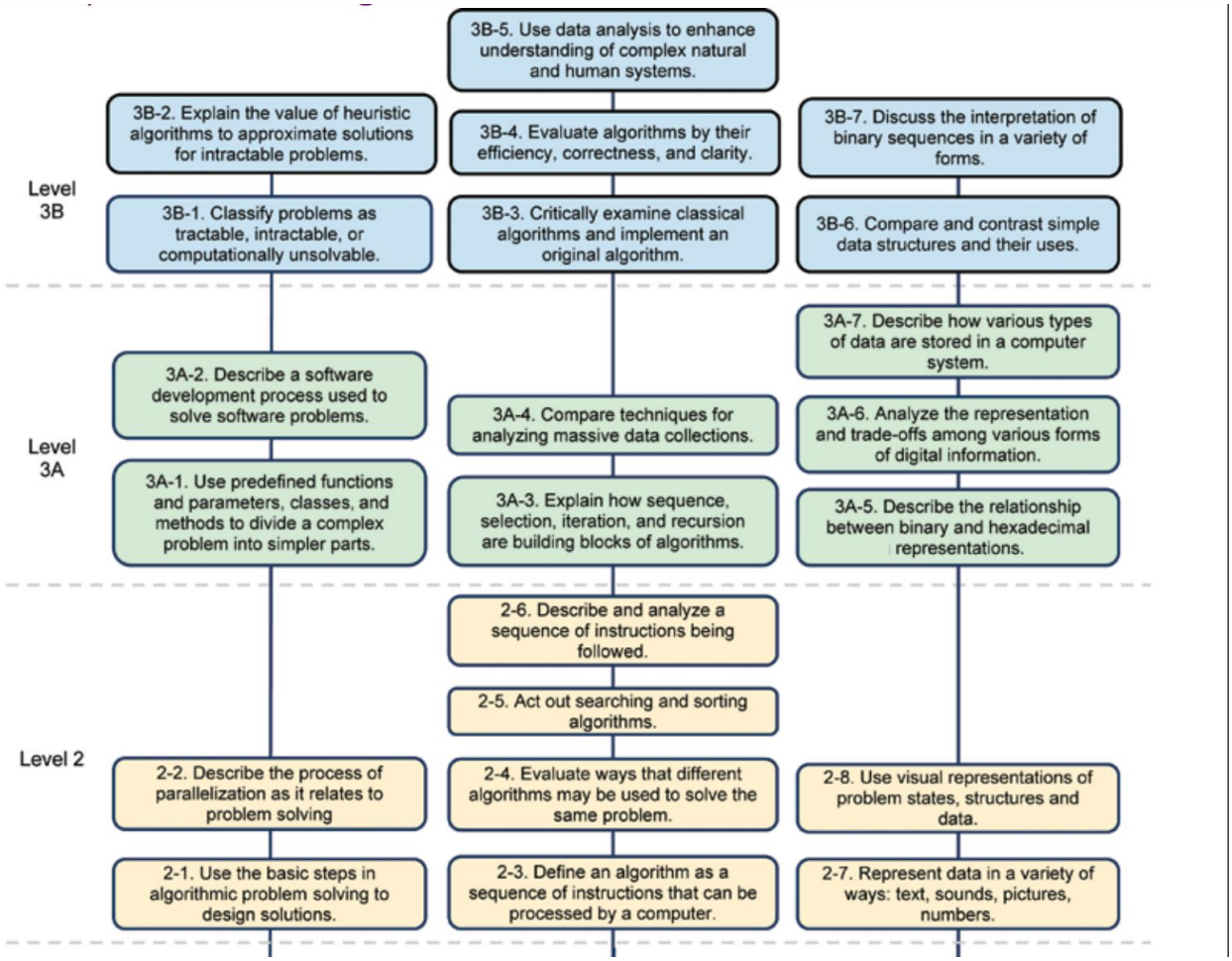


instructor_addFocus.html

```
<form novalidate class="simple-form" name="LPform">
    <div class="c1">
        <div><img src="/public/images/logo.png" height="60"></div>
        <div class="head1">
            <h2>Add New Focus</h2>
        </div>

        <input type="text" name="AFconcept" ng-model="AF.concept"  placeholder="Computational Thinking Concept" ng-required="true"/><br />
        <input type="text" name="AFitem" ng-model="AF.item"  placeholder="Curricular Item" /><br />
    </div>
    <button ng-click="submitForm()" value="Save" >Save</button>
</form>
```

3.2 Curricula Analysis

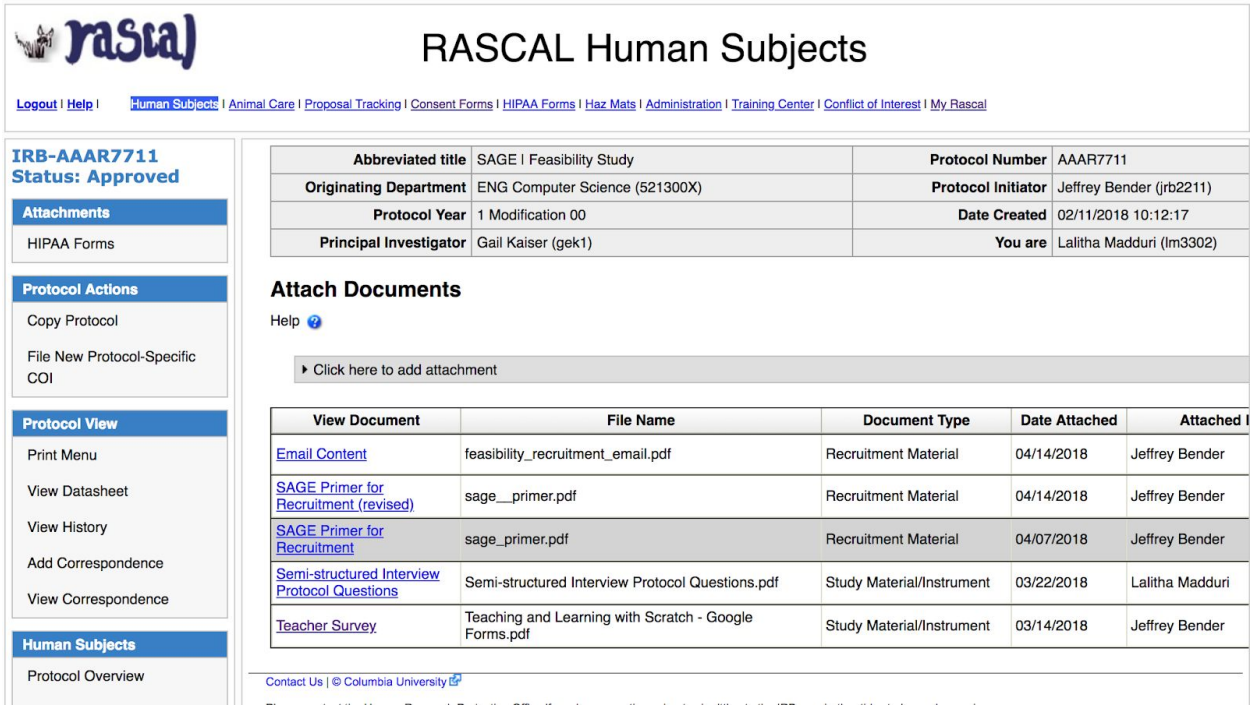


Before going into classrooms, we compiled a list of computational thinking standards across the United States and noted similarities and differences within each state's official procedures in order to better understand the needs of teachers when we do receive responses regarding the direction of Scratch. These standards are currently stored within the Field Study and Design Publication 2018 google drive folder, with various graphics such as the one above created in order to better-visualize the national standards we hope to align and improve with the use of SAGE.

3.3 IRB Training and Protocol

3.3.1 IRB Documentation

The Institutional Review Board (IRB) is an administrative body established to protect the rights and welfare of human research subjects recruited to participate in research activities conducted under the auspices of the institution with which it is affiliated. In order to conduct studies on-site with human subjects, a few members within SAGE have gone through the training and approval process, which takes roughly one to two hours per researcher to complete. The field studies we conduct as well as the materials we use in conducting the studies must closely follow IRB protocol and thus several iterations of the IRB protocol have been underway in terms of solidifying specific questions, interview personnel and the environment in which the interview must take place. Throughout the progression of these changes and decisions, the protocol has been tracked and amended on RASCAL and the documents used were regularly uploaded and updated so expedited review process could be granted. SAGE therefore currently has expedited IRB approval for its fieldwork.



In finalizing the IRB Protocol, certain systems such as medium of contact and informed consent policies had to be established. For example, grounded theory, a general way of thinking about conducting qualitative research that describes, either explicitly or implicitly, the purpose of the qualitative research, the role of the researcher(s), the stages of research, and the method of data analysis, is being used to gain specific insights on user feedback from volunteers, instructors and students. The largest factor in establishing an analysis on qualitative data surrounds selecting what qualifies for further analysis and what does not. The selection of the most frequently occuring patterns are those that may qualify for further investigation. In particular our tool of choice happened to be the qualitative analysis tool NVivo.

3.3.2 Distance Partnership with codeHER and TeCanal

In conducting in-person field studies, two student-run nonprofits that teach computer science to a varied demographic of youths, both incorporated and located within Maryland. Field studies will be conducted in the outreach venues through which these nonprofits operate.

CodeHER is a non-profit organization that aspires to close the gender gap in computer science by targeting girls in their late elementary and middle school years. All of our programs have one goal: to expose girls to technology and spark a passion in engineering that stays with them throughout their high-school and college years. CodeHER utilizes websites such as Code.org in order to manage classrooms and teach computational thinking skills, and dabbles in using Scratch, Python, and Blockly.

TeCanal is also a student-founded nonprofit organization tackling the issues of poor STEM education and lack of access to technology in low-income communities. As its name suggests, Te(ch)Canal aims to be a "canal", or link, between technology and the communities that will benefit from it, connecting youth in all communities together in pursuit of a unified goal: STEM education for all. TeCanal mainly teaches computational thinking through Scratch and Code.org's Hour of Code modules in simplified lessons. The target demographics include children of all genders, individuals with disabilities, low-income minorities, and immigrants ranging from elementary school-middle school. The teachers thus have more problems to tackle than simply teaching computational thinking, but their experience is extremely valuable in ensuring that SAGE is an inclusive platform that addresses issues in various communities.

3.4 Survey Development and Distribution

3.4.1 Pre-Distribution Survey

Before we began distribution and development of the formal survey we will be sending out for the purposes of using within the publication, we created and distributed a pre-distribution survey and the SAGE official survey for teachers and volunteers through codeHER and TeCanal. The pre-distribution survey was developed by referencing the UCD HURIE literature as well as a compilation and analysis of CT curriculum standards across US public schools to give the researchers a bit of foresight before the distribution of the survey in order to examine sample responses regarding the way teachers approach CT lessons.

3.4.2 Teaching and Learning with Scratch Survey

The survey itself was adapted by Jeff, Johan, our team, and others in the group in order to assess educators teaching styles, familiarity with Gameful Learning and overall impression of Scratch as a tool to use in education computational thinking. The survey aims to surface opportunities for improving teaching and learning with Scratch with a focus on grade 6-8 students. The responses are reported in an aggregated and anonymized manner regardless of whether identifying information is provided. The South Bronx Early College Academy teachers, who SAGE partnered with to conduct semi-structured interviews and user centered design sessions this semester, all took part in this survey, as well as various student teachers from non-school-affiliated organization.

The distribution of this survey began in February, but currently it has been incentivized and pushed to various Columbia computer science organizations, Facebook groups, e-mailing lists, and beyond in order to reach the goal of garnering 200 responses so that proper statistical analysis can be done on the data for the publication. The survey is currently at around fourty responses, including all versions distributed.

3.5 Semi-Structured Interviews

Semi-structured interview protocol was developed by researchers who had experience teaching computational thinking to children through various organizations. These interviews employed questions that elaborated on the Teaching and Learning with Scratch survey, drawing upon the personal experiences of each interviewee to see if SAGE's features would be received by current teachers with positive feedback. The interview questions spanned instructional paradigm, gameful learning, formative assessment, summative assessment, and a call to action. The direction in which the interviewer wanted to take these questions was up to them, making each interview unique in its delivery and information gathered. The questions used can be found in the google drive as well as the RASCAL page.

Following the distribution of the pre-distribution survey and survey, field studies comprising one-on-one semi-structured recorded interviews were initiated in the codeHER and

TeCanal classrooms in Maryland as well as remotely. During this time, seven or eight full-length interviews were completed, some recorded through a microphone and others transcribed if the interviewee did not provide consent to record the session. Within this timespan, we were able to get interviews from individuals who taught through TeCanal, codeHER, Girls Who Code, and independent tutors. After the semi-structured interviews in codeHER and TeCanal were completed and analyzed without much breadth, various researchers visited the South Bronx Early College Academy to conduct one-on-one semi-structured interviews with teachers from the school. These interviews lasted, on average, 30-45 minutes. As a whole, instructors were hopeful of the direction SAGE was going in and eager to use the features that SAGE is currently implementing, especially intelligent hinting and class creation and management.

3.6 User-Centered Design Sessions

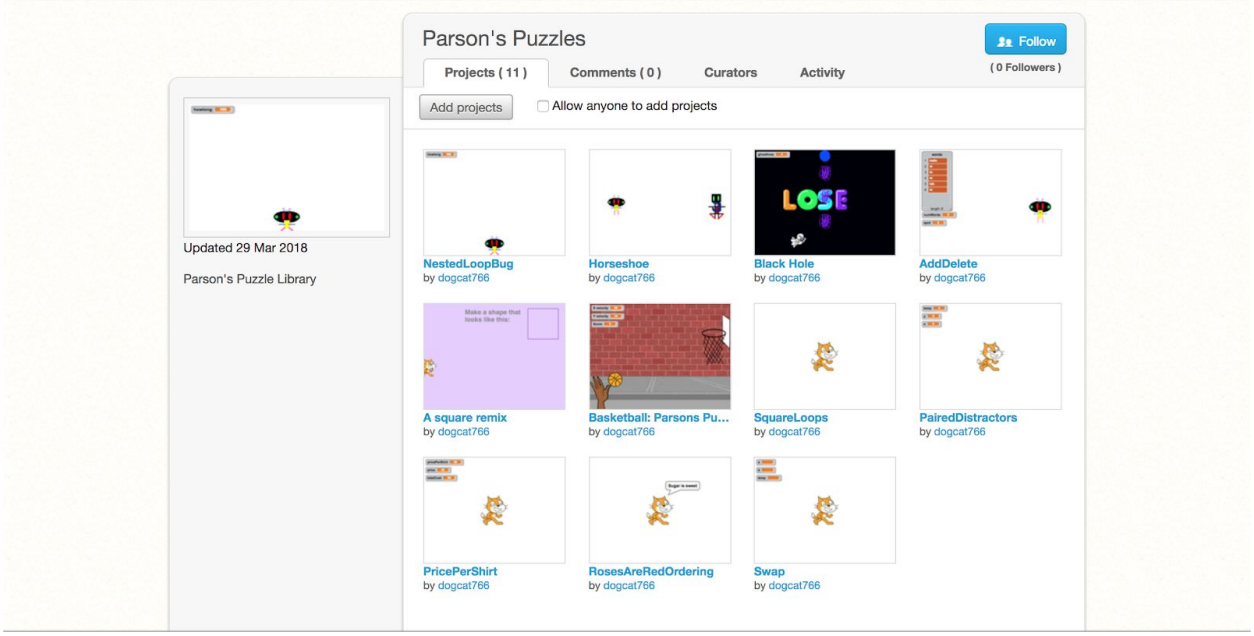
User-centered design is an iterative design process in which designers focus on the users and their needs in each phase of the design process. UCD calls for involving users throughout the design process via a variety of research and design techniques so as to create highly usable and accessible products for them. User-centered design sessions included a pre-prepared course to take the interviewee through the SAGE instructor and student dashboard, exploring SAGE's features while getting real-time feedback on which features they would or would not use. The basic skeleton used for the UCD sessions is included in the feasibility report google drive folder.

The demonstrations in UCD sessions are limited by the bugs present in the instructor dashboard on the developer website. When we conducted UCD, we emphasized the working portions as well as the potential that some of the in-progress tools, such as intelligent hinting, have. UCD sessions were conducted with the CEO and COO of codeHER, the CEO of TeCanal, two of the researchers within this group, experienced TeCanal volunteers, and a Columbia student well-versed in teaching Scratch at a coding bootcamp for children. After these initial sessions, the researchers conducted an hour-long UCD session with the SBECA teachers, exploring the student and instructor dashboard and fielding questions and suggestions for the framework. In total, around five UCD sessions were done with varying group size, with overwhelmingly positive feedback towards SAGE's features. Aspects that teachers liked as a

whole were the class creation and management features, the instruction blocks, the ability to limit the palette, and the prospect of a video library for independent student tutorials and learning.

3.7 Parson's Puzzle Library

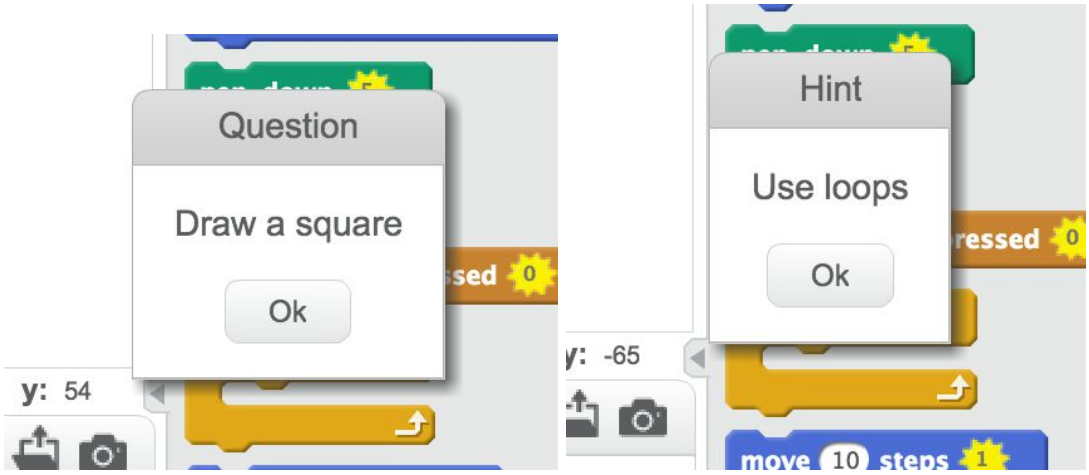
In a Parsons puzzle, the student is provided scrambled code and asked to reassemble the program based on a problem specification. Parson's Puzzles were created within the affinity space and quests on the instructor dashboard in design mode to build a gameful direct instruction library and use as demonstrations when time comes with further UCD sessions. These puzzles are created referencing literature as well as a thorough compilation and analysis of the computer science curriculum state standards across the United States. Parson's Puzzles are being built such that they are embedded within SAGE. The puzzles are created in design mode with their purpose and answers documented in a separate document and saved as .sb2 files. This puzzle library aims to hit various CT concepts, such as looping, conditionals, and more.



The original Parson's Puzzle library was built on the original Scratch website, and the .sb2 files were saved while there was no saving functionality within SAGE. Then, with the help of Johan, the puzzles were migrated over to SAGE with the appropriate point values assigned

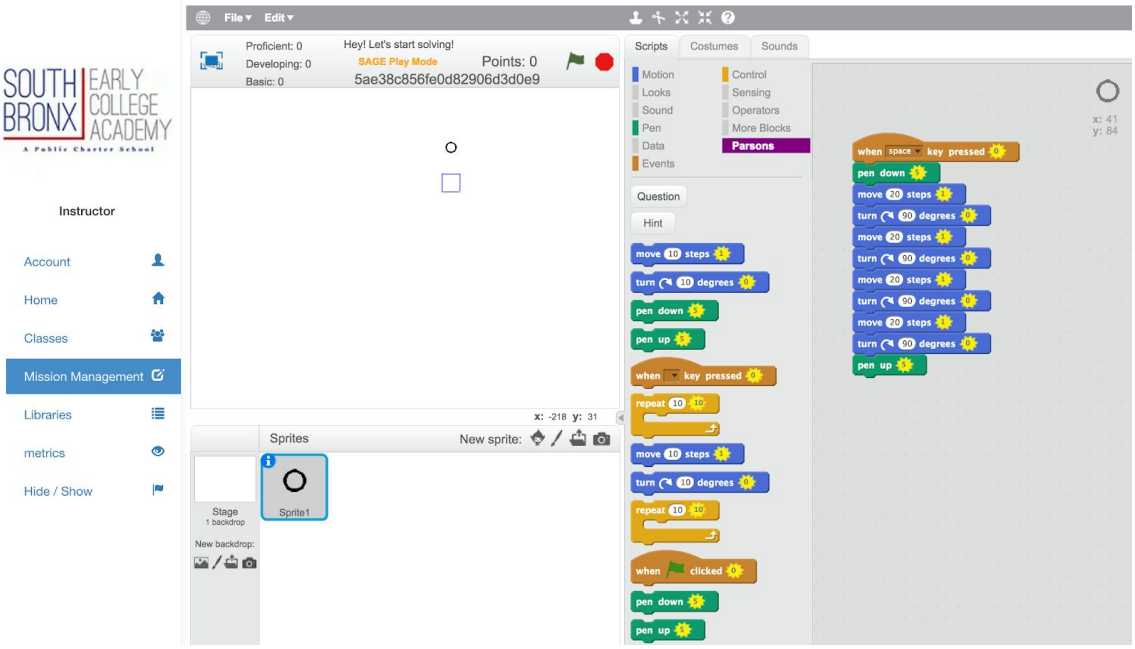
and instructions included, with documentation currently being updated as the Parson's Puzzles are being migrated over for use in future UCD sessions.

A simple example of a Parson's puzzle is shown below. Below is the Parson's section of the palette provided to the student, displaying the objective and providing hints for the student to get a higher number of points.



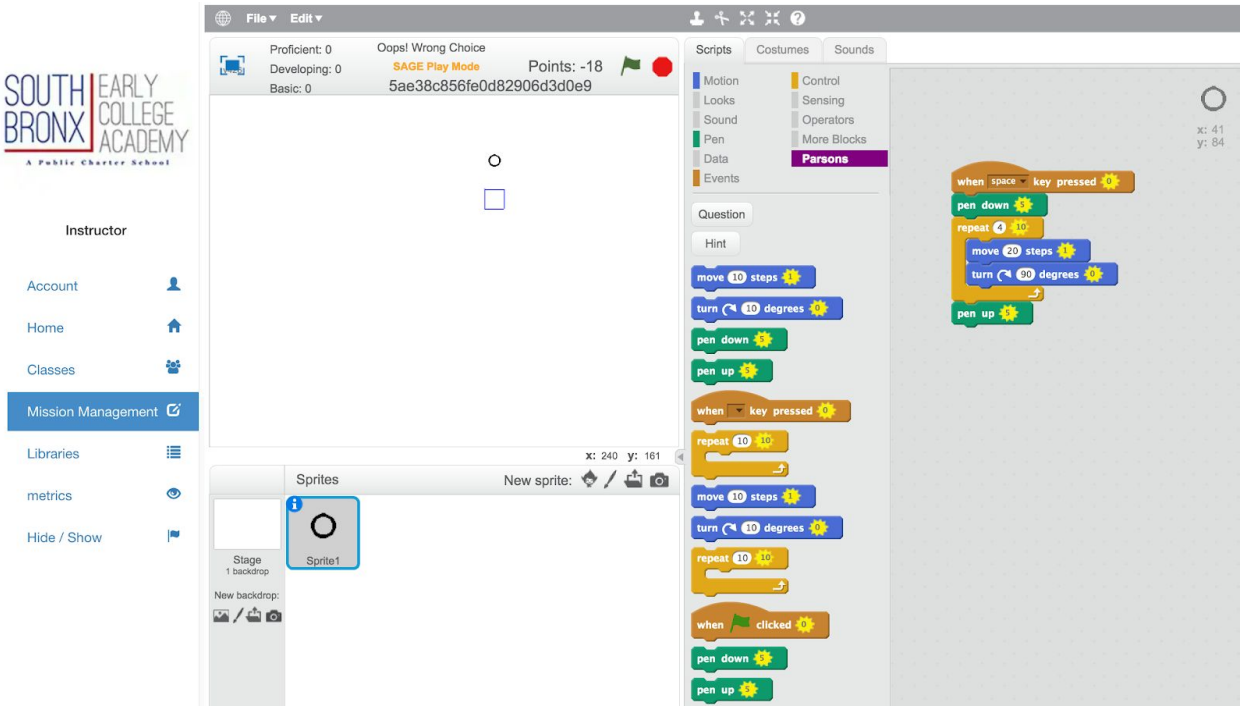
The student will have a selection from the blocks chosen by the instructor, shown on the Parson's pane. These blocks need to be used and arranged properly to solve the objective. First

shown is the naive solution, repeating the same blocks to draw the square by drawing each side



four times.

After that is the optimal solution for this Parson's puzzle, utilizing a loop, which gives the student more points.



3.8 Field Study Analysis

Building from the progress set in place from semi-structured interviews is the analysis of the data, semi-structured interviews and user centered design sessions were be leveraged for their empirical data. The first and most intuitive step was documenting participants responses. Manually written transcription in addition to natural language processing techniques were leveraged in order to assess feedback, using a Watson Speech-to-Text API to transcribe the recorded field studies and compiling them to analyze later on. In terms of analysis, in some ways, transcripts require abstraction of data. In addition to assessing what is recorded on the transcript, there are elements that are insightful from the actual interaction. Shown below is a word cloud created from compiling a stripped and cleaned compilation of all the statements said during each semi-structured interview protocol, showing the frequency each term came up in the interviewee's responses. Though informal, this cloud helps analyze the direction to be heading in when conduction future UCD sessions as well as what to expect.

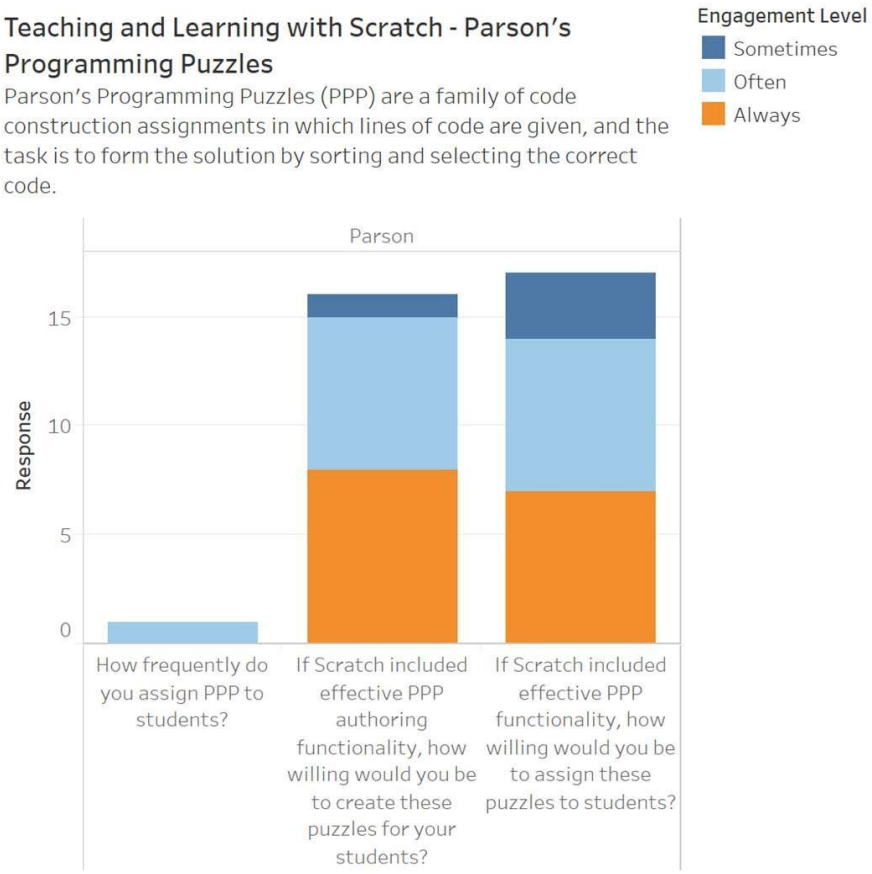
The data has more formally been analyzed through Grounded Theory, a general method of comparative analysis where we made consistent comparisons across interviews and grasp general themes from users' responses. Moreover, general trends and commonalities were identified particularly pertaining to easily comparable points of reference (e.g. age groups, teaching environments, teaching history among others.)

The fieldwork was conducted through in-person voice recordings as well as transcriptions. Meanwhile, surveys were distributed through Google Forms. The data will be analyzed through Grounded Theory, a general method of comparative analysis where we can make consistent comparisons across interviews and grasp general themes from users' responses.

A major step in configuring the selection of tools and establishment of processes is in selecting tool of choice. Our option were IBM's Speech to Text, Mozilla Deep Speech, or Carnegie Mellon Sphinx Toolkit. Through trial and error ultimately IBM's Speech to Text tool, NVivo and Carnegie Mellon's sphinx toolkit were selected and utilized. There are still several things to work toward in preparation for the coming semesters as moments have been slightly disheveled. Though success was made a streamlined process for recording and transcribing sessions with both teachers and volunteers should be established given we utilized several tools over the course of Sprint 5. However, there was in fact use of an established toolkit / establishment of processes that operates with the already existing SAGE ecosystem. After data analysis was complete, we derived clear and concise graphs as well as commentary for the data we've gathered for the publication. A similar, if not exact, process was used for the User Centered Design sessions as well.

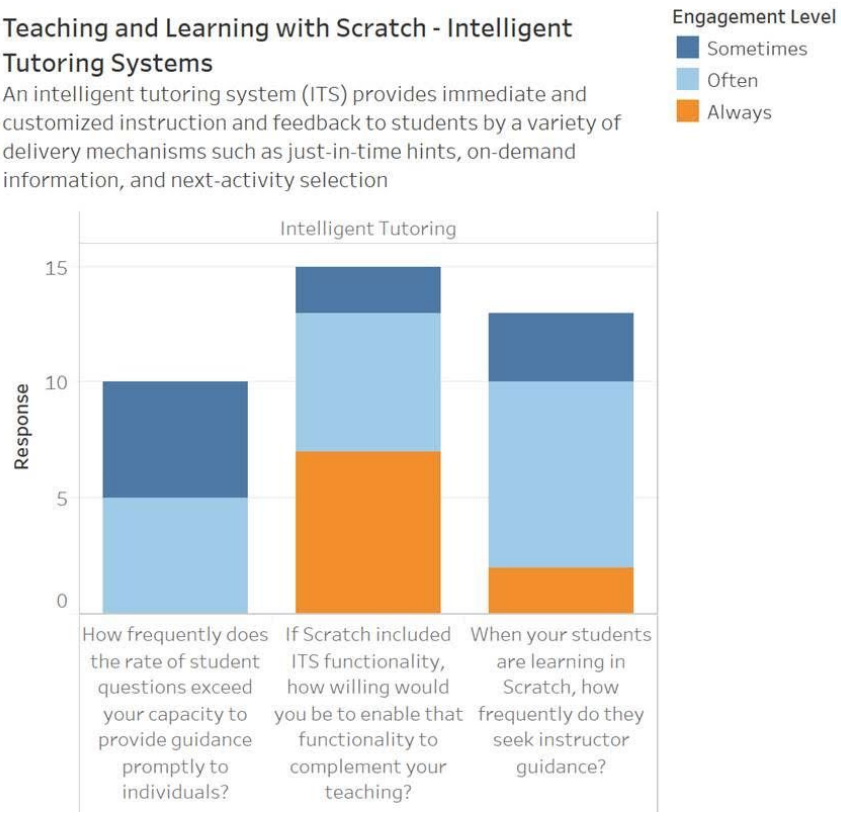
Shown below are some of the graphs created from the Teaching and Learning with Scratch survey developed in conjunction with Johan. A majority of the respondents heavily incorporate Scratch in the curriculum (>50%), and most aim to teach Computer Science, while Data Analysis, Representation, and Collection are underrepresented among the 9 CT concepts. Virtually nobody uses Parsons Puzzles but virtually all are interested to create and utilize them for the sake of teaching computational thinking, something that other websites have not incorporated in their frameworks for virtual playgrounds with code. Instructors overall feel

somewhat capable to provide prompt guidance to individuals. For ITS to truly bring value, we should aim for a bigger pain point since being overwhelmed does not seem to be that much of a hindrance. Constructionist games seem to be a staple of instructor's curriculum, but deeper questions can be asked on the level satisfaction and perceived effectiveness of the instructor-created CVG. Another good question might be asking the amount of time instructors spent in making/testing/iterating through a CVG and the quantity of CVGs deployed per class.



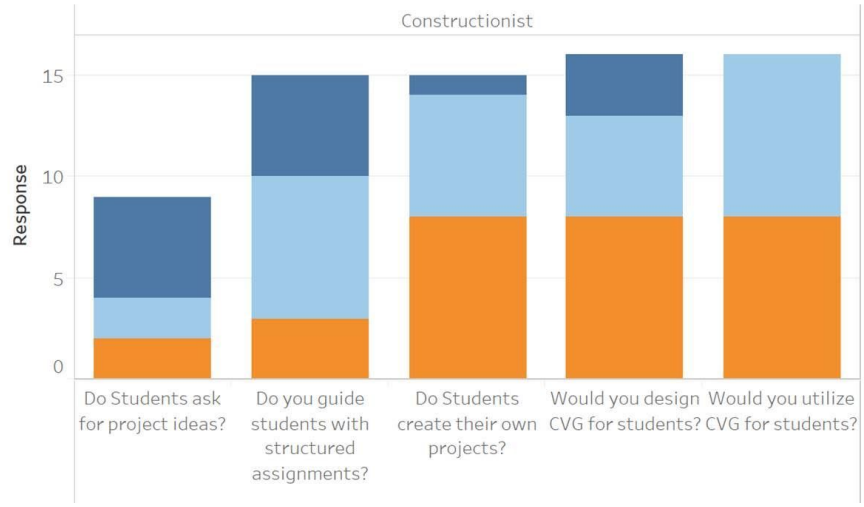
Teaching and Learning with Scratch - Intelligent Tutoring Systems

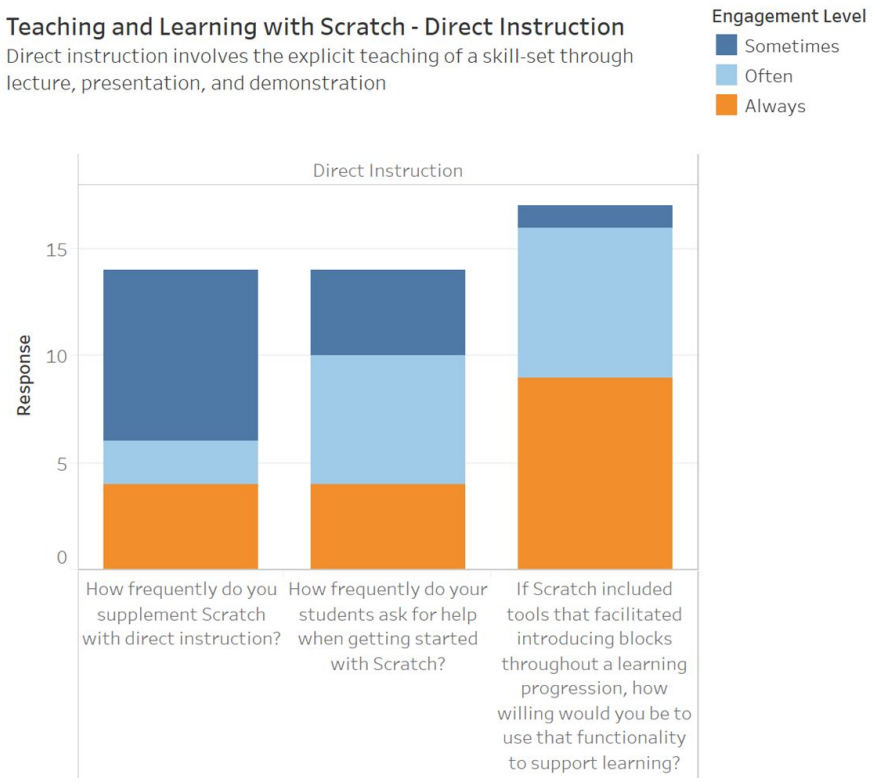
An intelligent tutoring system (ITS) provides immediate and customized instruction and feedback to students by a variety of delivery mechanisms such as just-in-time hints, on-demand information, and next-activity selection



Teaching and Learning with Scratch - Constructionist Video Games

Constructionist Video Games (CVG) are designed environments in which players construct personally meaningful artifacts in order to overcome conflicts or obstacles resulting in quantifiable outcomes.





Documentation

All the materials used within the field studies can be found in the 2018 Sage Feasibility Study folder within Google Drive. Included within are transcriptions, audio files, the surveys, UCD and semi-structured interview protocol, and locations in which we conducted field studies in Maryland and beyond.

In terms of Parson's Puzzles, the Parson's Puzzle Library stored on Scratch itself can be found at https://scratch.mit.edu/studios/4945809/. Instructions to play the Parson's Puzzles are documented in the instructor dashboard. One example is http://dev.cu-sage.org/instructor/#/coursePage/59f8c6fdc1bfb23c4ced8e20/course/5ae26cadc27be8135c0f946a/Assignment/1/id/5ae26d8cc27be8135c0f946b/Instruction. Documentation for the Parson's Puzzles can be found here: https://gudangdaya.atlassian.net/wiki/spaces/SAGE/pages/396918803/Parson%27s+Puzzle+Library.

Future Work

Future work includes converting the UI of the dashboard to the Material-UI framework. This semester marked the transition to Material-UI from its traditional Bootstrap and AngularJS frameworks. In order to forge a sense of cohesion, the use of Material-UI must be consistent across the platform. Standardizing the UI not only introduces consistency so the user can focus on their goal as opposed to the differing UI implementation, it also enforces familiarity since the framework borrows from Google's Material Design [**3**]. Other teams have already followed this paradigm. Other future work also includes the full integration of the "add focus" feature within SAGE so that instructors have the ability to curate their own curriculum paths.

The next steps in field studies are to test SAGE's features within the student populations at various schools. codeHER and TeCanal are enthusiastic to utilize this framework as soon as possible, and impact studies made in the classroom long-term and short-term should be assessed between the ways the students learn without the aid of SAGE and with the aid of SAGE. Inspired from Code.org's impact studies, we might also see how various demographics respond to the use of SAGE and tweak it to make it more accessible for all communities that will be using this framework in the future. Additionally, a video library on how to use SAGE and its features, especially Parson's Puzzles, is currently in the works due to high demand from instructors during the field studies. This will hopefully facilitate future UCD sessions.

In terms of analysis of the field studies, we can further refine the UCD, semi-structured interview protocol, and survey distribution processes to gain more insight into the data that we currently have. Asking more nuanced questions not only to gauge the current impact of SAGE, but also utilizing that data to improve the workflow of the teams in future semesters, would be extremely helpful to ensure that SAGE is going in the right direction. The data that we currently have needs to be graphed and analyzed in a visually appealing and persuasive format for the design publication to come.

Following the completion and assessment from the student progress information and individualized student interaction would be research into the effect of making this change.

Following the new change to the Learning Metrics suite would be research and development into the next steps and how this added feature impacts the user space. After embedding computational thinking within the interface of the Affinity Space, further research should be done to investigate the effects of updated interface. Student assessments, as well as student and teacher feedback, would be evaluated and analyzed to see the correlation, if any, between the updated interface and student performance. In addition, feedback should be evaluated to determine areas of improvement.

Conclusion

References

[1] "Khan Academy." *Khan Academy*, https://www.khanacademy.org/

[2] "Wire frames." https://github.com/cu-sage/sage-frontend/blob/master/wireframes/_img/games/2%403x.png

[3] "Material-UI." *Material-UI*, https://www.material-ui.com/#/

[4] "Code.org Research" https://code.org/research

[5] L. Seiter, B. Foreman. "Modeling the Learning Progressions of Computational Thinking of Primary Grade Students." John Carroll University.

[6] Serious Educational Game Assessment: Practical Methods and Models for Educational Games, Simulations and Virtual Worlds. Sense Publishers, 2011.