Understanding Collaborative Practices in the Scratch Online Community: Patterns of Participation among Youth Designers

Deborah A. Fields, Michael Giang[*], & Yasmin B. Kafai[†]
Utah State University, 2830 Old Main Hill, Logan, UT 84321
[*]Mount St. Mary's College, 12001 Chalod Rd., Los Angeles, CA 90049
[†]University of Pennsylvania, 3700 Walnut Street, Philadelphia, PA 19104
deborah.fields@usu.edu, mgiang@msmc.la.edu, kafai@upenn.edu

Abstract: Most research in massive online youth communities has focused on understanding patterns of participation and collaboration in games, social networks, and virtual worlds. Few studies have examined the nature and dynamics in amateur design communities where youth contribute content they have designed themselves. In this paper, we examine quantitative trends of participation in a youth design site focused on programming. Scratch is an online community with over 1 million registered youth designers 11-18 years of age. Drawing on a random sample of 5,000 youth programmers and their activities over three months in early 2012, we examined log files that captured the frequency of their contributions and comments on the site, making visible distinct classes of users who engaged in different sets of practices that support design on a collective scale. In the discussion we discuss implications for the design of collaborative spaces, tools, and communities.

Introduction

A growing body of research in massive online communities has sought to understand patterns of participation through collaboration in online sites, games, social networking sites, and virtual worlds. Research into these sites has provided insights into how people develop collaborations in massive communities within and beyond the designed structures, for instance developing fluid social networks for information gathering and gameplay (Williams, Contractor, Poolec, Srivastad, & Cale, 2011), building trust in long-term relationships that promote more effective teamwork (Chen, 2012), and engaging in knowledge sharing and problem solving in game forums (Steinkuehler & Duncan, 2009). Our own studies illustrated how younger users (tweens) drew on many social resources to learn secret commands in a virtual world, navigating relationships in-person as well as with familiar and unfamiliar people in the online community (Fields & Kafai, 2009).

Increasingly however, interest is growing in online communities where users contribute the main content through collaborative or "cooperative" work (Benkler, 2006). In such communities, often organized by community members rather than companies, the content is generated by members themselves. Research into these communities begins to reveal the motivation behind such volunteer collaboration as well as mechanisms for holding together both small and massive collaborative work. Finding fellow collaborators to work on a project and hold that collaboration together in an amateur design site can be a challenge. Indeed, some studies have noted in these cases that most groups fail (e.g., Luther, Caine, Zigler & Bruckman, 2010). Communication by leaders both to coordinate work (Luther et al, 2010) and to maintain socio-emotionally sustaining personal and social discussions (Aragon, Poon, Monroy-Hernandez, & Aragon, 2009) is key in nurturing collaborative design work. This importance of high levels of communication has also been noted in larger scale studies of native "in the wild" collaborations. For instance, Benkler (2006) noted the importance of recognition and communication by key leaders in sustaining vast unpaid volunteer contributions in Wikipedia and Linux. Yet it is unclear what other site-wide practices and design structures might support collaborative learning in amateur design sites. In addition, studies are rare for *youth* amateur design communities that are the focus of this paper. Understanding youth amateur design communities writ large can allow us to make more informed design decisions on how to sponsor collaborative learning at a collective level as well as which users may need scaffolds in participating in collective design communities.

In this paper, we tackle two challenges related to understanding collaboration in massive websites: (1) to understand *site-wide* group dynamics and behaviors that depict and promote collaboration at a collective level and (2) to study a *youth* programmer community. We examine broad trends of participation in the Scratch online community (http://scratch.mit.edu), with over 1 million registered youth predominantly aged 11-18 years who share creative programming projects. Drawing on a random sample of 5,000 active designers and their activities over a three-months time period in early 2012 we examined log files that captured the frequency of different kinds of

contributions and comments on the site and addressed the following research questions: What are the profiles of users on Scratch.mit.edu and what is their participation over time? How do users engage in the social and/or creative aspects of Scratch.mit.edu? Are there users who engage in one versus another? What role does gender play in users' participation in the website? In the discussion we review our approach to profile analysis and outline implications for the design and study of collaborative online spaces and tools.

Background

Prior CSCL work has focused on understanding various dimensions of smaller group work including different group arrangements (e.g., Engelmann & Hess, 2010), scaffolds for promoting group work (e.g., van der Pol, Admiraal, & Simons, 2006), and interactions between online and offline collaborations (e.g., Birchfield & Megowan-Romanowicz, 2009). With some exceptions (e.g., Fields & Kafai, 2009; Rick & Guzdial, 2006), there is one assumption about collaboration underpinning many of these efforts, which is the idea that collaboration happens in small groups, often of dyads and triads, as they engage in computer-supported collaborative tasks. Research is now starting to examine collaboration in *collective* levels found in massive websites, as can be found in amateur design communities that are the focus of this paper. This may involve the smaller enterprises of individual collaboratives who work together on shared projects as well as the broader dynamics of participation in amateur design communities. Although there are growing numbers of such communities where youth share art (e.g., Deviant Art, Bitstrips), mods of games (e.g., Little Big Planet, the Sims), or stories (e.g., Fanfiction.net, Storybird), we don't know much about who is participating in these productions or who engages in which aspects of or combinations co-designing, sharing and commenting.

We do know from observational studies that a number of collaborative practices such as small group collaborative design and remixing, broadly dispersed constructive feedback, and social incentives for design have sprung up on amateur design sites. Co-design in user-created small groups is closest to the type of small group collaboration often studied in CSCL. In Scratch, many groups of kids or "companies" gather together to work on creating games (Aragon, Poon, Monroy-Hernández, & Aragon, 2009) or interactive stories (Brennan, Valverde, Prempeh, Roque & Chung, 2011). This is similar to the collaboration amongst small groups (collabs) of adult Flash video designers in Newgrounds studied by Luther and colleagues, but as mentioned earlier, most collabs fail to produce a final product (e.g., Luther et al, 2010). Another collaborative practice is remixing, where users download designs made by others, edit them, then re-post. Remixing plays a double role as a collaborative practice. Users can learn by studying and editing others' projects and they also create social links through the traces left by remixing one another's projects, sometimes creating networks of thousands of remixes from just one generative project (Monroy-Hernandez, 2012). Finally, social feedback is type of collaborative practice that supports design in online design communities. These can take the forms of peer reviews on fanfiction sites (Black, 2008), constructive criticism as well as "flames" with negative feedback (Brennan, 2011). Successful small scale efforts have been made to educate select website members to leave more positive, constructive criticism in targeted design challenges (e.g. Roque, Kafai & Fields, 2012) but without measures for how widespread certain practices are (i.e., who leaves comments and who does not) it is difficult to measure change with any certainty.

In developing a framework for parsing computer supported collaborative learning at multiple scales, we need to analytically bring together different practices that support collaborative design on a collective scale, from creating to remixing to commenting, and investigate who engages in these practices, in what combinations of activity, and for what duration. This means that we need to focus not only on the artifacts of networked collaborations but also on the "networking residues" meaning the traces left on projects or profiles such as "love-its," friend requests, "favorites," "likes" and even gifts are types of that show that users have viewed and appreciate projects (Grimes & Fields, 2012). Networking residues can even become a type of commodity as they elevate the virtual presence of a person or project through signs of popularity. In Scratch, members leverage networking residues to support user-created design contests, offering projects, illustrations, love-its, and friending as prizes (Nickerson & Monroy-Hernandez, 2011). Yet while we see evidence of a range of collaborative practices supporting design in amateur online communities, we cannot judge how widespread or distributed these practices are across a full range of users on these sites, nor what patterns of activities users take up over time. To understand connections between practices of creating and sharing that traditionally have been seen as the cornerstone of collaborative design together with collective practices that create the underlying social fabric that encourages and supports continuing and iterative design practices, we examine participation patterns through log files collaborative activities (such as designing, remixing, commenting) of a random sample of users in the Scratch online community, thus complementing prior case studies of individuals, groups of designers (collabs), or common practices of activities (e.g., remixing).

Context & Methods

Scratch.mit.edu is an online massive community where participants, mostly youth ages 11-18 years share their computer programs (Resnick et al., 2009). Kids who share an interest in programming post animations, games, stories, science simulations, and the interactive art they have made in the visual programming environment of Scratch. Launched in May 2007 out of the MIT Media Lab, the Scratch site has grown to more than 1.2 million registered members with nearly 1500 Scratch projects uploaded everyday. As a type of social networking site, activity centers around sharing user-created projects. User profiles are portfolio based, showing individuals' created projects, "favorite" projects, and links to user-created galleries (collections) of projects and recent "friends" on their home page. While there are small spaces for a thumbnail picture and city/country information, projects dominate the user profile: one gets to know others through the quality of their projects or the comments they leave. Networking residues show up in comments, inclusion in someone's "favorites," and descriptive statistics listed under a project, including the number of views, taggers, "love-its," remixes, downloads, and the user-curated galleries in which the project is located. Projects with more views, comments, and love-its may eventually make it to the front page of Scratch through categories like "Featured Projects," "What the Community is Loving," "What the Community is Viewing," and other sections. The front page is a prized area for Scratchers; having one's project on the front page (or linked from the front page) means getting more views, more feedback, and more visibility. Yet even though the Scratch site is primarily project-based, project creation and social networking are deeply intertwined and the site allows for a number of forms of participation.

Data Collection and Analyses

To understand the group dynamics and behaviors of Scratch users, we used latent class analysis (LCA) to identify communities of similar Scratch users based on their participation behavior (for more details, see Muthen & Muthen, 2000). LCA identifies the maximum number of latent classes (groups of similar individuals) based on a set of observable categorical and/or continuous variables that can be observable online activities. This process can uncover different patterns of activity in various "classes" of users in addition to casual users, social users, and hard-core users that have been identified with more traditional cluster analyses (Giang, Kafai, Fields & Searle, 2012). To do so, LCA relies on model fitting statistics and theoretical interpretation of each class to identify the optimal number of latent classes. This approach avoids the risk of identifying classes with only a few users or a class that is generally similar to another except for minor statistical differences in specific observed activity. The second advantage of LCA is its ability to create unique profiles for each latent class. When dichotomous categorical data, each latent class contains a probability of answering 'yes' to questions about participation. For example, Scratch users can choose to participate by posting a bulletin board comment. The probability of answering 'yes' to this might be 80% for the first class, 20% for another, and 10% for the third class. The first class may consist of social users, while other classes may be different types of casual users who do not socialize.

We used LCA to examine the patterns of relations amongst a set of six Scratch participation variables: *1) Remixing*: downloading, editing and reposting a project that someone else originally posted. *2) Downloading*: clicking on a project to download it. This is indicative of looking into the inner workings (i.e., programming) of a project, since all projects can be played online without downloading them. *3) Commenting*: leaving a comment on a project or a set of projects (a gallery). *4) Favorites*: clicking "favorite" on a project. Favored projects show up on the profile of the user who clicked "favorite," meaning that others users can see an individuals' "favorite" projects by others. *5) Love-its:* clicking "love-it" on a project, which leaves a heart on the project. *6) Friend Request*: sending a friend request to a user. Friend requests on Scratch are unidirectional – they do not have to be reciprocated. Friends' latest projects show up on a users' homepage, making it a way to keep track of favorite designers.

Participants included a random sample of 5004 users from amongst more than 20,000 users who logged into Scratch during the month of January 2012. This sample reflected the broader population on Scratch in regard to self-reported gender and age. As there are no definitive indicators for the correct number of latent classes, both statistical and substantive criteria were used to identify the best model fit for each wave of analyses – one for each of three months (January, February, and March 2012). However, prior to conducting LCA, we had to transform the continuous variables as they were highly skewed, with many participants engaging in no or very few instances of a practice and a few participants engaging in hundreds of instances of a practice. For instance, 4101 users left no comments, 163 users left one comment, 104 users left 2 comments, but six individuals left more than 1000 comments. As a solution, we dichotomized each variable to indicate no activity (0) or activity (1).

Findings

This section reveals different aspects of participation in collaborative practices: (1) the impact of project creation in participation, (2) variability in participation over time, and (3) differences in participation between user groups.

Project Creation Influences Participation

The first discovery we made was that making a project was a gateway for other forms of visible participation on the website. Previous studies have found that Scratch project creation and commenting are not equally distributed amongst the users. Only about 29% of Scratch site participants, primarily male users, share projects. Of these, about half contribute only to a single project (Grimes & Fields, 2012). Some Scratchers prefer activities like commenting, live role-playing or forum posting over project creation (see Brennan, 2011). Our analyses revealed that creating at least one project in a given month was a gateway to all other activities. For instance, in the month of January, there were no users who posted comments who did not create at least one project, whereas there were many users who created projects but did not post comments. This meant that users who did not create projects also did not participate in any other activities (social or otherwise) represented by the variables, making that group difficult to study with the data available. From our sample of 5004 users, 1379 created an original project in one of the three months (January – March 2012), 533 created a project in each of two months, and 313 created a project in all three months. Thus, 2225 users (67% boys, 33% girls) who created at least 1 project across a three-month period formed the new sample from which all further analyses reported in this paper are drawn. This sample represents about 44.5% of the initial random sample of users.

Participation Varies over Time

We then examined participation in the other collaborative practices (remixing, downloading, commenting, favorites, love-its, and friend requests) for each of three months, January, February, and March. To unpack these rather complex analyses, we take first a look at the findings for January (see Figure 1) that resulted in a 5-class model.
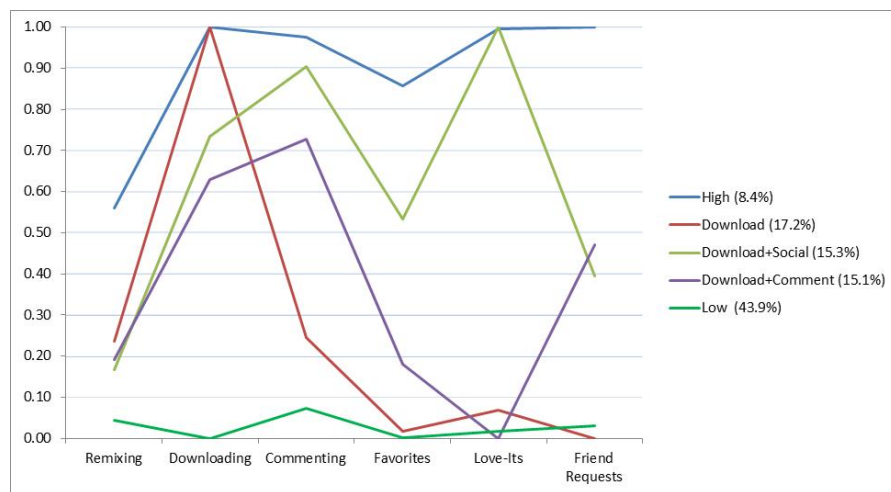

Figure 1: The 5-Class Model for January and the percentage of users for each class.

Here we see a "*high class*" profile of users (8.4% of all users) who are likely to participate in nearly every type activity studied. They have a 55% chance of posting a remix, a 100% chance of downloading a project, and very high (above 85%) chances of commenting on a project, favoriting a project, and 100% chance of loving a project and making a friend request. It is worth noting that the existence of this high class profile is consistent across all three months, as we will show later. A second profile we nickname the "download class" (17.2%) because there is a 100% chance that they will download projects from the Scratch site. There is also a profile of users who are likely to participate in most of the social networking activities available on Scratch a "*social + download class*" (16.3%), including commenting, favoriting, loving, and less likely friending. They are also very likely to download projects (75%). A fourth profile is the "*download + comment class*" (16.1%) who has a strong likelihood of participating in downloading project, commenting and projects, and friending, with low likelihoods of favoriting or loving projects. The fifth class is the "*low-level class*" (class 5, 43.9%) who are unlikely to do anything except post a project during the month.

The high class in particular is interesting, because it suggests that the core users on the Scratch site are the ones most likely to participate in remixing projects and in commenting, favoriting, loving, and issuing friend requests. In addition, the next class of users (download+social) most likely to participate in the social activities on the website are also very likely to download projects (75% likelihood). Indeed, against our expectations, there was no class of individuals who were likely to participate in commenting, favoriting, loving, or friending others without also having a strong likelihood of downloading projects, an activity which suggests that kids are not just playing projects but investigating and looking into them. In other words, besides posting a project, downloading a project is a second gatekeeper to social activity on the Scratch site. Although leaving networking residues such as favorites, love-its, and friend requests would seem to have the lowest bars for participating in a social networking website, this actually appears to be a practice in which only those who are most involved in a full range of practices on the site participate.

Another one of our key findings is that rather than having the same class model for each month, the statistical analyses suggested a different class model for each month (for details of the LCA analysis, see Appendix A). In the models for February and March we see the continuation of a high class of users who were very likely to participate in all forms of activities we studied, from remixing to friending. The percentage of users in this class stayed steady from January (8.4%) to February (8.6%) and increased into March (11.6%), perhaps because of the decrease in number of classes in the model. A class of downloaders+commenters also continued across all three months and remained relatively consistent (from 15.1% in January, 12.5% in February, and 11.4% in March), and the low class of users who were only likely to post one project also remained present (increasing from 43.9% to 70.1% to 77.1% in each subsequent month). This supports the idea that clicking love-it or favorite on projects or sending friend requests are high-end activities among a relatively small group of users compared to practices like downloading projects and commenting on projects. It also suggests that users who were active in one month (our sample was drawn from active users) are likely to drop off in activity in subsequent months.
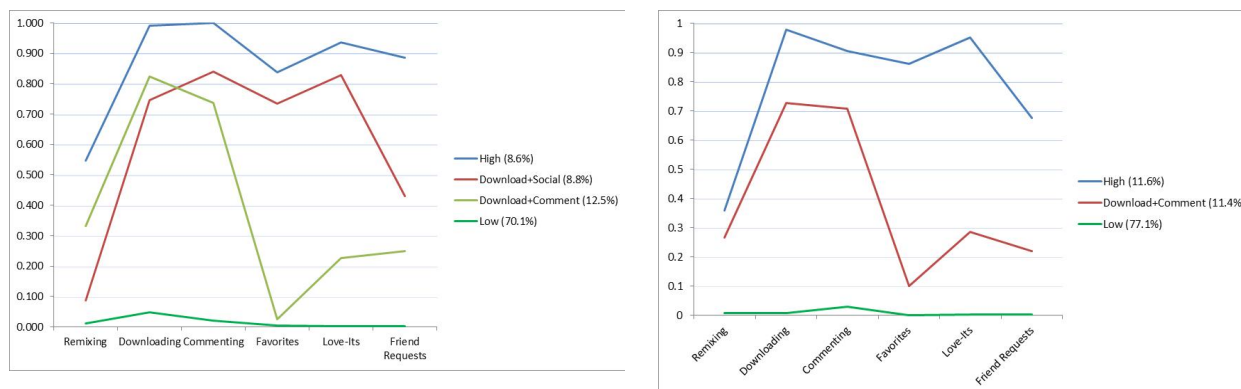

Figure 2: Four Class Model for February (left) and Three Class Model for March (right).

Absence of Gender Differences in Participation

Girls only represent one-third of all registered members on the Scratch site. The gender distribution in our overall sample reflected the distribution of self-reported gender on the Scratch site: 67% male, 33% female. Given this prior knowledge about differential membership in Scratch community, we tested whether gender was proportionately represented in each of the latent class. The distribution of gender within each class model was generally insignificant with only two exceptions: a higher proportion of girls in the high class in the month of January and a higher proportion of boys in the download+comment class in the month of March (for more detail on statistical analyses, see Appendix A). These analyses suggest that while males dominate the population of Scratch at large, within class profiles gender differences are minimal, an interesting finding for a youth amateur design site focused on programming.

Discussion

This paper examined broad quantitative trends of participation in a youth amateur design site focused on programming, making visible sets of practices that support design on a collective scale. Prior qualitative studies have documented that many of the activities on the Scratch site are collaborative in nature and support youths' programming designs though social supports (love-its, favorites, friend requests, comments) and constructive criticism as well as through opening up youths' designs to each other through the opportunity to play, download, and remix others' projects (e.g. Brennan, 2011; Roque et al, 2012; Aragon et al, 2009; Kafai et al, 2012). Before this

study we had little idea of what kinds of participation patterns users exhibited in the variety of collaborative design practices common on Scratch. Our findings suggest that there are several classes of users in Scratch and that making a project and downloading others' projects are gateways for other more prominently social activities. This is a surprising finding given our own prior assumptions that clicking "love-it" or "favorite" on a project were the lowest bars of activity. Indeed only a small percentage of the active Scratch population engaged in those activities: those already engaged in creating, downloading, and commenting on projects. Instead, active Scratch users prioritize designing projects and downloading others' designs, a finding that suggests studying social networking forums focused on design may be qualitatively different from the social networking sites discussed more prominently in popular media (e.g. Facebook, MySpace) that center on relationships.

Our results suggest several future steps for deeper analyses. First, the unavailability of data such as views of projects, home page, and notifications make the activities of a large proportion of users (55%) in the study invisible. The development of Scratch 2.0 is designed to capture these kinds of data, which should help illuminate the activities of this hidden group of participants. In addition, we need better ways of documenting the relative richness and sophistication of projects and comments at a quantitative scale. This will allow us to differentiate between types of project creators and comment posters, for instance between users who post many relatively simple projects or users who share one or two highly sophisticated or complex projects. Further, although the current class models of participation we presented here are static, in future analyses we will investigate which users transition between classes in each month's model, seeing what proportion of high-class (and other) users remain in that class or move to other classes. We also plan to investigate the role of experience in Scratch in seeing which users move between classes as well as analyzing whether other users' activities (clicking love-it on a project, commenting on a project, etc.) have an effect on users moving between classes. These analyses would provide richer information on who is participating in Scratch and what kind of users' activities may influence the likelihood of their changing their involvement on the site. Establishing models of participation in Scratch was an important precursor to these other analyses.

The larger goal of this research is to illuminate collaborative practices in massive communities that support learning and design, to see who is participating in those activities, and to evaluate how to sustain those types of activities. In the end we want to engineer online sites to more productively encourage computer supported collaborative learning on a collective level. Analyses presented in this paper are a first step toward a broader understanding of who participates in which activities over time in Scratch. Future interventions at a website-design level and at various local levels will be able to more intentionally target certain classes of users to involve them in more collaborative kinds of work. Some models for this already exist albeit in a small scale, for instance supporting highly involved designers to learn to post more positive, constructive criticism or helping other users connect to each other for collaborative co-design through semi-structured design challenges (see Roque, Kafai & Fields, 2012; Kafai, Fields, Roque, Burke, & Monroy-Hernandez, 2012). Yet these types of interventions reach out primarily to the "high class" of users who already engage in most opportunities on the site. Analyses in this paper suggest opportunities to target other classes of users, for instance reaching out to those in a "downloading" class to engage them in one additional collaborative practice such as commenting. The importance of design to this community also suggests that interventions should be targeted on youths' designs, perhaps by drawing attention to new users' designs who might benefit from feedback or making designs more easily searchable so that users can connect with others who have common interests. More broadly, further research is needed on site-wide analyses of other youth amateur design sites to see if similar trends prevail.

Appendix A

The LCA analysis resulted in a 5-class in January, a 4-class in February, and a 3-class model in March based on the indicators for each model fit (Hagenaars & McCutcheon, 2002; Muthen, 2002). For each participant, LCA generates probabilities for membership into each class, and generally one class has the highest probability of members. For instance, results show the students classified into their highest probability latent class had a 76-94% probability of being in that class, and a 0-24% of being in the other classes. Among the latter, 2 of the 20 classification were above .19. Taken together, these results support a 5-class model. LCA results for February indicate that a 4-class model provides the best fit for the data (see Figure 2). This is supported by a non-significant LMR p-value at the 5-class model (p= .08). In addition, aBIC value dipped to its lowest point at the 4-class model. Further, the substantive interpretation of the four model provided greater clarity. Additional statistical support can be provided by the average most likely latent class membership probability. For the 4-class model, students classified into their highest probability latent class had a 75–97% probability for being in that class, and a 0 to 21% of being in the other class (with only 2 of the 12 other class probabilities being above 10%). LCA results for March indicate that a 3-class model provides the best fit for the data (see Figure 2). This mainly supported by lowest BIC and aBIC values for

that class. In addition, the most likely class probabilities ranged from 85.5% to 98.0%, and 0 to 8.7% for the other classes. Interpretation of the 3-class model provided a better, more parsimonious interpretation of the data. A 4 class model was substantively rejected because one class consisted of a small number of users (1.4%) and it was only distinguished to another class by 1 of the 6 indicator items. Due to space constraints we have not put in the model fit indices table as we did for January.

Table 1: Model Fit Indices for January.

DICH	**January:** N = 2225 (Includes project creators for any of the 3 months)						
	likelihood	free par	BIC	aBIC	LMR p-value	Entropy	AIC
1	-6456.55	6	12957.97	12938.91	N/A	N/A	12925.10
2	-5527.54	13	11152.31	11111.01	0.0000	0.828	11081.09
3	-5477.94	20	11105.44	11105.44	0.0176	0.681	10995.87
4	-5448.24	27	11098.41	11012.63	0.0000	0.818	10950.48
5	-5426.66	34	11107.59	10999.58	0.0000	0.804	10921.31
6	-5416.54	41	11139.71	11009.46	0.0466	0.773	10915.08

Note. The Lo-Mendel-Rubin (LMR) likelihood ratio tests whether the current model is an improvement over the model with one less class. For instance, the p-values indicate the 5-class model provides the best fit. The non-significant p-value at the sixth class (.1129) suggests that it is not an improvement over a 5-class model. In addition, information criteria (i.e., Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), sample size Adjusted BIC (aBIC)) were used to compare models, wherein models with additional model parameters were penalized in the search for the most parsimonious model and information criteria with the lowest values indicate the best model fit. As values begin to level off (especially if there values do not increase), substantive interpretation and selection criteria take a greater role. For this analysis, the lowest value is at the 5-class model, wherein BIC and aBIC values begin to increase. In addition, a 5-class model provides a more meaningful and distinct interpretation of the class than models with fewer or great classes.

For assessing the distribution of gender, we performed multiple chi-square tests for independence analyses. For January, 2 (gender) x 5 (latent classes) chi-square analyses initially revealed that gender was distributed differently across the 5 latent classes, χ^2 (4) = 9.635, p = .047. However, upon close inspection of standardized residual scores comparing difference between the observed and expected, only 1 marginally significant difference emerged (z > 1.96). That is, there were more women in the class 1 (high class) than expected, z = 2.030. For February, chi-square test for independence suggest that each latent classes had a similar proportion of boys and girls, χ^2 (3) = 5.613, p = .132. For March, chi-square test for independence initially revealed that gender was distributed differently across the 3 latent classes, χ^2 (2) = 10.040, p = .007. However, standardized residual comparisons revealed showed only one significant difference within Class 1 (download+comment), z = -2.152; that is, there were fewer girls found in this group than expected.

References

Aragon, C., Poon, S., Monroy-Hernandez, A., & Aragon, D. (2009). A tale of two online communities: Fostering collaboration and creativity in scientists and children. *In Proceedings of the Creativity and Cognition Conference* (pp.9–18). Berkeley, CA. New York: ACM Press.

Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom.* New Haven and London: Yale University Press.

Birchfield, D. & Megowan-Romanowicz, C. (2009) Earth science learning in SMALLab: A design experiment for mixed reality. *International Journal of Computer Supported Collaborative Learning*, 4 (4), 403- 421.

Black, R.W. (2008). *Adolescents and online fiction*. New York: Peter Lang.

Brennan, K. (2011). Mind the gap: Differences between the aspirational and the actual in an online community of learners. *Proceedings of the International Conference of Computer Supported Collaborative Learning*.

Chen, M. (2012). *Leet noobs: The life and death of an expert player group in World of Warcraft.* New York: Peter Lang.

Engelmann, T. & Hesse, F.W. (2010) How digital concept maps about the collaborators' knowledge and information influence computer-supported collaborative problem solving. ijcscl 5 (3), pp. 299-320.

Fields, D. A. & Kafai, Y. B. (2009). A connective ethnography of peer knowledge sharing and diffusion in a tween virtual world. *International Journal of Computer Supported Collaborative Learning,* 4(1), 47-68.

Giang, M. T., Kafai, Y. B., Fields, D. A., & Searle, K. A. (2012). Social interactions in virtual worlds: Patterns and participation of tween relationship play. In J.Fromme & A. Unger (Eds.), *Computer games and new media cultures: A handbook on the state and perspectives of digital games studies.* New York: Springer Verlag. 543-556.

Grimes, S. & Fields, D. (2012). *Kids online: A new research agenda for understanding social networking forums.* New York. The Joan Ganz Cooney Center at Sesame Workshop. Available online at http://www.joanganzcooneycenter.org/reports-38.html.

Hagenaars J, & McCutcheon A. (Eds). (2002). *Applied Latent Class Analysis*. Cambridge, UK: Cambridge University Press.

Kafai, Y. B., Fields, D. A., Roque, R., Burke, W. Q., & Monroy-Hernández, A. (2012). Collaborative agency in youth online and offline creative production in Scratch. *Research and Practice in Technology Enhanced Learning*, 7(2), 63-87.

Luther, K., Caine, K., Ziegler, K., & Bruckman, A. (2010). Why it works (when it works): Success factors in online creative collaboration. In *GROUP '10: Proceedings of the ACM Conference on Supporting Group Work* (pp. 1–10) New York: ACM Press.

Monroy-Hernandez, A. (2012). Designing for remixing: Supporting an online community of amateur creators. *Unpublished dissertation*. MIT.

Muthen B. 2002. Statistical and substantive checking in growth mixture modeling. Retrieved January 2007 from http://www.gseis.ucla.edu/faculty/muthen/full_paper_list.htm.

Muthen B, Muthen L. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. Alcohol Clinical Experimental Research, 24 (6), 882–891.

Nickerson, J.V., Monroy-Hernández, A. (2011). Appropriation and Creativity: User-initiated Contests in Scratch. In *Proceedings of the Hawaii International Conference on System Sciences (HICSS '11)*.

Resnick, M., Maloney, J., Hernández, A. M., Rusk, N., Eastmond, E., Brennan, K., Millner, A. D., Rosenbaum, E., Silver, J., Silverman, B., & Kafai, Y. B. (2009). Scratch: Programming for everyone. *Communications of the ACM, 52(11),* 60–67.

Rick, J. & Guzdial, M. (2006). Situating CoWeb: A scholarship of application. International Journal of Computer-Supported Collaborative Learning, 1(1), 89-115.

Roque, R., Kafai, Y. B., & Fields, D. A. (2012). From tools to communities: Designs to support online creative collaboration in Scratch. *Proceedings of IDC 2012,.* Bremen, Germany.

Steinkuehler, S., & Duncan, S. (2009). Scientific habits of mind in virtual worlds. *Journal of Science Education and Technology, 17(6),* 530–543.

van der Pol, J., Admiraal, W. & Simons, P. R. J. (2006). The affordance of anchored discussion for the collaborative processing of academic texts. *International Journal of Computer-Supported Collaborative Learning*, 1(3), 339-357.

Williams, D., Contractor, N., Poolec, M. S., Srivastad, J., & Cale, D. (2011). The Virtual Worlds Exploratorium: Using Large-Scale Data and Computational Techniques for Communication Research. *Communication Methods and Measures*, 5(2), 163-180.

Acknowledgments