Designing for Remixing: Supporting an Online Community of Amateur Creators

by

Andrés Monroy-Hernández

Thesis proposal submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Media Arts and Sciences

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

April 2011

Author_____

Program in Media Arts and Sciences
April 19, 2011

Certified by_____

Mitchel Resnick
LEGO Papert Professor of Learning Research
Program in Media Arts and Sciences
Thesis Supervisor

Designing for Remixing:
Supporting an Online Community of Amateur Creators

by

Andrés Monroy-Hernández

Thesis proposal submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
on April 19, 2011, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Media Arts and Sciences

Abstract

In this dissertation proposal I describe a framework to design and study an online community of amateur creators. I focus on remixing as the lens to understand the social, cultural and technical infrastructure of a social media environment that supports creative expression. I am motivated by three broad questions: 1) what are the *structural* properties of an online remixing community? 2) what is the *functional* role of remixing in cultural production and social learning? 3) what are amateur creators' *attitudes* towards remixing? As part of this work, I conceived, developed and studied the Scratch Online Community: a website where young people share and remix their own video games and animations, as well as those of their peers. In three years, the community has grown to close to 800,000 registered members and more than 1.7 million community-contributed projects.

Thesis Supervisor: Mitchel Resnick
Title: LEGO Papert Professor of Learning Research, Program in Media Arts and Sciences

Designing for Remixing:
Supporting an Online Community of Amateur Creators

by

Andrés Monroy-Hernández

The following people will serve as readers for this thesis:

Thesis Reader_____

Yochai Benkler
Berkman Professor of Entrepreneurial Legal Studies
Harvard University

Thesis Reader_____

Robert C. Miller
Associate Professor of Electrical Engineering and Computer Science
MIT Computer Science and Artificial Intelligence Laboratory

Designing for Remixing: Supporting an Online Community of Amateur Creators

Andrés Monroy-Hernández

ABSTRACT

In this dissertation proposal I describe a framework to design and study an online community of amateur creators. I focus on remixing as the lens to understand the social, cultural and technical infrastructure of a social media environment that supports creative expression. I am motivated by three broad questions: 1) what are the *structural* properties of an online remixing community? 2) what is the *functional* role of remixing in cultural production and social learning? 3) what are amateur creators' *attitudes* towards remixing? As part of this work, I conceived, developed and studied the Scratch Online Community: a website where young people share and remix their own video games and animations, as well as those of their peers. In three years, the community has grown to close to 800,000 registered members and more than 1.7 million community-contributed projects.

INTRODUCTION

In recent years, there has been an explosion of online communities where people share their ideas and creations. Participants of these communities not only share their own work, they also remix, re-create, remake, re-tweet, fork, sample, mash-up, edit or appropriate other people's work to produce derivatives or remixes that are shared back to the community. Previous work has described how these communities and the activities they support, have far-reaching implications for the way cultural and economic production work. In this dissertation proposal, I outline a research framework to examine the design and usage of an active online remixing community built to engage young people in social production and creative learning.

Online remixing communities span a wide range of creative endeavors, there are communities for sharing and remixing *code* (for instance, GitHub and Sourceforge), *articles* (for instance, Wikipedia and Wikia), *videos* (for instance, YouTube and Vimeo), *images* (for instance, Flickr and DeviantArt), *status updates* (for instance, Twitter and Facebook), *music* (for instance, ccMixter and IndabaMusic), and even the design of *physical objects* (for instance, Ponoko and Etsy). In this dissertation, I focus on the Scratch Online Community, a website developed to enable amateur creators, especially children aged between eight and sixteen, to share and remix animations and video games. I plan to, first, describe the motivations and socio technical infrastructure of the Scratch website, then examine how people have used the website to collaborate with others through remixing and how the design has influenced these activities.

Thesis proposal submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Media Arts and Sciences.

BACKGROUND

Remixes, mash-ups, forks and derivative works are terms used to refer to a form of social production based on the idea of creating something out of existing creations. Wikipedians editing one another's articles, Free and Open-Source programmers reusing existing code libraries, and teenagers mixing their favorite videos on YouTube, are all examples of extensive remixing practices found today in the digital landscape.

I build on existing work that places remixing in new forms of economic and cultural production (Benkler, 2006; Jenkins, 2006; Manovich, 2005; Sinnreich et al., 2009), and describes some challenges that remixing poses to our existing legal and moral assumptions (Lessig, 2008; Posner, 2007).

Furthermore, I build on work that advocates remixing as a new media literacy skill (Ito, 2010; Jenkins et al., 2009; Livingstone, 2008; Perkel, 2008) and that positions remixing as legitimate form of participation in social learning environments, building on learning philosophies such as Constructionism (Papert, 1980) and Situated Learning (Lave and Wenger, 1991).

Finally, I look at the implications of remixing for the design of social computing systems. I do this by connecting to research on human-computer interaction that examines the participation patterns and the design of peer-production communities that engage people in sharing and remixing videos (Diakopoulos et al., 2007; Shaw and Schmitz, 2006), images (Seneviratne et al., 2009), music (Cheliotis and Yew, 2009) and hypertext (Viégas et al., 2004).

Economic and Cultural Production

Network information technologies have enabled social production to emerge as a new form of economic production parallel to markets and firms. People engaging in this new form of production, called Commons-based Peer Production (Benkler, 2002), rely on the availability of existing common

resources that are often re-purposed for the creation of new products that go back to that communal source for others to reuse as well.

Culture has also been influenced by these new forms of peer-to-peer participation where the boundaries between producers and consumers blur, what Jenkins (2006) calls Participatory Culture.

These creative practices, based on the idea of building new things by combining existing ones, are not necessarily new. Artists, especially postmodernist, have engaged in similar practices through "appropriation art", "pastiche", "collage", "sampling" and "bricolage". Furthermore, folk culture and oral traditions rely on this idea of re-creating and remixing what others have made. For example, as Manovich (2005) argues, remixing is a part of cultural evolution as one can see how ancient Rome was a remix of ancient Greece. Despite this common feeling that everything old is new again, the influence of digital technologies in remixing-like practices is undeniable. These technologies have enabled people to a) create perfect copies and b) go beyond just being inspired by other's creations but to remix the original works themselves (Sinnreich et al., 2009).

Learning and New Media Literacy

Remixing is a social activity by nature. Hence, some discussions around remixing and learning stem from learning theories that look at the social context in which learning happens. Situated Learning is one of these models that has advocated legitimizing peripheral forms of participation, in particular in apprenticeship-like learning environments (Lave and Wenger, 1991). This apprenticeship model is also central to Constructionism (Papert, 1980), a learning philosophy based on the idea that some of the most valuable learning experiences occur when children engage in building personally meaningful objects in a social context. Building on this theory, Turkle and Papert (1990) argued for "bricolage" as a legitimate learning style, alternative to planning, where learners "construct theories by arranging and rearranging, by negotiating and re-negotiating with a set of well-known materials", a model that fits well with the process of remixing.

Similarly, Wenger (1998) stressed the importance of also having access to a "shared repertoire of communal resources" to help shape a "community of practice."

More recently, cultural anthropologists and media scholars have documented the ways young people engage in creative practices through remixing. For example, Ito (2007) described how children relate to media franchises, such as Pokemon. She found that children not only consume media but they engage in remixing and re-creating it, "demonstrating that children can master highly esoteric content, customization, connoisseurship, remixing." Livingstone (2008) has described similar practices concerning Internet use and presented the challenges and possibilities that these creative practices promise. Similarly, Jenkins (2006) has narrated how children have become active participants in media creation by remixing their favorite literary characters in Harry Potter and creating rich fan fiction. Later, Jenkins et al. (2009) argued that the ability to remix, is a core competency that children must possess to be fluent with new media.

Ethical and Legal Challenges

Creative appropriation or remixing has been confronted on moral and legal grounds. In his description of children's fan fiction, Jenkins (2006) describes how the young girl who created a Harry Potter fan fiction website, was challenged by the company who owns the legal rights to the book and the movie, on the grounds of copyright infringement. Eventually the company dropped the legal action and reached an agreement.

More recently, appropriation art has been under legal scrutiny (Greenberg, 1992; Landes, 2000). For example, a federal court judge determined that photographer Richard Price violated Patrick Cariou's copyrights for remixing a set of pictures by putting them on frames, "painting over some portions" (Batts, 2011).

Similarly, countless video remixes on the video-sharing website YouTube have been removed under the Digital Millenium Copyright Act, for being identified as remixes of commercial videos (Seneviratne and Monroy-Hernández, 2010).

Posner (2007) has articulated how plagiarism is highly context-dependent but that one can assess the ethics of appropriation by thinking through issues of deception, perception and social expectation. Benkler (2006) argues that if we want peer-production to flourish, we must figure out how to enable remixing: "If we are to make this culture our own, render it legible, and make it into a new platform for our needs and conversations today, we must find a way to cut, paste, and remix present culture. And it is precisely this freedom that most directly challenges the laws written for the twentieth-century technology, economy, and cultural practice."

Similarly, Lessig (2008) has reported on cases of "copyright extremism" that, he argues, "chill" innovation and creativity, especially among young people who often engage in these practices. In response to such copyright extremism, the Creative Commons licenses allow creators to have more control of their copyright and release their work under more permissive licenses that would foster amateur creativity.

At the core of examining the ethics of remixing lies the understanding of cooperation, that is, the idea of how much people are willing to sacrifice their selfish and rational desires, to obtain monetary or reputation gains, to behave in altruistic and cooperative ways. These questions lay at the core of literature on human cooperation which is beginning to be translated to social system design.

Social Computing System Design

Human-computer interaction research has studied the use and design of social computer systems that foster cooperative practices that allow remixing to take place. Wikipedia has been perhaps the most widely researched of those systems. For example, Viégas et al. (2004) developed a visual-

Figure 1. The home page of the Scratch website

ization of Wikipedia edits that led to insights into the nature of the system and its editors' ability to collaborate. Later Kittur and Kraut (2008) studied the quality of Wikipedia's articles in relationship to various types of coordination mechanisms. Similarly, analyses on open source software development have led to insights into the mechanisms that lead to successful cooperative projects. Raymond (1999), for example, argued that one of the lessons to be learned from open source software programmers, is the importance of knowing what to rewrite and reuse. He describes how Linus Trovalds (the creator of Linux) did not "try to write Linux from scratch" instead "he started by reusing code and ideas from Minix, a tiny Unix-like OS for PC clones."

Researchers have also developed web mash-up tools that allow people to remix web content (Bolin et al., 2005; Wong and Hong, 2007). A study on one of those web mash-up tools found that many of its users lacked programming skills and found that web mash-up are an effective way of searching and aggregating information (Zang and Rosson, 2008).

More specifically on online communities for remixing, Shaw and Schmitz (2006) developed a video remixing platform and studied the nature of the most generative video segments intending to understand how to integrate automatic recommendation systems with user-driven suggestions. Diakopoulos et al. (2007) analyzed user's participation in a video remixing website and documented how participants developed specific norms for appropriating other people's work that were not encoded in the architecture of the web site. Additionally, a study of the music remixing online community ccMixter.org, looked at the impact of a remixing contest in the community dynamics. The study found that the contests increased participation among new comers but that they did not continue using the website after the contest (Cheliotis and Yew, 2009).

THE SCRATCH ONLINE COMMUNITY

The empirical setting for this work is the Scratch Online Community, a website (Figure 1) I conceived and developed over the past four years in collaboration with others at the Lifelong Kindergarten research group, for this and other lines of research. The website allows anyone, especially young people between ages eight and sixteen, to share their animated stories, interactive art, and video games. Participants use the Scratch programming environment, a desktop application, to create or remix projects by putting together images, music and sounds with visual programming command blocks (Resnick et al., 2009). Projects range from interactive greeting cards, physics simulations, animations of popular songs to homemade video games, just to name a few.

Scratch projects are organized in sprites (e.g. a character in a game). Each sprite has a set of "costumes" or images that represent its different visual states, for example, a sprite of a bird flying could have multiple costumes, each one representing the different positions of the wings. Sprites can also have sounds associated to them, these sounds can be either recorded with the microphone or imported from the hard creator's hard drive. Finally, sprites' behavior is controlled by "scripts" which are stacks of visual programming blocks.

In my dissertation, I plan to document in detail the motivations that led to the design of the Scratch website as it is now and the various iterations that it went through as a result of internal and user-driven demands and participation patterns. I expand on the tight relationship between the technical capabilities of the website and the social dynamics that it supported or intended to support. I take a critical look at how the original goals of the community were or were not achieved and in what ways that happened. I narrate scalability challenges with the technology and moderation of the community. In the following two sections I provide a glimpse of this.

More broadly, I try to tackle questions such as: How was the culture of sharing seeded and maintained in a public space? What were some important incidents that changed policies or architecture? How did the architecture and management of the site influence the culture of the site and how did users help shap this? Was it the design of the space or the work of users to create the culture? What were the lessons learned along the way by administrators? What kinds of learning outcomes were achieved? To answer some of these questions, I primarily rely on participant observation data, experiences collected during the three years and I support these arguments with descriptive statistics and case studies.

Motivations

From its start (Monroy-Hernández, 2007; Monroy-Hernández and Resnick, 2008), I set the goal for the Scratch Online Community as to give creators access to 1) a *network of peers* that functions as an audience and as potential collaborators and, 2) a *repository* of inspirational creations that can be creatively appropriated by anyone.

More broadly, the Scratch Online Community was created with the idea of supporting a Community of Practice around Scratch where novices and experts would come together, in the spirit of Papert's Samba school's metaphor (Papert, 1980).

Additionally, the idea of the Scratch Online Community was conceived under the umbrella of embodying the ethos of Participatory Culture of empowering people to become producers rather than just consumers of media.

Last, the Scratch Online Community was motivated to support the various iterative stages of the creative process (Resnick, 2008), namely: Supporting creators' *imagination* by giving access to a pool of inspirational projects and ideas; supporting *creation* by allowing people to reuse and remix; supporting *play* by letting people interact with others and their creations within a community; supporting *sharing* by allowing people to easily upload their creations to the platform; supporting *reflection* by providing a space to receive comments and discussion forums for more in-depth discussions.

Sociotechnical Infrastructure

The Scratch website platform, called ScratchR, is broadly composed of the following components:

1. A repository of projects and metadata. Projects can be downloaded by anyone, modified and re-uploaded to the website as a remix. Each project has its web page where it is displayed and where people can interact with it and other people.

2. A social network consisting of profile pages and unidirectional friendship connections. Profile pages list the friends, projects and "favorited" projects for each user with his or her avatar image and the Country of origin (all self-reported data).

3. Social features for interacting with people's creation such as commenting, tagging, "loving", "favoriting".

4. Galleries, which are pages where people can group projects and talk about them. It is important to note that galleries have been repurposed by the community as group spaces where people collaborate to create projects or use it as a space to talk to one another or play Role Playing Games.

5. Discussion forums where community members help one another with technical problems, find collaborators and talk about non-Scratch related activities that foster a sense of community on the website.

6. External services supported by an API[1] such as a website where people can link projects or a Wiki where people can document their experiences with Scratch and the community.

The website runs on a hybrid model for moderation that combines user-driven moderation through flagging and appointed moderators working in parallel with a full-time staff member and other part-time ones that review the flags and ensure that the social dynamics are kept as civil as possible.

[1]API stands for Application Program Interface. In ScratchR, they are a set of web-accessible functions that let people retrieve and submit data such as login authentication and information about projects.

This model has allowed for scaling the community management at a relatively low cost, however, much of the architecture and software development during the three years has been put into mechanisms for preventing antisocial behavior.

Three years after its official release, the Scratch Online Community website I developed, handles more than ten million page views and six hundred-thousand people monthly. This web traffic is more than half the page views of websites like newsweek.com[2]. As of April of 2011, more than 1.7 million projects have been uploaded at an average of 1 MB per project. Every second, the website receives up to 180 requests and it transfers 4MB.

To handle this level of activity, ScratchR, the website's underlying platform, uses a caching engine for static content called Varnish and another for dynamic content and database queries called Memcached. ScratchR runs on a completely Free and Open Source software stack that includes Apache for its web server and MySQL for its database running on CentOS Linux. ScratchR is written using a PHP-based framework called CakePHP. Additionally, ScratchR supports external web applications such as a discussion forum, a user-driven Wiki, a statistics visualization website, a Scratch sprites sharing website, and a few other websites that provide additional support to community members. These extra websites are supported through an API that has allowed scalability.

RESEARCH

I propose a framework to examine remixing in an online community of amateur creators. I approach this by first studying the design of the online community as a remixing system, and then analyzing what people do and how they react to what others do. More specifically, I focus on the structural, functional and attitudinal characteristics of an online community's sociotechnical infrastructure and its participants activities. This framework derives from and is examined through design interventions, three-years of participant observation data, case studies, interviews with community members, quantitative and network analysis of a large corpus of data that includes more than 700,000 registered accounts and a repository of more than 9 million comments and 1.6 million interactive media objects, 30% of which are remixes.

Structure of a Remixing System

I plan to study the sociotechnical architecture of the Scratch Online Community by examining its following structural attributes (Figure 2): 1) granularity of the remixable components, 2) modularity of the remixable components, 3) decomposability of finished projects, 4) attributability mechanisms, and 5) openness to remix across systems.

In this proposal I briefly define each structural attribute using one or two examples, and I explain how I am planning to go about studying it. I plan to analyze these structural properties of the system using varied approaches including: 1) case studies that give a rich description of various scenarios that explain the influence of each attribute, 2) analyses of

[2]04/2011 data from http://www.quantcast.com/newsweek.com

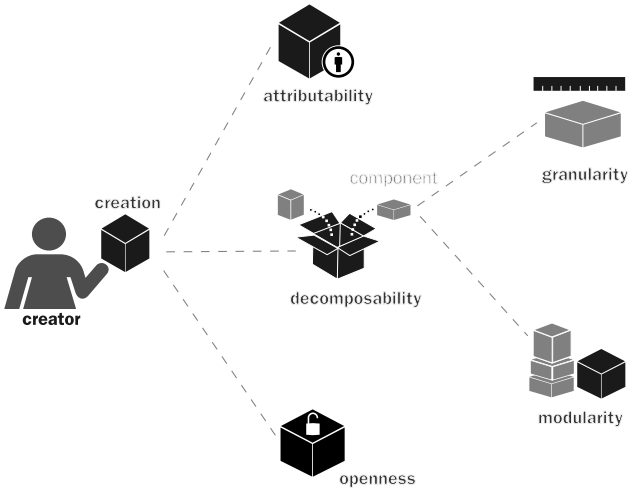

Figure 2. Structural dimensions of a remixing system (some icons taken from creativecommons.org, openclipart.org and thenounproject.com).

the data corpus to understand the frequency of the scenarios and the relationships among the structural attributes, 3) natural design experiments to study the effect of varying some structural dimensions.

Granularity

Granularity is the size of project's modules (Benkler, 2002). In Scratch, projects have several levels of granularity. Projects are made of "sprites" such as characters in a game or elements in the user interface. Each sprite can have "scripts" or stacks of programming blocks that control the sprite's behavior such as its position on the screen, its look, sound and interaction with other sprites and the user (by mouse, keyboard, or other sensors). Each sprite also has one or more costumes or images that represent the various visual states of a sprite. Sprites can also have sounds that are played programmatically; for example, a character in a game could make a sound each time it jumps.

Proposed Work. The Scratch Online Community, by default, only allows sharing at the coarsest degree of granularity: only projects can be shared. I plan to analyze the implications of this design decision and the ways people get around the limitations of the architecure.

I have anecdotal evidence that Scratch participants get around the granularity limitations by sharing full projects with the sole intention of sharing a single sprite, script, image or sound. This need for finer granularity often stems from a desire to engage in sharing practices that the original design did not anticipate. I plan to look at these practices in more detail to understand how a finer granularity mechanism would help support various types of creative and collaborative learning practices beyond the existing research that suggests that finer granularity is correlated with the number of people engaged in cooperative activities.

To analyze the effect of explicitly supporting finer degree of granularity through a natural experiment, I plan to analyze the adoption of "Scratch Resources" [3], a website created by members of the Scratch community to support sharing sprites, images and sounds.

Some of my motivating questions are: How common is sharing and remixing across the various levels of granularity available (even if not explicitly supported)? What are the types of participants and motivations for sharing finer grained components? What role does granularity play in the likelihood of a project being remixed controlling for all other factors? How do various levels of granularity affect the type and quality of remixes? How do various levels of granularity support various levels of familiarity with Scratch – are novices more likely to rely on coarser granularity when engaging in remixing as a scaffolding mechanism?

Modularity

Benkler (2002) defines modularity as the "property of a project referring to the extent to which it can be broken down into smaller components, or modules, that can be independently and asynchronously produced before they are assembled into a whole." For this work, I separate two aspects of modularity: 1. The ease of *integrating* such components into new creations or remixes. 2. The ease of *decomposing* an existing project into smaller components, typically for remixing. In this section I plan to analyze the term "modularity" by focusing on the first aspect. In the next section, I focus on the second aspect under the term "decomposability".

Proposed Work. Some components or projects can be more easily integrated into new projects. One of the questions I am interested in exploring is what makes some components more modular than others.

For example, an image generated programmatically might be harder to integrate than one in bitmap format. A module that represents a cultural icon, such as Mario[4], is perhaps easier to integrate into other projects than an image of a less well-known character. However, there are situations when new subcultural icons emerge within the Scratch community. For example, a community member from New Zealand created a character called Maki-Tak. Maki-Tak became so popular in the community that other people started to create projects that included Maki-Tak itself of that character or other characters inspired by it, often called Takis. Situations exist where some components are remixed despite their internal complexity, which might indicate that these could be built in a way that their internal complexity is hidden and remixing them is easy. For example there is a sprite created by an advanced Scratch user that represented a physics simulation of a string. This project was later remixed in a project where it represented a necklace.

I plan to operationalize the assessment of modularity by measuring its adoption through a number of remixes. The assumption is that more modular components are remixed more. I examine this in two types of components: the sample sprites,

[3] Available at http://resources.scratchr.org

[4] Mario is a character in a popular video game called Mario Bros. from Nintendo Inc.

images and sounds that come with Scratch, and the ones that members of the community have created.

The type of questions I will to answer are: 1. What technical or cultural attributes are linked with component modularity? 2. Are modular components used more often by newcomers, and do they provide scaffolding in their learning of Scratch programming? 3. Are community-generated components more often created by advanced users?

Decomposability

Building on the concept of modularity examined before, in this section I examine decomposability as the ease of decomposing a compiled project. Decomposability is the ease of breaking something apart for remixing. Therefore decomposability depends on the internal complexity of a project, which in itself depends on the expertise of the person attempting to decompose the project. For example, one can argue that images in Flickr.com are harder to decompose than those in OpenClipart.org, which provides the source vectors, or Aviary.com which provides the bitmap images that were used to create an image. The same happens with software applications that provide the source code in contrast with those that only provide the final compiled executable. But even if the source code is provided, there are some cases where projects are "impenetrable" because of their complexity or a mismatch between the expertise of the person trying to decompose a project and the complexity of the project. For example, for a novice programmer having access to the source code of the Linux kernel might not allow for easy decomposability.

In Scratch, all the sources of a project are provided so the decomposability of a project depends, among other things, more on how interconnected its various components are. For example, sometimes sprites "broadcast" messages back and forth making their decomposability much harder than those that are self-contained. Also some creators add instructions to their projects explaining how they can be broken apart, while others obfuscate their code to prevent remixing.

Proposed Work. I first do a manual analysis of a sample of projects to observe patterns of decomposability. Using the findings of this analysis I devise mechanisms to automate the evaluation of decomposability. For example, a decomposability metric could depend on the use of particular blocks (the "broadcast" block could reduce its ranking, while comments in the code would increase it), explicit obfuscation or matching between the typical expertise level of community members and the expertise required to understand a project.

I also look at remixes to analyze how different they are from their original project. The assumption would be that "impenetrable" projects would be correlated with no remixing or to superficial remixes, such as slight changes in the images, while highly decomposable projects are associated with significant differences between the remix and the original.

I also analyze the practices of code obfuscation and the strategies people use to discourage decomposability of their projects.

Attributability

Creators often want to get credit for their work. For example, the Creative Commons license originally had attribution as one of the options of their licenses, but after analyzing several years of usage of the licenses they found that few people waived the attribution clause. This led them to include attribution by default in all their licenses (and create a separate license for completely public domain works) (Brown, 2004). To further our understanding of attribution when designing a remixing system it is important to know the role attribution plays in supporting cooperative behavior among members of a remixing online community.

In Scratch, we have run some design experiments playing with attribution. For example, we found evidence that suggests that about 20% of creators object to seeing their projects remixed (Hill et al., 2010). We also have anecdotal evidence that people, who objected to remixing of their project, referred to a lack of credit as one of the problems. To address this, I added a mechanism that would automatically give credit to the creator of the original project whenever a remix was uploaded (for example, "Shared by John, based on Mary's project"). Using that design intervention as a natural experiment, we found evidence that suggests people not only want to get credit but that they prefer the credit given by another person over the automatic attribution given by the system Monroy-Hernández et al. (2011).

Proposed Work. I plan to extend this work with additional experiments. For example, one of the common complaints against remixing is that the textual description of projects often gets copied from an original project onto its remixes. This often gives the wrong impression that those notes were by the remixer rather than by the original creator. I plan to add a mechanism to show a distinction between the notes by the originator and those added by the remixer. These additional notes come with a mechanism to encourage remixers to explain how the original project was changed. This will serve as a natural experiment to test the additional value of not only giving binary representation of attribution but also give a qualifier explanation to the connection between a remix and the original project.

Openness

In this section I focus on analyzing a system's openness from the perspective of its norms and interactions with other systems. The Scratch Online Community allows any of its members to download and remix any project, in that sense Scratch is more open than similar websites such as Kongregate.com or Newgrounds.com, which do not provide either the license, the environment, nor the technical features to freely open its content. The ethos of openness is present in Scratch through its terms of use, the license used by all projects shared, and the community moderation styled enforced by administrators. However, this openness is not always understood or embraced by its members. Also this openness is not always compatible with other systems. For example, there are a some members of the community, who even after discussions with the administrators of the site, do not see enough value of openly sharing their work and de-

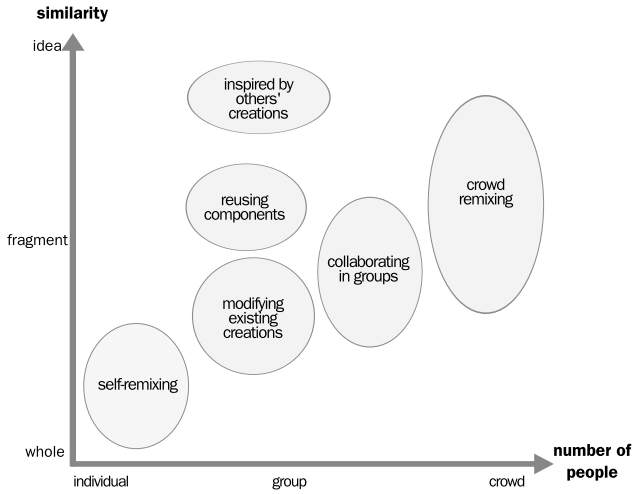

Figure 3. Functional roles of remixing in the social and product-focus dimensions

cide to stop using the website. Additionally, there are other members of the community who encounter conflicts when remixing content from other online spaces that are not compatible with Scratch's openness. These conflicts often involve art-sharing websites like DeviantArt.com, from where Scratch creators sometimes get images for their projects.

Proposed Work. I plan to investigate the ways young people in Scratch embrace, understand and reject the way openness is interpreted in Scratch. I try to examine whether young people see a value in openness and how a system is or is not able to surface these values through sociotechnical design. I present a set of case studies that show the range of ways people embrace and reject openness and how the social and technical design of the website influence them.

Functional Role of Remixing

The previous section helped describe the structural properties of the Scratch Online Community as a remixing system. This section is intended to examine what people do. In particular, the goal is to understand what are the various roles that remixing plays and how it is part of collaborative and creative practices.

Remixing represents many activities in Scratch. For example, people occasionally remix to fix software bugs on someone's project; in other cases people want to add to an existing project to customize it or make it more personal; in others, people want to start a trend or "meme" so they create a "template" project just so others remix it, add to it and pass it on.

To map these varied types of remixing, based on previous analysis of other remixing online communities (Seneviratne and Monroy-Hernández, 2010), I focus on two dimensions:

1. Similarity. This indicates how different a remix is from the original work on which it is based. Remixes can range from merely being inspired by someone's idea, to tweaking someone's project, to making an exact copy. This is an important aspect of remixing because it gets at the core of some of the tensions around remixing, and it also helps us understand the degree of interaction and creative output between a remix and its antecedent project.

2. Number of people. This indicates the number people involved or intended to be involved in the remix. Remixes can range from a single author (for example, a creator remixing herself as a form of version control), to a small group of people (for example, a team collaborating in making projects), to a crowd (for example, a large group of people participating in a remix chain). The number of people involved is a useful dimension because it helps us identify the distinct roles remixing plays in Scratch.

Using these two dimensions, one can identify a set of remixing categories (Figure 3) useful for the analysis of the functional roles of remixing in Scratch. Broadly speaking, the types of remixing I plan to use to guide my analyses are the following:

Incremental Remixing

This type of remixing often involves downloading someone's project to customize it or to fix a bug. For example, replacing the costumes of a sprite in a game for a different one. This form of remixing often occurs when people make small modifications to the sample projects that come preinstalled with Scratch. Also anecdotal evidence suggests that this form of remixing tends to be a useful way for newcomers to get started with Scratch, modifying an existing project instead of creating a completely new project. This type of remixing sometimes leads to controversy when the changes are subtle or nonexistent. Project creators sometimes argue that types of remixing in which the changes are minimal should not be allowed.

Component Remixing

Remixing also occurs when people use pieces of others' projects to produce something new, rather than building on top of existing work. In these cases, often one cannot quickly tell in what way the remix and the original are related. This type of remixing typically involves some sample sprites that come preinstalled with Scratch, such as the "jetpack girl" example mentioned above or some templates, images or sounds that members of the community created for others to reuse.

Group Remixing

Remixing is also used by groups or "collabs", as they are often referred to by Scratch members, to work collaboratively by remixing one another's work almost as a form of version control (Tichy, 1985). For example, we found in that one of the "collabs" we analyzed, one member of the group would start by creating a first version of the project they had decided to make while others would work on varied parts such as the images, sounds or the engine of a game (Aragon et al., 2009). In that particular "collab" there were, on average, seventeen remixes before officially releasing a project. In this case, remixing is part of an explicit collaborative activity.

Crowd Remixing

The home page of the Scratch website has a section called "What the Community is Remixing" that features the three top remixed projects in the past two weeks. Although these projects could represent any type of remixing, they are typically projects that consist of a compelling template that encourages people to add something small to it and pass it on to someone else, like a chain. For example, one of those projects titled "Add yourself to the party" shows a character dancing on a colorful dance floor by itself. The creator leaves a note in the description of the project inviting people to download the project and add a sprite representing their avatar or profile picture. Other examples are "Remix if you care about animal rights", representing a type of crowd remixing with a social cause, and "Coloring contest", representing popular genre among more artistic-inclined community members that consists of downloading a still image (often with music playing in the background), coloring it and submitting it to a contest for the best-colored image. Although typically the intention of the creator of this type of project is to create a chain, often the structure of the remix network looks more like a star because of people not following the rules or not checking where the last element of the chain is. These types of remixes are often incremental like the category described above, but the relationship between the creator and its intended audience is significantly different. Crowd remixing explicitly invites people to remix en masse following a specific template. The purpose of the creation is not the creation itself but the collection of many remixes created by many individuals. Evidence suggests that this type of crowd remixing emerged after the top remixed projects section was added to the website, but more research is needed to understand these phenomena.

Inspirational Remixing

On the far end of the similarity spectrum are projects that are inspired by other projects that do not use any particular component of the original work. These projects are often trying to replicate popular software, video games or even genres that have emerged in the community. For example, some people create remixes of the Microsoft Windows operating system, or the iPhone, or operating systems that other members of the community have created such as SynOS, a popular Scratch-based pseudo-operating system created by a teenage member of the community. Other examples of inspirational remixes are the "fan art" projects created after popular characters created by other members of the Scratch community such as "Maki-Tak", a lizard that dances and sings in animations created by a teenage girl, or "Mr. Happy Man", a grumpy character invented by a teenage boy. This type of inspirational remixing is a lot harder to identify automatically because the original and the remix do not share any bytes but their names, descriptions and log data might help.

Self-remixing

About 13% of remixes are projects created based on the same person's projects. This is typically used as some form of version control. For example, sometimes children create a Scratch project at school, upload it to the website, download it at home, continue working on it, then re-upload it to the website under a name like "My video game v2". When this happens, the website identifies them as remixes and links them back to their previous version. A few people, mainly popular Scratch creators who care about the projects displayed on their profile page, have two accounts: one for testing and one for sharing final version of their projects. For example, there is an adult member of the community who has two usernames: "Paddle2See" and "Paddle2SeeTest". As the name implies, one is often used for sharing drafts or work in progress.

Proposed work

In this section of my dissertation I plan to describe the way people engage in remixing on the Scratch Online Community. In particular, I plan to answer questions such as: 1. Are the types of remixing described above representative for Scratch and generalizable for other online remixing communities? For this, I plan to connect with other researchers investigating other online communities, especially ccMixter.org. 2. What kind of people engage in each type of remixing? For example, do novices engage in incremental and crowd remixing more often than component remixing? Are the originators of crowd remixing longtime members of the community and how do they manage to get more people to remix their work? 3. In what way does each type of remixing help scaffold Scratch participation? For example, are some types of remixing better at learning the technical aspects of Scratch while others are more helpful in learning how to collaborate and socialize online? Are people who engage in remixing initially more likely to stay active on the community and in what ways? 4. How have these types of remixing evolved over the course of three years? Are the social networks created through various types of remixing different in their structure and quality? 5. How do the projects in each type of remixing differ technically and in content? For example, in what ways might some types be more complex than others? The distinctions amongst the types are subjective at this point, and part of the work I plan to do is to clearly identify the boundaries among categories.

Attitudes Toward Remixing

I investigate remixers' and originators' attitudes toward remixing. In particular, I analyze how participants perceive remixing and how they do or do not cooperate by letting or encouraging others to reuse their work. This analysis is driven by an interest in understanding how the system design may influence these attitudes. I plan to analyze these issues from the perspective of people whose projects are remixed as well as from those creating the remixes.

Previous work studying people's attitudes toward remixing in the Scratch Online Community (Hill et al., 2010; Monroy-Hernández et al., 2011) found evidence that people are as likely to react negatively as they are to react positively when someone remixes their work. Additionally, we have found originators are more likely to respond negatively when their projects are more complex.

Broadly speaking, people whose projects are remixed react either by being indifferent to it, accepting it, condition-

ally accepting it (for example, specific rules or norms have to be followed), or by explicitly opposing to it (for example, posting a negative reaction comment like "you stole my project!"). Similarly, remixers go about remixing by either being oblivious of the norms (for example, giving credit or asking for permission has emerged as a norm in the community), or cautiously doing it by asking for permission first, or even being confrontational and using remixing as a form of "trolling" (Donath, 1998).

Proposed Work

I plan to analyze people's attitudes toward remixing in the Scratch Online Community through case studies and experiments. Additionally, I expect these metrics of people's responses and attitudes will serve to understand the health of the community and to motivate further design interventions. In particular, I propose two studies for studying Scratch participant's attitudes toward remixing that complement the two recently published articles (Monroy-Hernández et al., 2011; Hill et al., 2010).

Plasticity of Virtue. For the past five months, I have been running an experiment aimed at ascertaining how people's attitudes toward remixing could be changed through a design intervention. The study consists of sending notification messages that try to appeal to various motivations for cooperating in the Scratch online community. As mentioned before, a contentious issue in the Scratch community has been the acceptance of remixing by those whose projects are remixed. People only learn about remixes of their projects by browsing their project pages and looking at the list of derivative works. The notification page is the principal form of communication between the system and the users. The experiment consists of informing people when one of their projects gets remixed. This experiment provides an opportunity to test: 1) the comparative effectiveness of various ways of communicating the same event and 2) the permanency of the behavioral change, if any.

The experiment consists of two phases. The first is the notification period. When someone's project gets remixed, the system automatically assigns the creator of that project to one of the treatment conditions. From then on, the person receives messages whenever someone remixes his or her project. The message depends on the treatment category. For example some messages are neutral (for example, "Your project FooBar was remixed. Check out the remix."); others are positive (for example,"Congratulations! Your project FooBar was remixed. Check out the remix ."); others try to elicit generosity (for example, "Your project FooBar has been remixed. Sharing your work is a generous thing to do and a great thing for the Scratch community! Check out the remix."). Other messages try to elicit conformity, reputation building and fairness. Finally, a control category is added where no notification is sent.

The second phase of the experiment is sending no notifications. In this phase of the experiment, the notifications stop and people's reactions continue to be logged.

After a few weeks we look at the results of pre- and post-phase one, to analyze the effectiveness of each treatment by measuring the number of positive and negative reactions that originators leave on the remixes. Additionally, I analyze the pre- and post-phase two, to see how malleable the behavior is.

Featuring Top Remixes. The home page of the Scratch website now has a section called "What the Community is Remixing" that features the three top remixed projects in the past two weeks. This section did not exist when the website was first unveiled. In fact, this section was added as measure to counter the backlash against remixing by showing that getting one's project remixed could increase social status in the community (being on the front page is an important reputation-building mechanism). This study aims at assessing whether a design intervention aimed at increasing the acceptance of remixing had a "crowding out" effect by decreasing the quality or effort people put into making their remixes. I have devised metrics to operationalize the complexity and effort a creator puts into creating his or her project. These metrics are based on the number of sprites, scripts, blocks, diversity of blocks and time it took from the first save to the hard disk until it gets shared on the website.

CONTRIBUTIONS

Among the contributions of this thesis are:

- Conceptualization, design and implementation of a large online community of amateur creators.
- Collection of a large corpus of research data that includes millions of interactive media projects and activity logs of more than half-a-million accounts.[5]
- A set of studies of sociotechnical design interventions.
- A set of descriptive case studies of the activity on the Scratch Online Community.
- A multidisciplinary framework to analyze and understand a new cultural phenomenon.

PLAN

For the past four years, I have led the development of the Scratch website and worked on cultivating an online community that has grown to close to 800,000 people. I have also published work describing the ways people participate, create, collaborate and remix on the Scratch Online Community (Monroy-Hernández, 2007; Monroy-Hernández and Resnick, 2008; Monroy-Hernández et al., 2011; Hill et al., 2010; Aragon et al., 2009; Nickerson and Monroy-Hernández, 2011; Brennan et al., 2010).

Over the next few months, I plan to continue analyzing the structural, functional and attitudinal aspects of remixing in Scratch to better understand this specific phenomenon as well as the community at large. I plan to integrate the findings into a cohesive framework for understanding remixing

[5]Currently working on ways to release an anonymized version of these data to other researchers.

in a way that can be generalizable to other online remixing communities.

Resources

The technical infrastructure to support the website is in already in place. In the coming days, I am planning to add two extra servers that will help further scale the website to handle more traffic and that will let me analyze the large corpus of data. We also have a community manager, an IT consultant and a database specialist on a retainer who are able to assist with this research. I already have the necessary approvals from COUHES[6].

Timeline

- May 2011 – Data aggregation. In the first month, I plan to put together the data sets necessary to answer some research questions described in this document. Much of the data has already been collected, but it needs to be collated in an appropriate format for answering each question. This phase is also exploratory, which will help tweak the framework in this proposal to better represent the patterns found in the data.

- May 2011 – Survey and experiment deployment. Also in the first month, I plan to deploy the necessary design interventions that will help get the necessary data that I do not have already. Most of these experiments are already programmed and just need to be tweaked.

- June, July 2011 – Data analysis. I plan to create the necessary visualizations and statistical analyses to examine some research questions outlined in this proposal. Additionally, I will build a set of representative case studies and execute the necessary interviews.

- August 2011 – Results. I plan to put together the findings in a structured form, and integrate all the findings into a cohesive document accessible to a broad audience.

- September 2011 – Draft – a first draft of the dissertation.

- January 2012 – Defense[7].

- February 2012 – Final version.

References

Aragon, C., Poon, S., Monroy-Hernández, A., and Aragon, D. (2009). A tale of two online communities: Fostering collaboration and creativity in scientists and children. In *Everyday creativity: shared langauges and collective action*, Berkley, CA. ACM.

Batts, D. A. (2011). Patrick cariou vs richard prince, gagosian gallery, inc., lawrence gagosian, and rizzoli international publications, inc.

Benkler, Y. (2002). Coase's penguin, or, linux and the nature of the firm. *Yale Law Journal*, 112(3):369.

Benkler, Y. (2006). *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. Yale University Press, New Haven, CT. Some chapters.

Bolin, M., Webber, M., Rha, P., Wilson, T., and Miller, R. C. (2005). Automation and customization of rendered web pages. *Proceedings of the 18th annual ACM symposium on User interface software and technology*, page 163–172. ACM ID: 1095062.

Brennan, K., Monroy-Hernández, A., and Resnick, M. (2010). Making projects, making friends: Online community as catalyst for interactive media creation. *New Directions for Youth Development*, 2010(128):75–83.

Brown, G. O. (2004). Announcing (and explaining) our new 2.0 licenses. http://creativecommons.org/ weblog/entry/4216.

Cheliotis, G. and Yew, J. (2009). An analysis of the social structure of remix culture. In *Proceedings of the fourth international conference on Communities and technologies*, pages 165–174, University Park, PA, USA. ACM.

Diakopoulos, N., Luther, K., Medynskiy, Y. E., and Essa, I. (2007). The evolution of authorship in a remix society. In *Proc. HT 2007*, pages 133–136, Manchester, UK. ACM.

Donath, J. S. (1998). Identity and deception in the virtual community. In Smith, M. A. and Kollock, P., editors, *Communities in Cyberspace*, pages 27–57. Routledge.

Greenberg, L. A. (1992). Art of appropriation: Puppies, piracy, and Post-Modernism, the. *Cardozo Arts & Entertainment Law Journal*, 11:1.

Hill, B. M. H., Monroy-Hernández, A., and Olson, K. (2010). Responses to remixing on a social media sharing website. In *Proc. ICWSM 2010*, pages 74–81, Washington, D.C. AAAI.

Ito, M. (2007). Technologies of the childhood imagination: Yu-Gi-Oh!, media mixes, and everyday cultural production. In Karaganis, J., editor, *Structures of Participation in Digital Culture*, pages 86–111. The Social Science Research Council, New York.

Ito, M. (2010). *Hanging out, messing around, and geeking out : kids living and learning with new media*. John D. and Catherine T. MacArthur Foundation series on digital media and learning. MIT Press, Cambridge Mass.

Jenkins, H. (2006). *Convergence Culture: Where Old and New Media Collide*. NYU Press, revised edition.

Jenkins, H., Purushotma, R., Clinton, K., Weigel, M., and Robinson, A. (2009). *Confronting the Challenges of Participatory Culture: Media Education for the 21st Century*. John D. and Catherine T. MacArthur Foundation Reports on Digital Media and Learning. The MIT Press, Cambridge, MA.

Kittur, A. and Kraut, R. E. (2008). Harnessing the wisdom of crowds in wikipedia: quality through coordination. In *Proceedings of the ACM 2008 conference on Computer supported cooperative work*, pages 37–46, San Diego, CA, USA. ACM.

[6]Committee On the Use of Humans as Experimental Subjects

[7]This date might change depending on whether whether I am in non-resident status or not

Landes, W. M. (2000). Copyright, borrowed images, and appropriation art: An economic approach. *George Mason Law Review*, 9:1.

Lave, J. and Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, 1 edition.

Lessig, L. (2008). *Remix: Making Art and Commerce Thrive in the Hybrid Economy*. Penguin Press, New York, NY.

Livingstone, S. (2008). Taking risky opportunities in youthful content creation: teenagers' use of social networking sites for intimacy, privacy and self-expression. *New Media & Society*, 10(3):393 –411.

Manovich, L. (2005). Remix and remixability.

Monroy-Hernández, A. (2007). ScratchR: sharing user-generated programmable media. *Proceedings of the 6th international conference on Interaction design and children*, page 167–168. ACM ID: 1297315.

Monroy-Hernández, A., Hill, B. M., Gonzalez-Rivero, J., and danah boyd (2011). Computers can't give credit : How automatic attribution falls short in an online remixing community. In *Proceedings of the 29th international conference on Human factors in computing systems*, Vancouver, British Columbia, Canada. ACM Press.

Monroy-Hernández, A. and Resnick, M. (2008). Empowering kids to create and share programmable media. *interactions*, 15(2):50–53.

Nickerson, J. V. and Monroy-Hernández, A. (2011). Appropriation and creativity: Userinitiated contests in scratch. pages 1–10.

Papert, S. (1980). *Mindstorms : children, computers, and powerful ideas*. Basic Books, New York.

Perkel, D. (2008). Copy and paste literacy: Literacy practices in the production of a MySpace profile. In Drotner, K., Jensen, H. S., and Schroder, K. C., editors, *Informal Learning and Digital Media*, pages 203–224. Cambridge Scholars Press, Newcastle, UK.

Posner, R. A. (2007). *The little book of plagiarism*. Pantheon Books.

Raymond, E. (1999). The cathedral and the bazaar. *Knowledge, Technology & Policy*, 12(3):23–49.

Resnick, M. (2008). Sowing the seeds for a more creative society. *Learning & Leading with Technology*, 35(4):18–22.

Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., Silverman, B., and Kafai, Y. (2009). Scratch: Programming for all. *Commun. ACM*, 52(11):60–67.

Seneviratne, O., Kagal, L., and Berners-Lee, T. (2009). Policy-Aware content reuse on the web. In Bernstein, A., Karger, D., Heath, T., Feigenbaum, L., Maynard, D., Motta, E., and Thirunarayan, K., editors, *The Semantic Web - ISWC 2009*, volume 5823 of *Lecture Notes in Computer Science*, pages 553–568. Springer Berlin / Heidelberg.

Seneviratne, O. and Monroy-Hernández, A. (2010). Remix culture on the web: A survey of content reuse on different User-Generated content websites. In *WebSci10: Extending the Frontiers of Society On-Line*, Raleigh, NC.

Shaw, R. and Schmitz, P. (2006). Community annotation and remix: a research platform and pilot deployment. *Proceedings of the 1st ACM international workshop on Human-centered multimedia*, page 89–98. ACM ID: 1178761.

Sinnreich, A., Latonero, M., and Gluck, M. (2009). Ethics reconfigured: How today's media consumers evaluate the role of creative reappropriation. *Information, Communication & Society*, 12(8):1242.

Tichy, W. F. (1985). Rcs — a system for version control. *Software: Practice and Experience*, 15(7):637–654.

Turkle, S. and Papert, S. (1990). Epistemological pluralism: Styles and voices within the computer culture. *Signs*, 16(1):128–157.

Viégas, F. B., Wattenberg, M., and Dave, K. (2004). Studying cooperation and conflict between authors with history flow visualizations. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 575–582, Vienna, Austria. ACM.

Wenger, E. (1998). Communities of practice learning as a social system. *1998*, 9(5).

Wong, J. and Hong, J. I. (2007). Making mashups with marmite: towards end-user programming for the web. *Proceedings of the SIGCHI conference on Human factors in computing systems*, page 1435–1444. ACM ID: 1240842.

Zang, N. and Rosson, M. B. (2008). What's in a mashup? and why? studying the perceptions of web-active end users. *Proceedings of the 2008 IEEE Symposium on Visual Languages and Human-Centric Computing*, page 31–38. ACM ID: 1550043.