Responses to remixing on a social media sharing website

Andrés Monroy-Hernández
Benjamin Mako Hill
Kristina Olson
Massachusetts Institute of Technology | Yale University

ICWSM 2010
Remixing

Creation of new content, such as songs and video, using existing content produced by others.

• Remixing important cultural phenomenon (Lessig 2009, Manovich 2005, Benkler 2006)

• Literature focuses on youth (Jenkins 2006, Ito 2006, Palfrey and Gasser 2008)
Empirical research on remixing communities

Remixing has been understudied (Cheliotis and Yew 2009):

- Largely exploratory and qualitative (e.g. Diakopolous et al. 2007, Cheliotis and Yew 2009)
- Focuses on adult remixers
- Emphasizes remixing of “external” content (e.g. Shaw and Schmitz 2006)
Research Questions

1. How do young people respond to remixing?

2. When do users react negatively to remixing?
Tool and online community
when [space] key pressed
repeat 10
move 10 steps
next costume
Download this project!

Download the 27 sprites and 176 scripts of "Neo" and open it in Scratch

Project Notes

The Instructions are in the game
Post your high score and difficulty level!

And please post a love it! because i will make more games if you do!

UPDATE: added ship customization!

UPADTE: added a boss on level nine!

Tags

game
animation
art
cool
Scratch online community

• Launched 2007
• 500,000+ users (median age: 12)
• 1,000,000+ projects
• Designed for remixing
  – All projects remixable
  – Remixing explicitly encouraged
  – Permissive licensing
Remixing in Scratch

Neo

Click To Start

star wars: Republican Gunship

Click To Start

comment by originator
Data and procedures

March 2007 ➔ April 2008
136,929 projects
11,861 remixes
Coded comments (2 coders)
Project metadata
Comment categories

plagiarism accusation
“Hello mr plagiarist”,
“Copy-cat!”

positive
“Love what you did with my code! Great idea!”

hinting plagiarism
“I mostly pretty much made this whole entire game”

negative
“Alright you crap eating thumb sucking baby”

none of the above
“b for peanut butter jelly time!”

no comment
“”
Study 1: How do people respond to remixing?

1. What proportion of users respond positively/negatively when their project is remixed?

2. What proportion of users accuse remixers of plagiarism when their project is remixed?
Study 1: Results

Proportion of original creators who left comments on remixes

Proportion of comments that accused remixers of plagiarism by the valance of those comments (n=3742)
Study 1: Take aways

- People just as likely to leave positive comments (21%) as they were to leave a direct or indirect complaint of plagiarism (22%).
Study 2: When do creators accuse remixers of plagiarism?

**H1.** Originators of more *complicated* projects are more likely to make plagiarism accusations.

**H2.** Originators are *less likely* to accuse remixers of plagiarism when their project *is a remix itself.*

**H3.** Originators are *less likely* to accuse remixers of plagiarism if they have *shared at least one remix.*
Study 2: Data

Coded comments from remix pairs that were viewed by originator. \(n=3,742\)

Question predictors:
- Project complexity (number of sprites) (H1)
- Was project a remix itself? (H2)
- Had originator remixed? (H3)

Controls for:
- age (creator and remixer)
- gender (creator)
- shared date
Study 2: Analysis

Logistic regression models of the likelihood of an originator accusing a remixer of plagiarism.

- Tested several alternative specifications that also included all negative reactions and excluding hinting accusations with very similar results.
Study 2: Results (H1)

H1: Authors of *larger* or *more complicated* projects will be more likely to accuse remixers of their work of plagiarism.

Finding:
Support
H2: Originators will be less likely to accuse remixers of their projects of plagiarism when the remixed project is itself a remix.

Finding: No Support
**Study 2: Results (H3)**

**H3:** Originators will be less likely to accuse remixers of their projects of plagiarism if *they have shared at least one remix themselves.*

**Finding:**
No Support
Study 2: Takeaways

• Smallness of contributions (not cumulative contributions) may drive acceptance of remixing

• Users seem to have a strong concept of good remixing and bad remixing (Diakopoulos et al, 2007)
Study 3: Original/Remix Similarity

H4. Originators are more likely to accuse remixers of plagiarism when the remixed project and its antecedent are more similar to each other.

Data/Procedure:
• 240 pairs of remixes
  • 40 for each category in the initial coding scheme
• Coded from:
  • 1 (can't tell they're related) to
  • 5 (can't tell they're different)
Study 3: Results

Finding:
Support

Details:
Projects marked as plagiarism (m=4.40, sd=.65) were more similar than:
- negative (m=3.93, sd=1.02) *
- positive (m=3.75, sd=.85) **
- hinting plagiarism (m=3.46, sd=1.30) ***
- no comment (m=3.65, sd=1.14) ***
- other (m=3.53, sd=.85) ***

(*** p < .001; ** p < 0.01; * p < 0.1)
Thank you!

http://scratch.mit.edu

andresmh@media.mit.edu | mako@mit.edu
Data Analysis Appendix
Study 2: Results

Other results:

- Female users accuse others of plagiarism much less than men (i.e., 0.6 times the odds)
- Younger originator are more likely to complain
- Younger remixers are more likely to elicit complaints
<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ACCUSE.PLAG</td>
<td>0.13</td>
<td>0.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. SPRITES</td>
<td>9.05</td>
<td>12.24</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. ORIG.REMIX</td>
<td>0.05</td>
<td>0.21</td>
<td>-0.01</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. HAS.REMIXED</td>
<td>0.79</td>
<td>0.41</td>
<td>-0.01</td>
<td>-0.07</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. FEMALE</td>
<td>0.25</td>
<td>0.43</td>
<td>-0.06</td>
<td>-0.11</td>
<td>0.04</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. WEEKS</td>
<td>35.15</td>
<td>14.49</td>
<td>0.01</td>
<td>-0.03</td>
<td>0.08</td>
<td>0.15</td>
<td>0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. AGE</td>
<td>17.08</td>
<td>10.15</td>
<td>-0.07</td>
<td>0.08</td>
<td>0.01</td>
<td>-0.08</td>
<td>-0.20</td>
<td>-0.16</td>
<td></td>
</tr>
<tr>
<td>8. REMIXER.AGE</td>
<td>15.01</td>
<td>9.58</td>
<td>-0.05</td>
<td>0.01</td>
<td>0.06</td>
<td>-0.09</td>
<td>-0.05</td>
<td>-0.05</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Table 1: Means, standard deviations, and correlations between variables used in the logistic regression analysis in Study 2. The sample includes all remixed projects that had been clicked on and viewed by the originators. (n=3742)
ICWSM Study 2: Correlation Table

<table>
<thead>
<tr>
<th></th>
<th>Model 0</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.904***</td>
<td>-2.159***</td>
<td>-2.301***</td>
<td>-2.302***</td>
<td>-2.287***</td>
<td>-2.155***</td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.312)</td>
<td>(0.315)</td>
<td>(0.315)</td>
<td>(0.325)</td>
<td>(0.335)</td>
</tr>
<tr>
<td>Female</td>
<td>-0.534***</td>
<td>-0.496***</td>
<td>-0.495***</td>
<td>-0.494***</td>
<td>-0.489***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.127)</td>
<td>(0.128)</td>
<td>(0.128)</td>
<td>(0.129)</td>
<td>(0.131)</td>
<td></td>
</tr>
<tr>
<td>Weeks</td>
<td>0.064***</td>
<td>0.064***</td>
<td>0.064***</td>
<td>0.064***</td>
<td>0.065***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.019)</td>
<td>(0.019)</td>
<td>(0.019)</td>
<td>(0.020)</td>
<td></td>
</tr>
<tr>
<td>Weeks²</td>
<td>-0.001***</td>
<td>-0.001***</td>
<td>-0.001***</td>
<td>-0.001***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.029***</td>
<td>-0.031***</td>
<td>-0.031***</td>
<td>-0.031***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.006)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sprites (std)</td>
<td>0.210***</td>
<td>0.210***</td>
<td>0.209***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.042)</td>
<td>(0.042)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orig.Remix</td>
<td></td>
<td>-0.059</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.247)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Has.Remixed</td>
<td></td>
<td></td>
<td>-0.022</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.124)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remixer.Age</td>
<td></td>
<td></td>
<td></td>
<td>-0.018**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.007)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>3742</td>
<td>3615</td>
<td>3615</td>
<td>3615</td>
<td>3615</td>
<td>3480</td>
</tr>
<tr>
<td>AIC</td>
<td>2888.162</td>
<td>2758.412</td>
<td>2736.849</td>
<td>2738.790</td>
<td>2738.816</td>
<td>2599.331</td>
</tr>
<tr>
<td>BIC</td>
<td>2913.072</td>
<td>2882.269</td>
<td>2885.477</td>
<td>2912.190</td>
<td>2912.216</td>
<td>2771.665</td>
</tr>
<tr>
<td>log L</td>
<td>-1440.081</td>
<td>-1359.206</td>
<td>-1344.424</td>
<td>-1341.395</td>
<td>-1341.408</td>
<td>-1271.665</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
† significant at *p < .10; *p < .05; **p < .01; ***p < .001

Table 2: Taxonomy of logistic regression models on ACCUSE.PLAG, a dichotomous construct representing whether project creators accused the mixer of their project of plagiarism in a non-positive manner.
Prototypical plots from fitted logistic model showing predicted probability of eliciting a plagiarism accusations for projects created by 12 year old males in Scratch's 35th week. Panel 1 (top) shows the effects of additional sprites (through 90th percentile). Panels 2 and 3 (bottom) hold sprites constant at the sample median (5). (n=3742)