Quality Control Mechanisms in Crowdsourcing Systems

Clemens Pollak
Improving Learning through Achievement Priming in Crowdsourced Information Finding Microtasks

- by Ujwal Gadiraju & Stefan Dietze
- from L3S Research Center (Leibniz Universität Hannover)
Improving Learning through Achievement Priming in Crowdsourced Information Finding Microtasks

- Problem of large batch size
- Research Questions and Hypotheses
- Experimental setup
  - Inspiring Quotes
  - Achievement Motivation
  - Variations of priming
  - Design decisions
- Results
  - Key Findings
  - Worker orientation
  - Learning Rate
Problem of large batch size

- Many crowdsourcing tasks are large batches of repetitive subtasks
  → More attractive because learning pays off
  → The fewer tasks left, the fewer contributions

**GOAL:** Retaining crowd workers in large batches of microtasks
Research Questions

1. How can achievement priming be used to increase worker retention and facilitate learning in information finding tasks?
2. How is the learning process of workers affected when they are retained in long batches of information finding tasks?

Hypotheses

1. Achievement primes can improve the worker retention rate and thereby facilitate learning
2. Retained workers learn more about the tasks and perform more effectively
Experimental setup

- Task type
  - Information finding
- Examining
  - Increase of motivation by *achievement priming*
- Measured by
  - Worker retention rate:
    - average number of tasks that workers complete in a given batch
  - Worker performance:
    - accuracy with which the workers complete the tasks
  - Worker learning rate:
    - average change of worker performance over the course of the batch completion
Inspiring Quotes

- CrowdFlower task with quotes related to achievement
- Top 25 Quotes used for priming (rating >= 4.5)
Achievement Motivation

- Divide workers into two groups
  - Fun oriented
  - Achievement oriented
- Evaluate crowdworkers on established test on motivation
  - Questioning from
    1: Not at all like me
    5: Extremely like me
Variations of priming

1. Passive Achievement Priming
2. Active Achievement Priming
3. No Priming (Baseline)
4. Random Quotes Passive
5. Random Quotes Active
Design decisions

- Attention checks
- Randomized orders
- No overlap in workers
- Top level workers on CrowdFlower
- 5 USD cents for 10 units
Web Science – Investigating the Future of Information and Communication

![Graph showing the percentage of workers against the number of units completed before dropping out.](image)

<table>
<thead>
<tr>
<th>Priming Variation</th>
<th>Avg. Acc. Per Worker (%)</th>
<th>Worker Retention Rate (%)</th>
<th>Overall Task Acc. (%)</th>
<th>Avg. TCT (in mins)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP-Baseline</td>
<td>76.71</td>
<td>36.23</td>
<td>78.88</td>
<td>9.11*</td>
</tr>
<tr>
<td>AP-Passive</td>
<td>77.72</td>
<td>36.08</td>
<td>77.60</td>
<td>10.67</td>
</tr>
<tr>
<td>AP-Active</td>
<td>76.35</td>
<td>44.64*</td>
<td>79.96</td>
<td>11.44</td>
</tr>
<tr>
<td>RQ-Passive</td>
<td>78.74</td>
<td>28.58</td>
<td>77.58</td>
<td>9.72</td>
</tr>
<tr>
<td>RQ-Active</td>
<td>74.81</td>
<td>32.61</td>
<td>75.41</td>
<td>9.87</td>
</tr>
</tbody>
</table>
Accuracy

- Learning effect
- Improved by active priming

(a) NP-Baseline

(b) AP-Active

(c) AP-Passive
### Learning Rate of Workers by Worker Groups

<table>
<thead>
<tr>
<th>Worker Groups</th>
<th>NP-Baseline</th>
<th>AP-Active</th>
<th>AP-Passive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short</td>
<td>2.96</td>
<td>3.43*</td>
<td>3.15</td>
</tr>
<tr>
<td>Medium</td>
<td>1.28</td>
<td>1.30</td>
<td>1.55*</td>
</tr>
<tr>
<td>Long</td>
<td>0.69</td>
<td>0.64</td>
<td>0.73</td>
</tr>
<tr>
<td><strong>Overall Avg.</strong></td>
<td>1.64</td>
<td>1.79</td>
<td>1.81</td>
</tr>
</tbody>
</table>
Results - Worker orientation

- 301/340 workers achievement oriented

- Active achievement priming

<table>
<thead>
<tr>
<th>Achievement Motivation</th>
<th>Avg. Acc. Per Worker (%)</th>
<th>Worker Retention Rate (%)</th>
<th>Overall Task Acc. (%)</th>
<th>Avg. TCT (in mins)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACHIEVE</td>
<td>79.60</td>
<td>36.78</td>
<td>80.96</td>
<td>10.74</td>
</tr>
<tr>
<td>FUN</td>
<td>75.85</td>
<td>26.19</td>
<td>74.25</td>
<td>9.60</td>
</tr>
</tbody>
</table>

- Passive achievement priming

<table>
<thead>
<tr>
<th>Achievement Motivation</th>
<th>Avg. Acc. Per Worker (%)</th>
<th>Worker Retention Rate (%)</th>
<th>Overall Task Acc. (%)</th>
<th>Avg. TCT (in mins)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACHIEVE</td>
<td>76.98</td>
<td>45.60</td>
<td>80.45</td>
<td>11.41</td>
</tr>
<tr>
<td>FUN</td>
<td>65.17</td>
<td>27.78</td>
<td>67</td>
<td>12.02</td>
</tr>
</tbody>
</table>
Conclusion

- highest quality level workers, from CrowdFlower are driven by achievement as opposed to fun
- active interaction with the primes led to an improvement of over 8% on comparison to the baseline method devoid of primes
Curiosity Killed the Cat, but Makes Crowdwork Better

- by Edith Law, Ming Yin, Joslin Goh, Kevin Chen, Michael Terry, Krzysztof Z. Gajos
- from University of Waterloo & Harvard University
Curiosity Killed the Cat, but Makes Crowdwork Better

- Curiosity
- Research Questions
- Curiosity interventions
- Experimental setup
- Results
  - Effectivity of curiosity interventions
  - Reasons for persisting
  - Accuracy
  - Guessing behaviour
  - Task characteristics and curiosity interventions
- Conclusion
- Discussion
Curiosity

- **Definition:** 
  *the desire to know, to see, or to experience that motivates exploratory behavior directed towards the acquisition of information*

- **Information gap theory**
  - knowledge baseline
  - information goal

- **Inducing curiosity**
  - people are aware of a salient information gap in their knowledge (guessing is increasing awareness)
  - people are provided with a means to help them close the gap
Research Questions

1. Can crowds be motivated by curiosity?

Hypotheses

If workers are presented with a curiosity-inducing stimulus, they will

a. complete more tasks
b. have a higher probability of completing all 30 tasks
c. have similar or better performance
Research Questions

2. How do individuals respond differently to curiosity interventions?

Hypotheses

Workers who make a correct guess will

- a. complete more tasks
- b. have a higher probability of completion
- c. have similar or better performance
Research Questions

3. What are the interactions between task characteristics and curiosity interventions?

Hypotheses

The effect of a curiosity intervention is larger when the intervention is combined with a task that is less interesting
Curiosity interventions

- Information Goal
  - Create an information gap by posing a question that is relevant to the current task at hand
- Gap Salience
  - Prompt workers to guess the answer to the question.
- Incremental Reveal
  - Reveal information as workers progress with their tasks
Experimental setup

- Audio transcription task
- 5 articles
  - crowdsourced for “boringness”
  - divided into 30 parts
  - random fixed order
- Workers have to complete min 3 tasks
- Additional 1 cent bonus for every additional task
Experimental setup

(a) Baseline

(b) Question Only

(c) Narrative

(d) Ordered Photo

(e) Scrambled Photo
(a) Question Only vs. Baseline

(b) Narrative vs. Baseline

(c) Ordered Photo vs. Baseline

(d) Scrambled Photo vs. Baseline
Results

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Median</th>
<th>Mean (SD)</th>
<th>Completion Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>13</td>
<td>15.86 (10.38)</td>
<td>31%</td>
</tr>
<tr>
<td>Question only</td>
<td>13</td>
<td>15.97 (9.76)</td>
<td>29%</td>
</tr>
<tr>
<td>Narrative</td>
<td>17.5</td>
<td>18.85 (11.14)</td>
<td>47%</td>
</tr>
<tr>
<td>Ordered Photo</td>
<td>13</td>
<td>16.45 (10.88)</td>
<td>36%</td>
</tr>
<tr>
<td>Scrambled Photo</td>
<td>18</td>
<td>19.03 (11.04)</td>
<td>47%</td>
</tr>
</tbody>
</table>

Table 2. Summary of quit index in different conditions.
Reasons for persisting

- **payment**
  - “I wanted to make more money from the bonuses”

- **learning**
  - “It was good practice, I felt I was getting better as I went on”

- **completionist attitude**
  - “I don’t like leaving things half-finished”

- **curiosity**
  - “I wanted to see if I could figure out the player’s name.”
  - “I persisted because I was curious about how the article would unfold.”
  - “I wanted to know who the article is about. It was like getting a puzzle piece and putting it all together.”
Accuracy

- Very high in baseline and experiment
- No statistical difference
Guessing behaviour

![Pie chart and graph showing guessing behaviour](image)

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Correct Guess</th>
<th>Incorrect Guess</th>
<th>No Guess</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quit Index</td>
<td>Median: 30</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Mean (σ): 22.63 (9.84)</td>
<td>13.60 (10.03)</td>
<td>15.37 (10.23)</td>
</tr>
<tr>
<td>Completion</td>
<td>%: 62</td>
<td>24</td>
<td>29</td>
</tr>
<tr>
<td>Error Rate</td>
<td>Median: 0.03</td>
<td>0.05</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>Mean (σ): 0.04 (0.04)</td>
<td>0.06 (0.04)</td>
<td>0.05 (0.03)</td>
</tr>
</tbody>
</table>
Task characteristics and curiosity interventions

- Most boring and 2 most interesting articles
- More interesting tasks being completed more reliably
Conclusion

- stimulating curiosity can be an effective way to incentivize crowd workers
  - improving worker retention
  - maintaining high level of performance
- workers who made a correct guess
  - complete a significantly larger number
  - have significantly higher quality

*Information can be used as currency for workers*
Discussion