The Effects of Performance-Contingent Financial Incentives in Online Labor Markets

Ming Yin (Harvard), Yiling Chen (Harvard), Yu-An Sun (Xerox)

AAAI-13, Bellevue, WA, USA, July 14-18, 2013

Various Incentives in Crowdsourcing

amazon mechanical turk™
Artificial Artificial Intelligence

freelancer.com

oDesk®
Changing How the World Works.

crowdSPRING

99designs®

[TopCoder]®

innovation exchange
where CREATIVITY is the CURRENCY

duolingo

memrise

ZOONIVERSE
REAL SCIENCE ONLINE

foldit
Solve Puzzles for Science

eBird
Where your bird sightings make a difference!

Related Work

- [Mason and Watts (2009), Rogstadius et al. (2011)] When the magnitude of performance–independent financial incentive increased, work quality was not improved.
  - Post–task surveys showed that worker’s perception of appropriate compensation amount was consistently higher than and increasing with the actual payment.

- [Harris (2011)] The quality of work was higher when performance–contingent financial incentives (both rewards and punishment) exists.

Research Questions

For performance–contingent financial incentives:

- Does the magnitude of financial incentives affect work quality and/or worker effort in online labor markets?

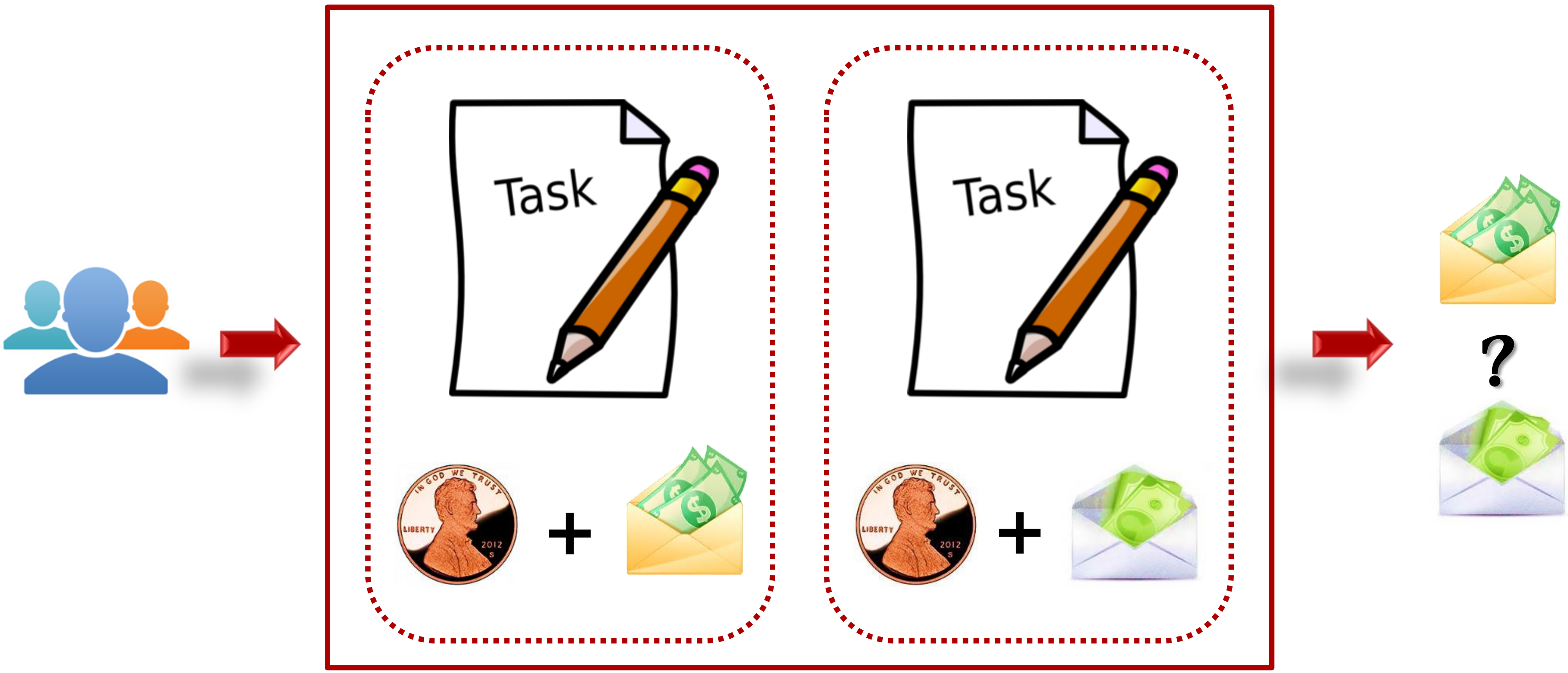
  Higher pay: higher effort & better performance?

- Will work quality and/or worker effort change if the magnitude of financial incentives changes in a sequence?

  Increasing pay: higher effort & better performance?

Experiment Method

- Package two tasks in one HIT (**H**uman **I**ntelligence **T**ask)



Payment in each task = Base payment  + Bonus  or 

Experiment Treatments

- Four bonus levels: 4 cents, 8 cents, 16 cents, 32 cents

- Ten treatments:
  - Base treatments: 4 cents—4 cents, 8 cents—8 cents, 16 cents—16 cents, 32 cents—32 cents
  - Treatments with increasing bonus: 4 cents—8 cents, 4 cents—16 cents, 4 cents—32 cents
  - Treatments with decreasing bonus: 8 cents—4 cents, 16 cents—4 cents, 32 cents—4 cents

Spotting Differences (SD) and Button Clicking (BC)

- Spotting Differences: Mainly requires cognitive skills
- Button Clicking: Mainly requires motor skills

Work Quality:

# of Differences Correctly Spotted

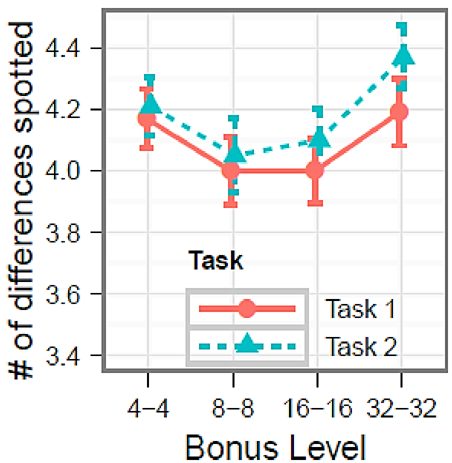
Worker Effort:

Total Time

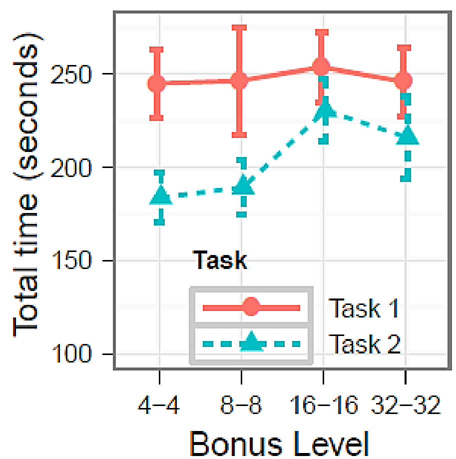
Longest Interval

Does the Magnitude Alone Matter?

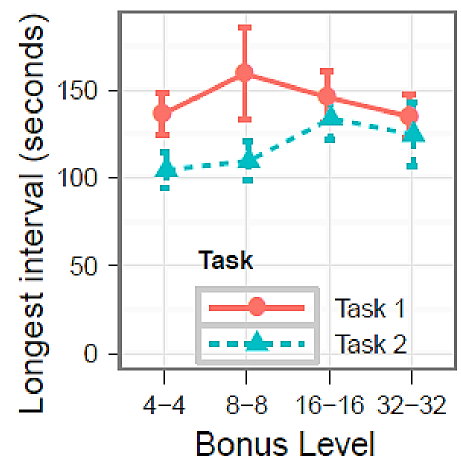
- Both work quality and worker effort have similar mean values across 4 base treatments.



(b) SD: Work Quality

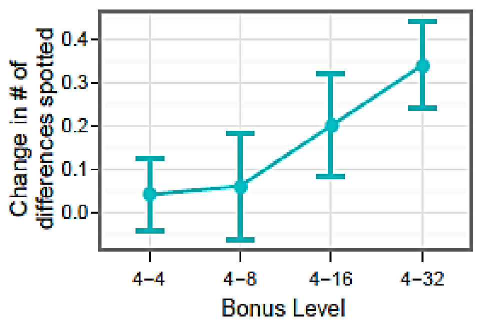


(c) SD: Total Time

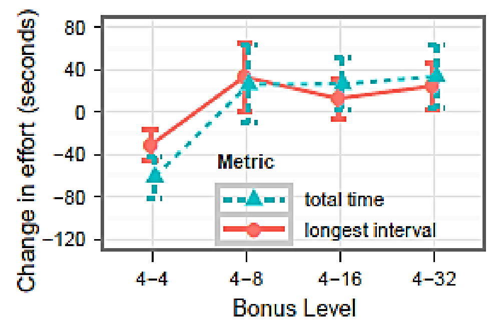


(d) SD: Longest Interval

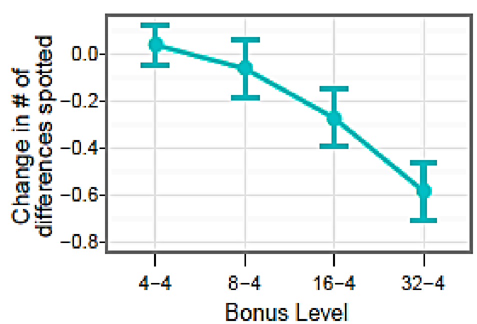
Can Incentives be More Effective in Sequence?



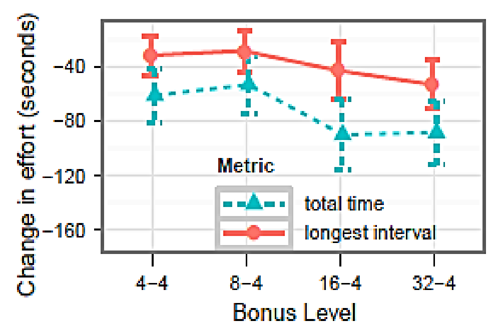
(b) SD: Change in work quality for treatments with increasing bonus



(c) SD: Change in worker effort for treatments with increasing bonus



(e) SD: Change in work quality for treatments with decreasing bonus



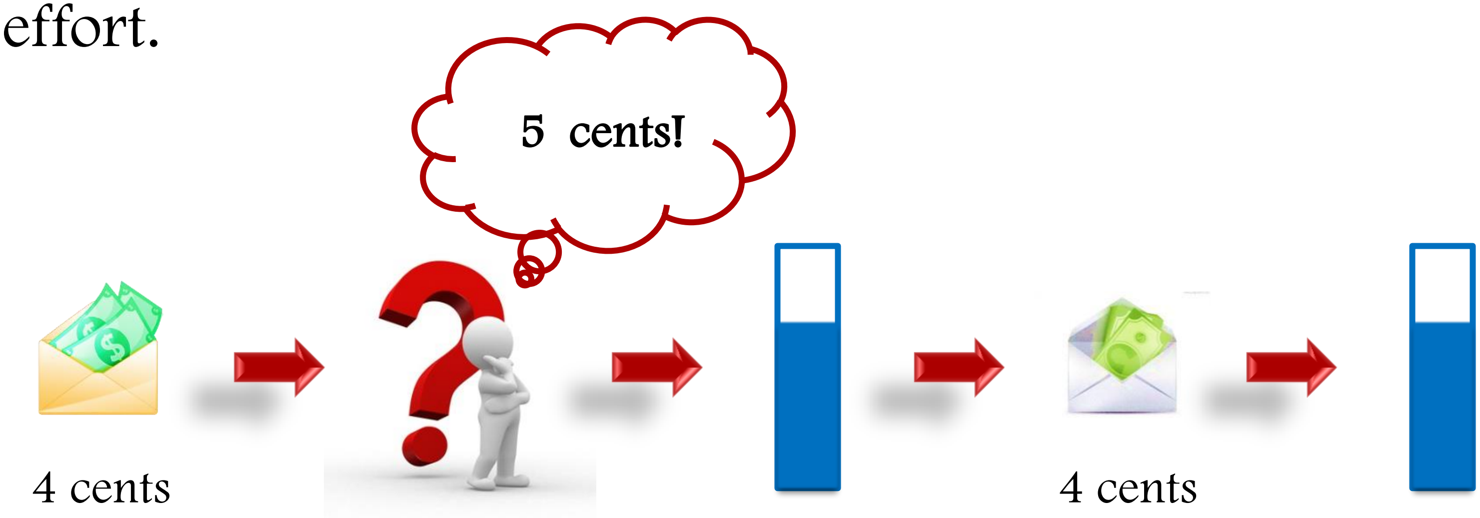
(f) SD: Change in worker effort for treatments with decreasing bonus

Change in the financial incentives over the subsequent two tasks affects both work quality and worker effort!

Interpretation of the Result

- **Anchoring effect** [Tversky and Kahneman (1974)]: People tend to rely heavily on the first piece of information to make judgment.

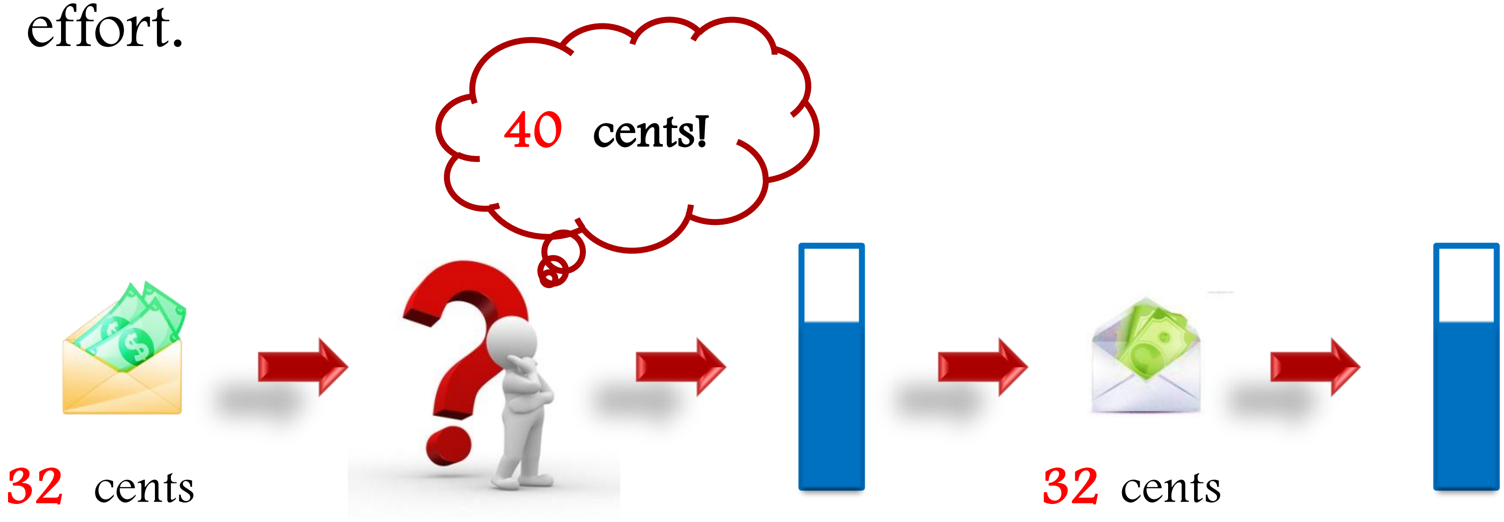
- **Fair wage–effort hypothesis** [Akerlof and Yellen (1988)]: If the actual wage is less than worker's perception of fair wage, they will supply a corresponding fraction of their normal effort.



Interpretation of the Result

- **Anchoring effect** [Tversky and Kahneman (1974)]: People tend to rely heavily on the first piece of information to make judgment.

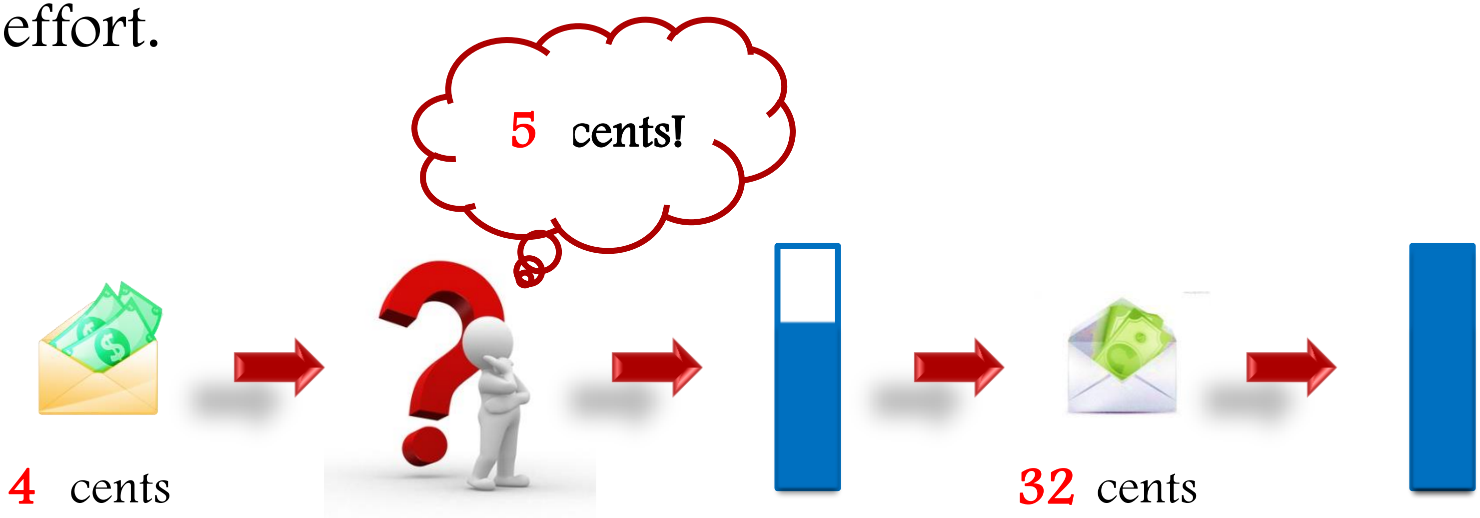
- **Fair wage–effort hypothesis** [Akerlof and Yellen (1988)]: If the actual wage is less than worker's perception of fair wage, they will supply a corresponding fraction of their normal effort.



Interpretation of the Result

- **Anchoring effect** [Tversky and Kahneman (1974)]: People tend to rely heavily on the first piece of information to make judgment.

- **Fair wage–effort hypothesis** [Akerlof and Yellen (1988)]: If the actual wage is less than worker's perception of fair wage, they will supply a corresponding fraction of their normal effort.



Future Work

- Sequence with more than two tasks: Which reward matters the most?
- Anchoring effect in other dimension: anchoring on the difficulty levels?

Thank you!

mingyin@fas.harvard.edu