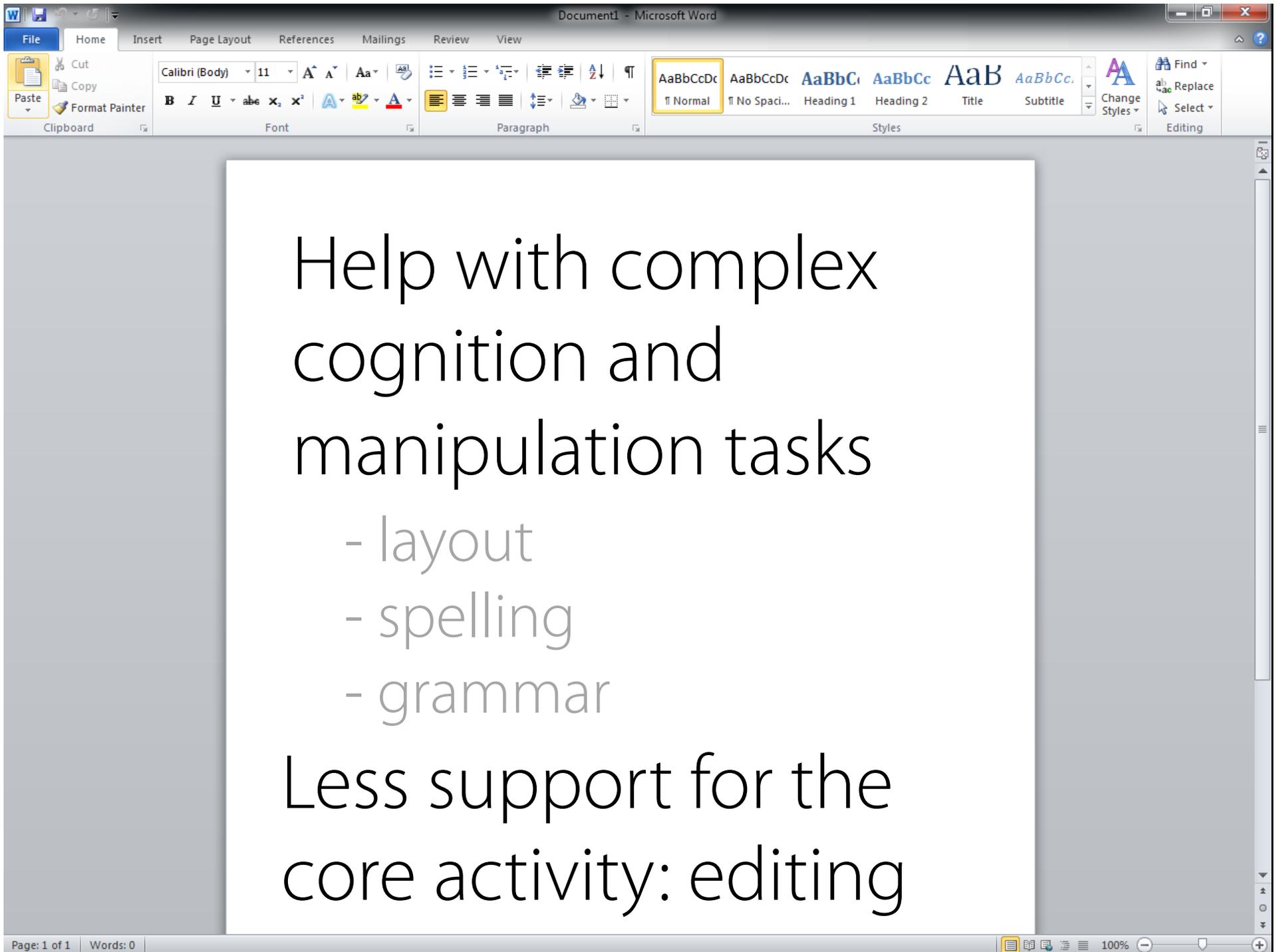


Crowd-Powered Systems



Michael Bernstein
Stanford Computer Science
HCI Group



Help with complex cognition and manipulation tasks

- layout
- spelling
- grammar

Less support for the core activity: editing

REFERENCES

1. Bernstein, M., Marcus, A., Karger, D.R., and Miller, R.C. Enhancing Directed Content Sharing on the Web. *CHI '10*, ACM Press (2010).
2. Bernstein, M., Tan, D., Smith, G., Czerwinski, M., et al. Collabio: A Game for Annotating People within Social Networks. *UIST '09*, ACM Press (2009), 177–180.
3. Bigham, J.P., Jayant, C., Ji, H., Little, G., et al. VizWiz: Nearly Real-time Answers to Visual Questions. *UIST '10*, ACM Press (2010).
21. Quinn, A.J. and Bederson, B.B. A Taxonomy of Distributed Human Computation.
22. Ross, J., Irani, L., Silberman, M.S., Zaldivar, A., et al. Who Are the Crowdworkers? Shifting Demographics in Amazon Mechanical Turk. *alt.chi '10*, ACM Press.
23. Sala, M., Partridge, K., Jacobson, L., and Begole, J. An Exploration into Activity-Informed Physical Advertising Using PEST. *Pervasive '07*, Springer Berlin Heidelberg (2007).
24. Simon, I., Morris, D., and Basu, S. MySong: automatic accompaniment generation for vocal melodies. *Proc. CHI '08*, ACM Press (2008).
25. Snow, R., O'Connor, B., Jurafsky, D., and Ng, A.Y. Cheap and fast—but is it good?: evaluating non-expert annotations for natural language tasks. *ACL '08*, (2008).
26. Sorokin, A. and Forsyth, D. Utility data annotation with Amazon Mechanical Turk. *CVPR '08*, (2008).
27. von Ahn, L. and Dabbish, L. Labeling images with a computer game. *CHI '04*, ACM Press (2004).

Shortening a paper

Supported by human editors

Related Work

Crowdsourcing: A Batch Platform

Data collection, machine learning training,
user studies, social science experiments

[Ipeirotis 2010, Heer et al. 2010, Kittur et al. 2008]

Games with a purpose

[von Ahn and Dabbish 2004, Cooper et al. 2011]

Collective action

[Wikipedia, Polymath Project, Search for Jim Gray]

Historical roots: distributed calculation
of mathematical tables

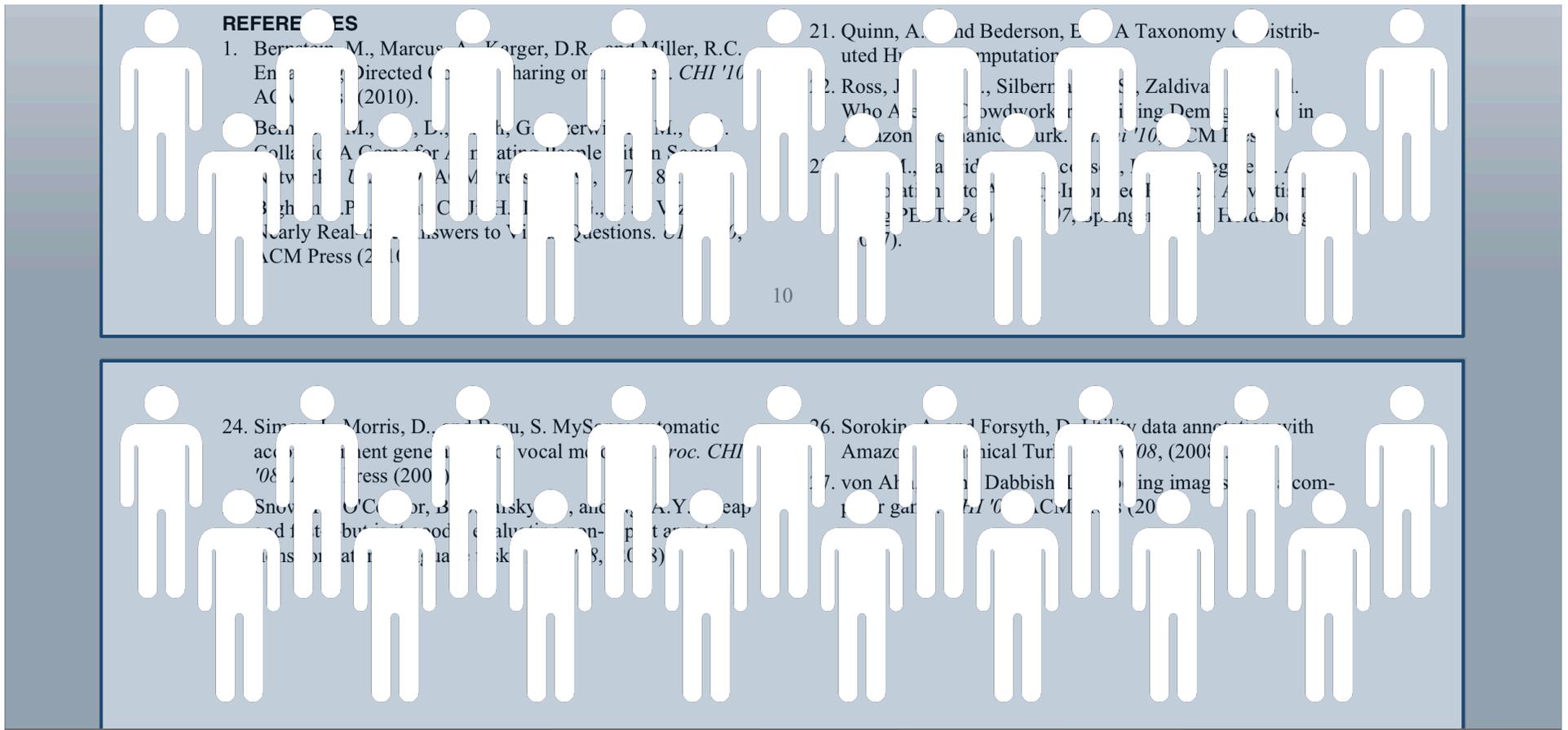
[Grier 2007]

REFERENCES

1. Bernstein, M., Marcus, A., Karger, D.R., and Miller, R.C. Enhancing Directed Content Sharing on the Web. *CHI '10*, ACM Press (2010).
2. Bernstein, M., Tan, D., Smith, G., Czerwinski, M., et al. Collabio: A Game for Annotating People within Social Networks. *UIST '09*, ACM Press (2009), 177–180.
3. Bigham, J.P., Jayant, C., Ji, H., Little, G., et al. VizWiz: Nearly Real-time Answers to Visual Questions. *UIST '10*, ACM Press (2010).
21. Quinn, A.J. and Bederson, B.B. A Taxonomy of Distributed Human Computation.
22. Ross, J., Irani, L., Silberman, M.S., Zaldivar, A., et al. Who Are the Crowdworkers? Shifting Demographics in Amazon Mechanical Turk. *alt.chi '10*, ACM Press.
23. Sala, M., Partridge, K., Jacobson, L., and Begole, J. An Exploration into Activity-Informed Physical Advertising Using PEST. *Pervasive '07*, Springer Berlin Heidelberg (2007).
24. Simon, I., Morris, D., and Basu, S. MySong: automatic accompaniment generation for vocal melodies. *Proc. CHI '08*, ACM Press (2008).
25. Snow, R., O'Connor, B., Jurafsky, D., and Ng, A.Y. Cheap and fast—but is it good?: evaluating non-expert annotations for natural language tasks. *ACL '08*, (2008).
26. Sorokin, A. and Forsyth, D. Utility data annotation with Amazon Mechanical Turk. *CVPR '08*, (2008).
27. von Ahn, L. and Dabbish, L. Labeling images with a computer game. *CHI '04*, ACM Press (2004).

Shortening a paper

Supported by human editors

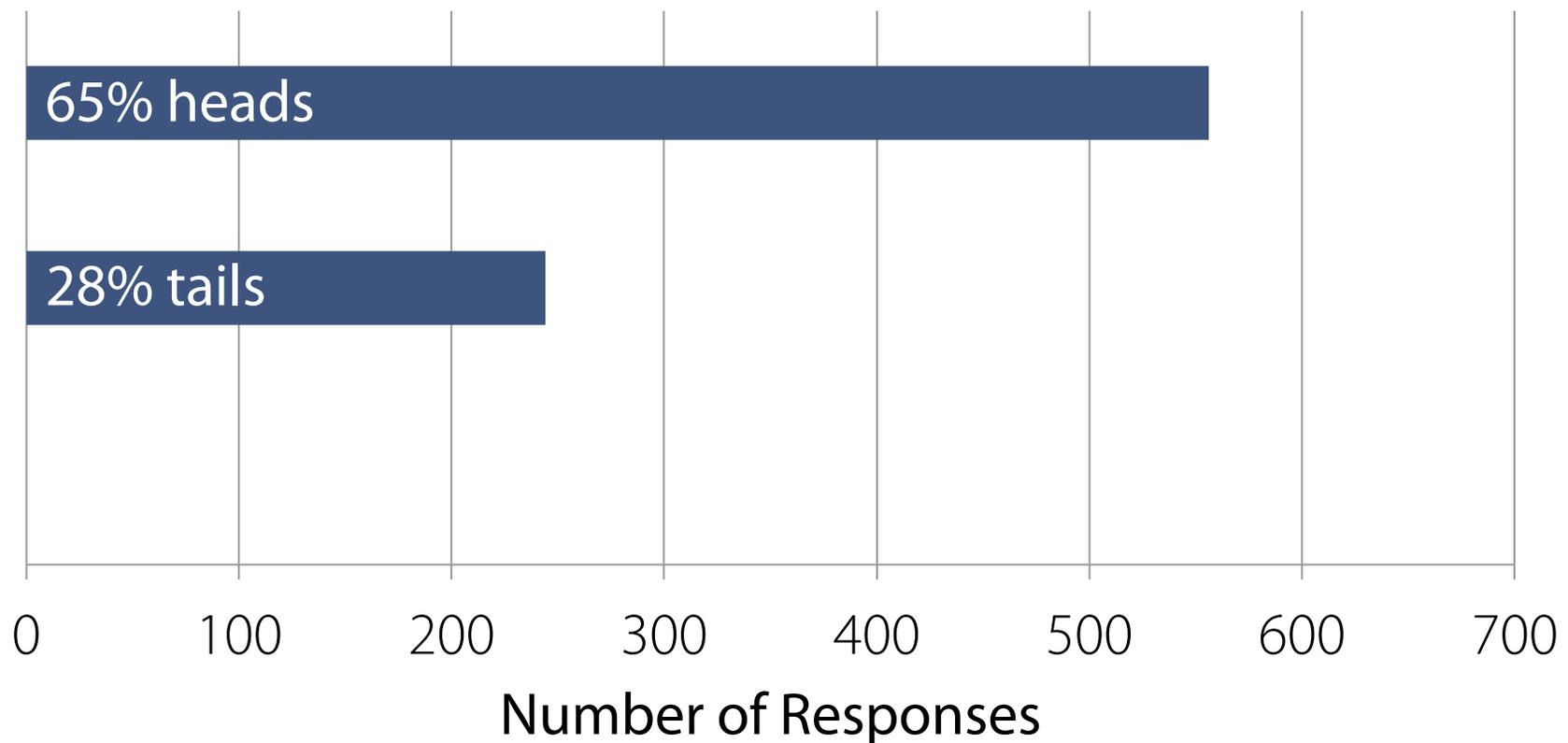


Crowd-powered system

Interactive computing system supported by human intelligence

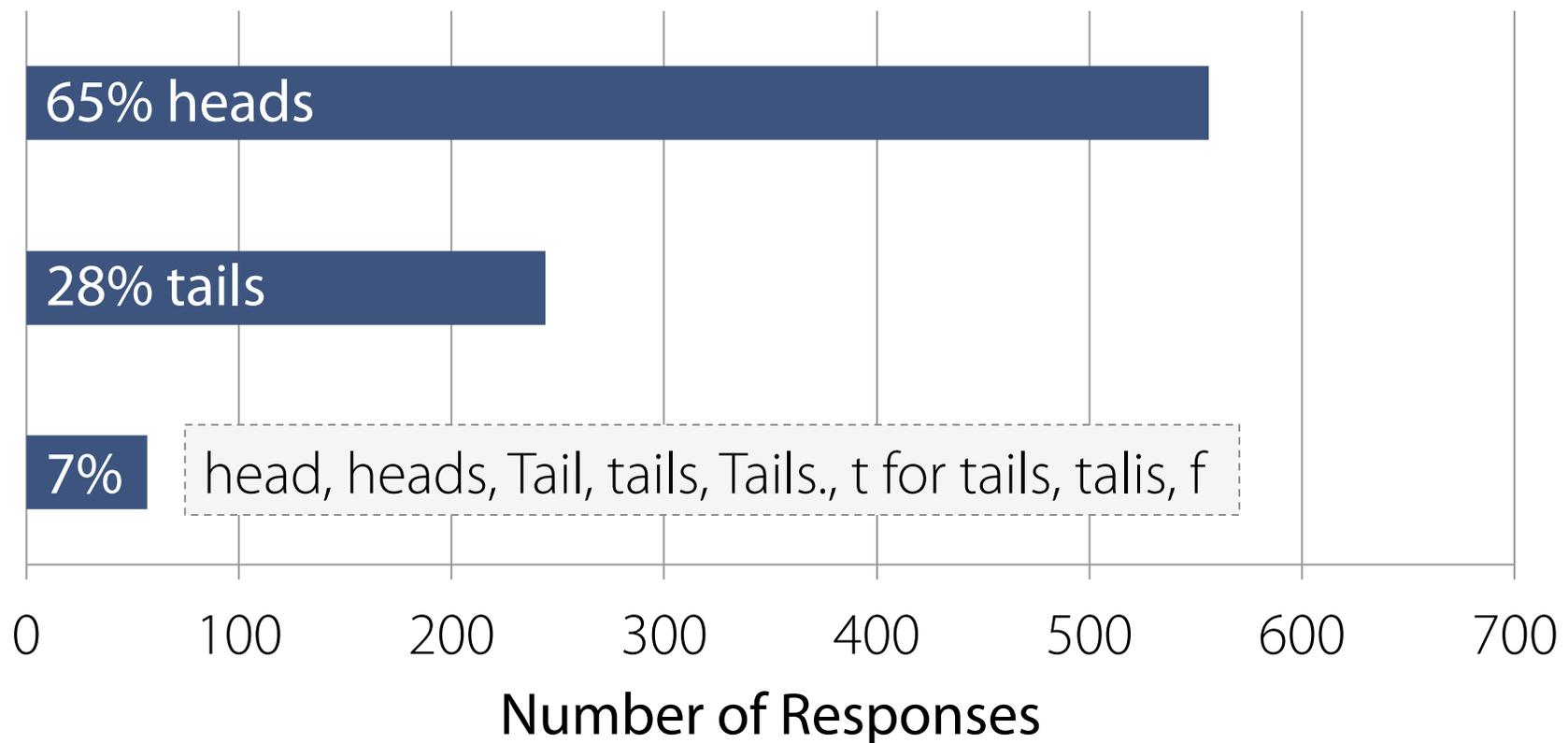
Challenge: Quality

1,000 participants on Amazon Mechanical Turk flip a coin and report "h" (heads) or "t" (tails)



Challenge: Quality

1,000 participants on Amazon Mechanical Turk flip a coin and report "h" (heads) or "t" (tails)



Challenge: Speed

Interactive applications need faster responses than crowds can provide

“User response was extremely fast”: 48 hours
[Kittur et al. 2008]

“Cheap and fast”: 190 hours
[Snow et al. 2008]

Half-life for 2.5¢ reward is 2 days,
Half-life for \$1 reward is 12 hours
[Wang et al. 2011]

**Interactive systems that
embed crowd intelligence**

**Computational techniques that
produce high-quality, fast results**

Paid Crowdsourcing

Pay small amounts of money for short tasks

Amazon Mechanical Turk: Roughly five million tasks completed per year at 1-5¢ each [Ipeirotis 2010]

Label an image

Requester: Matt C.

Reward: \$0.01

Transcribe short audio clip

Requester: Gordon L.

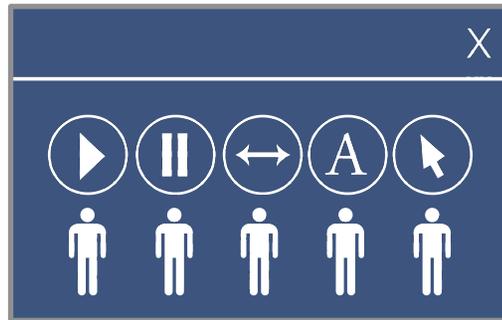
Reward: \$0.04

Population: 40% U.S., 40% India, 20% elsewhere

Gender, education and income are close mirrors of overall population distributions [Ross 2010]

Outline

1 Soylent



Word processor
with a crowd inside

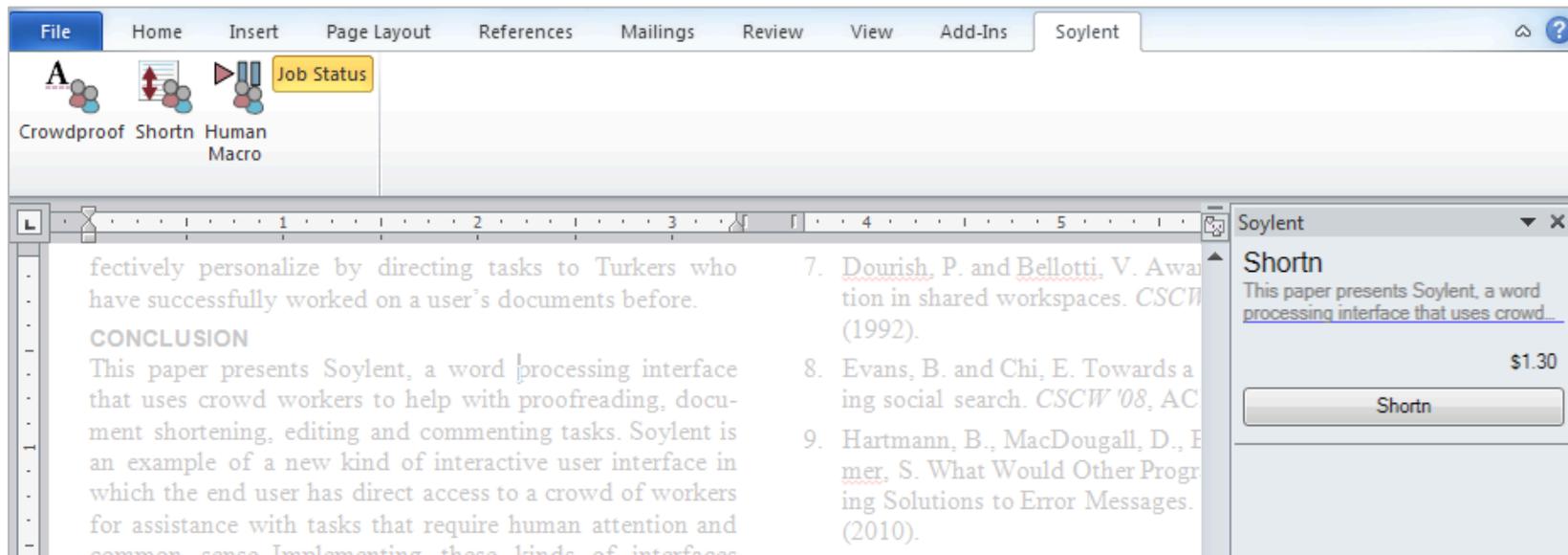
2 Adrenaline



Realtime
crowdsourcing

Soylent

Word processor that recruits crowds to aid complex writing tasks



M. Bernstein et al. Soylent: A Word Processor with a Crowd Inside. UIST 2010.

Soylent

Word processor that recruits crowds to aid complex writing tasks

Embeds crowds as first-order building blocks in a software system

Decomposes open-ended tasks via a new design pattern

demo

Challenges in Programming Crowds

Soylent has interacted with ~10,000 workers on > 2000 different tasks

Key Problem: crowd workers often produce poor output on open-ended tasks

30% Rule

**~30% of the results
in open-ended tasks
will be unsatisfactory**

Two Personas — An Example

Proofread and correct the following paragraph:

The theme of loneliness features throughout many scenes in Of Mice and Men and is often the dominant theme of sections during this story. This theme occurs during many circumstances but is not present from start to finish. In my mind for a theme to be pervasive is must be present during every element of the story. There are many themes that are present most of the way through such as sacrifice, friendship and comradship. But in my opinion there is only one theme that is present from beginning to end, this theme is pursuit of dreams.

Two Personas — An Example

Proofread and correct the following paragraph:

*The theme of loneliness features throughout many scenes in *Of Mice and Men* and is often the dominant theme of sections during this story. This theme occurs during many circumstances but is not present from start to finish. In my mind for a theme to be pervasive is must be present during every element of the story. There are many themes that are present most of the way through such as sacrifice, friendship and comradship. But in my opinion there is only one theme that is present from beginning to end, this theme is pursuit of dreams.*

Persona One: The Lazy Worker

Does as little work as necessary to be paid

The theme of loneliness features throughout many scenes in Of Mice and Men and is often the dominant theme of sections during this story. This theme occurs during many circumstances but is not present from start to finish. In my mind for a theme to be pervasive is must be present during every element of the story. There are many themes that are present most of the way through such as sacrifice, friendship and comradship. But in my opinion there is only one theme that is present from beginning to end, this theme is pursuit of dreams.

Persona One: The Lazy Worker

Does as little work as necessary to be paid

The theme of loneliness features throughout many scenes in Of Mice and Men and is often the dominant theme of sections during this story. This theme occurs during many circumstances but is not present from start to finish. In my mind for a theme to be pervasive is must be present during every element of the story. There are many themes that are present most of the way through such as sacrifice, friendship and comradeship. But in my opinion there is only one theme that is present from beginning to end, this theme is pursuit of dreams.

Persona Two: The Eager Beaver

Goes beyond task requirements to be helpful, but introduces errors in the process

The theme of loneliness features throughout many scenes in Of Mice and Men and is often the dominant theme of sections during this story. This theme occurs during many circumstances but is not present from start to finish. In my mind for a theme to be pervasive is must be present during every element of the story. There are many themes that are present most of the way through such as sacrifice, friendship and comradship. But in my opinion there is only one theme that is present from beginning to end, this theme is pursuit of dreams.

Persona Two: The Eager Beaver

Goes beyond task requirements to be helpful, but introduces errors in the process

The theme of loneliness features throughout many scenes in Of Mice and Men and is often the dominant theme of sections of this story. \n

This theme occurs during many circumstances but is not present from start to finish. \n

In my mind, for a theme to be pervasive it must be present during every element of the story. \n

There are many themes that are present most of the way through such as sacrifice, friendship and comradeship. \n

But in my opinion there is only one theme that is present from beginning to end: this theme is pursuit of dreams.

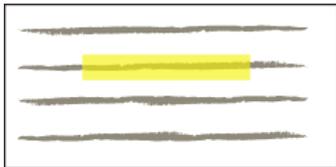
The Result: Low-quality Work

Programming with crowds today is haphazard: we lack design patterns

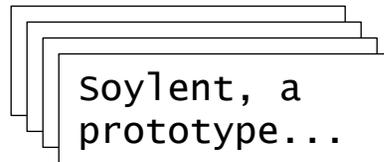
Solution: Find-Fix-Verify

Find-Fix-Verify is a design pattern for programming with crowds in open-ended tasks.

Find a problem



Fix the problem



Verify the quality of each fix

- Soylent ~~is,~~ a prototype...
- Soylent ~~is-a~~ prototypes...
- Soylent is a ~~prototypetest~~...

Find

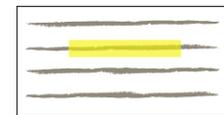
“Identify at least one area that can be shortened without changing the meaning of the paragraph.”



Independent agreement to identify patches

Fix

“Edit the highlighted section to shorten its length without changing the meaning of the paragraph.”



Soylent, a prototype...



Randomize order of suggestions

Verify

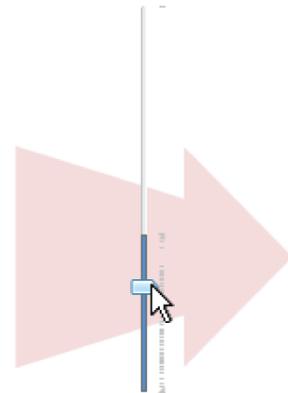
“Choose at least one rewrite that has style errors, and at least one rewrite that changes the meaning of the sentence.”

- Soylent ~~is,~~ a prototype...
- Soylent ~~is a~~ prototypes...
- Soylent is a ~~prototypetest~~...



Keep suggestions that do not get voted out

Automatic clustering generally helps separate different kinds of records that need to be edited differently, but it isn't perfect. Sometimes it creates more clusters than needed, because the differences in structure aren't important to the user's particular editing task. For example, if the user only needs to edit near the end of each line, then differences at the start of the line are largely irrelevant, and it isn't necessary to split based on those differences. Conversely, sometimes the clustering isn't fine enough, leaving heterogeneous clusters that must be edited one line at a time. One solution to this problem would be to let the user rearrange the clustering manually, perhaps using drag-and-drop to merge and split clusters. Clustering and selection generalization would also be improved by recognizing common text structure like URLs, filenames, email addresses, dates, times, etc.



Automatic clustering generally helps separate different kinds of records that need to be edited differently, but it isn't perfect. Sometimes it creates more clusters than needed, because the differences in structure aren't relevant to a specific task. | Conversely, sometimes the clustering isn't fine enough, leaving heterogeneous clusters that must be edited one line at a time. One solution to this problem would be to let the user rearrange the clustering manually using drag-and-drop edits. Clustering and selection generalization would also be improved by recognizing common text structure like URLs, filenames, email addresses, dates, times, etc.

changes the meaning of the sentence.”

Find-Fix-Verify Discussion

Why split Find and Fix?

Focus Lazy Workers on a problem of our choice
Group suggestions by core problem

Why add Verify?

Quality rises when Turkers are in productive tension

Crowds and Algorithms

[Little et al. 2010, Kittur et al. 2011, Shahaf & Horvitz 2010, Franklin et al. 2011, Marcus et al. 2011, Dai et al. 2010, Parameswaran et al. 2011, Chilton et al. 2013]

Evaluation Goals

Is Soylent's approach of crowdsourced interactive systems feasible?

- 1** How high is the quality?
- 2** How long is the delay?
- 3** How much does it cost?

Blog

Print publishers are in a tizzy over Apple's new iPad because they hope to finally be able to charge for their digital editions. But in order to get people to pay for their magazine and newspaper apps, they are going to have to offer something different that readers cannot get at the newsstand or on the open Web.

Classic HCI Paper

The metaDESK effort is part of the larger Tangible Bits project. The Tangible Bits vision paper introduced the metaDESK along with two companion platforms, the transBOARD and ambientROOM.

Draft HCI Paper

In this paper we argue that it is possible and desirable to combine the easy input affordances of text with the powerful retrieval and visualization capabilities of graphical applications. We present WenSo, a tool that uses lightweight text input to capture richly structured information for later retrieval and navigation.

Technical Writing

Figure 3 shows the pseudocode that implements this design for Lookup. FAWN-DS extracts two fields from the 160-bit key: the i low order bits of the key (the index bits) and the next 15 low order bits (the key fragment).

Rambling E-mail

A previous board member, Steve Burleigh, created our web site last year and gave me alot of ideas. For this year, I found a web site called eTeamZ that hosts web sites for sports groups. Check out our new page: [...]

Blog – 83%

Print publishers are in a tizzy over Apple's new iPad because they hope to *finally* be able to charge for their digital editions. But in order to get people to pay for their magazine and newspaper apps, they ~~are going to~~ have to offer something different that readers cannot get at the newsstand or on the open Web.

Classic HCI Paper – 87%

The metaDESK effort is part of the larger Tangible Bits project. *The Tangible Bits*

Cut 15% of original paragraph length on average.

graphical applications. We present WenSo, ~~a tool that~~ which uses lightweight text input to capture richly structured information for later retrieval and navigation.

Technical Writing – 82%

Figure 3 shows the pseudocode that implements this design for Lookup. FAWN-DS extracts two fields from the 160-bit key: ~~the i low order bits of the key~~ (the index bits) and the next 15 low order bits ~~(the key fragment)~~.

Rambling E-mail – 78%

~~A previous board member,~~ Steve Burleigh, created our web site last year and gave me a lot of ideas. ~~For this year,~~ I found a web site called eTeamZ that hosts web sites for sports groups. Check out our new page: [...]

Blog – 83%

Print publishers are in a tizzy over Apple's new iPad because they hope to *finally* be able to charge for their digital editions. But in order to get people to pay for their magazine and newspaper apps, they ~~are going to~~ have to offer something different that readers cannot get at the newsstand or on the open Web.

Focus on unnecessarily wordy phrases

But in order to get people to pay for their magazine and newspaper apps, they ~~are going to have to~~ offer something different that readers cannot get at the newsstand or on the open Web.



Technical Writing – 82%

Figure 3 shows the pseudocode that implements this design for Lookup. FAWN-DS extracts two fields from the 160-bit key: *the i low order bits of the key* (the index bits) and the next 15 low order bits *(the key fragment)*.

Rambling E-mail – 78%

~~A previous board member,~~ Steve Burleigh, created our web site last year and gave me alot of ideas. ~~For this year,~~ I found a web site called eTeamZ that hosts web sites for sports groups. Check out our new page: [...]

Blog – 83%

Print publishers are in a tizzy over Apple's new iPad because they hope to *finally* be able to charge for their digital editions. But in order to get people to pay for their magazine and newspaper apps, they ~~are going to~~ have to offer something different that readers cannot get at the newsstand or on the open Web.

Classic HCI Paper – 87%

The metaDESK effort is part of the larger Tangible Bits project. ~~The Tangible Bits vision paper, which~~ introduced the metaDESK ~~along with~~ and two companion platforms, the transBOARD and ambientROOM.

Merge sentences when patches span sentence boundaries

The metaDESK effort is part of the larger Tangible Bits project. ~~The Tangible Bits vision paper, which~~ introduced the metaDESK ~~along with~~ and two companion platforms, the transBOARD and ambientROOM.

web sites for sports groups. Check out our new page: [...]

Blog – 83%

Print publishers are in a tizzy over Apple's new iPad because they hope to *finally* be able to charge for their digital editions. But in order to get people to pay for their magazine and newspaper apps, they ~~are going to~~ have to offer something different that readers cannot get at the newsstand or on the open Web.

Classic HCI Paper – 87%

The metaDESK effort is part of the larger Tangible Bits project. ~~The Tangible Bits vision paper, which~~ introduced the metaDESK ~~along with~~ and two companion platforms, the transBOARD and ambientROOM.

Draft HCI Paper – 90%

~~In this paper we argue that~~ it is possible and desirable to combine the easy input

Introduced style errors when workers were not part of the community of practice

~~In this paper we argue that~~ it is possible and desirable to combine the easy input affordances of text with the powerful retrieval and visualization capabilities of graphical applications.

Blog – 83%

Print publishers are in a tizzy over Apple's new iPad because they hope to *finally* be able to charge for their digital editions. But in order to get people to pay for

Parallelism can introduce inconsistent changes

FAWN-DS extracts two fields from the 160-bit key: ~~the i low order bits of the key~~ (the index bits) and the next 15 low order bits ~~(the key fragment)~~.

affordances of text with the powerful retrieval and visualization capabilities of graphical applications. We present WenSo, ~~a tool that~~ which uses lightweight text input to capture richly structured information for later retrieval and navigation.

Technical Writing – 82%

Figure 3 shows the pseudocode that implements this design for Lookup. FAWN-DS extracts two fields from the 160-bit key: ~~the i low order bits of the key~~ (the index bits) and the next 15 low order bits ~~(the key fragment)~~.

Rambling E-mail – 78%

~~A previous board member,~~ Steve Burleigh, created our web site last year and gave me a lot of ideas. ~~For this year,~~ I found a web site called eTeamZ that hosts web sites for sports groups. Check out our new page: [...]

Blog – 83%**3 para., 158 people, \$1.52/para**

Print publishers are in a tizzy over Apple's new iPad because they hope to ~~finally~~ be able to charge for their digital editions. But in order to get people to pay for their magazine and newspaper apps, they ~~are going to~~ have to offer something different that readers cannot get at the newsstand or on the open Web.

Classic HCI Paper – 87%**7 para., 264 people, \$1.06/para**

The metaDESK effort is part of the larger Tangible Bits project. ~~The Tangible Bits vision paper,~~ which introduced the metaDESK ~~along with~~ and two companion platforms, the transBOARD and ambientROOM.

Draft HCI Paper – 90%**5 para., 284 people, \$1.49/para**

~~In this paper we argue that~~ it is possible and desirable to combine the easy input affordances of text with the powerful retrieval and visualization capabilities of graphical applications. We present WenSo, ~~a tool that~~ which uses lightweight text input to capture richly structured information for later retrieval and navigation.

Technical Writing – 82%**3 para., 188 people, \$1.61/para**

Figure 3 shows the pseudocode that implements this design for Lookup. FAWN-DS extracts two fields from the 160-bit key: ~~the i low order bits of the key~~ (the index bits) and the next 15 low order bits (~~the key fragment~~).

Rambling E-mail – 78%**6 para., 362 people, \$1.62/para**

~~A previous board member,~~ Steve Burleigh, created our web site last year and gave me a lot of ideas. ~~For this year,~~ I found a web site called eTeamZ that hosts web sites for sports groups. Check out our new page: [...]

How Fast Is Shortn?

Soylent
Posts Task



Worker
Accepts Task



Worker
Submits Task

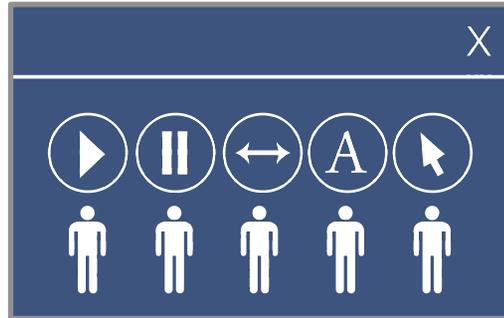
Wait time is the longest:
Median 18.5 minutes

Summed medians across Find, Fix and Verify
 $Q_1=8.3$ minutes, $Q_3=41.6$ minutes

Actual work time is shorter:
Median 2.0 minutes

Summed medians across Find, Fix and Verify
 $Q_1=60$ seconds, $Q_3=3.6$ minutes

Soylent



Word processor
with a crowd inside

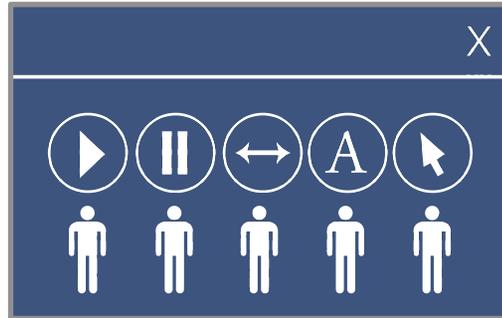
New class of paid, on-demand
crowd-powered systems

Find-Fix-Verify design pattern

Lazy Worker and Eager Beaver

Outline

1 Soylent



Word processor
with a crowd inside

2 Adrenaline



Realtime
crowdsourcing

Applications are constrained by crowd latency.

Design [Yu and Nickerson 2011, Xu and Bailey 2011]

Health and nutrition [Noronha et al. 2011]

Open-world databases [Franklin et al. 2011, Marcus et al. 2010]

Crowd algorithms [Little et al. 2010, Parameswaran et al. 2011]

Assistive technology [Bigham et al. 2010]

Robotics [Sorokin et al. 2010, Lasecki et al. 2011]

Maps [Stranders et al. 2011]

Task decomposition [Kulkarni et al. 2012]

Machine vision [Rodriguez and Davis 2011, Yan et al. 2010]

Feedback and collaboration [Kittur 2010, Dow et al. 2012]

One unverified response
in 56 seconds

[Bigham et al. 2010]

— but —

The user loses focus
after 10 seconds

[Nielsen 1993, Card et al. 1991]

**Our goal is on-demand,
realtime crowds.**

Adrenaline

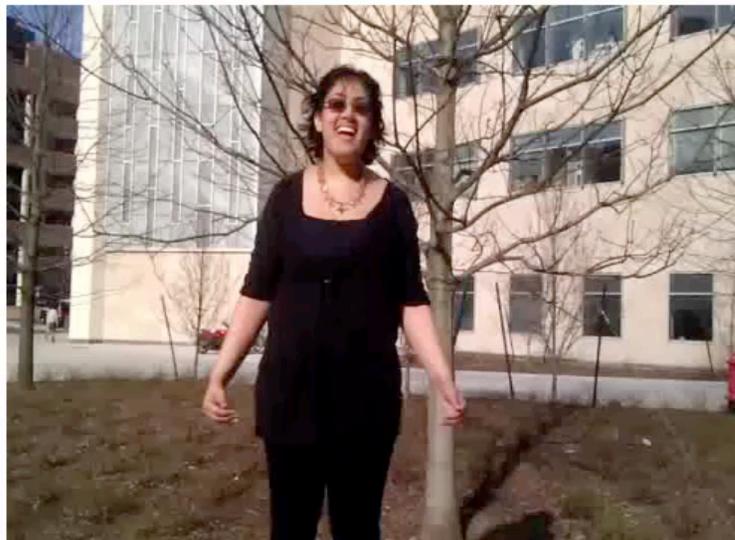
Realtime crowd-powered camera



M. Bernstein, J. Brandt, R. Miller, and D. Karger. Crowds in Two Seconds: Enabling Realtime Crowd-Powered Interfaces. UIST 2011.







**How do we
recruit crowds quickly?**

**Approach:
Retainer model**

Retainer Model

Workers sign up in advance

Offer ½¢ per minute to remain on call

Alert when task is ready



Task:

Move the playback head to find the best moment.

Wait at most:

5 minutes



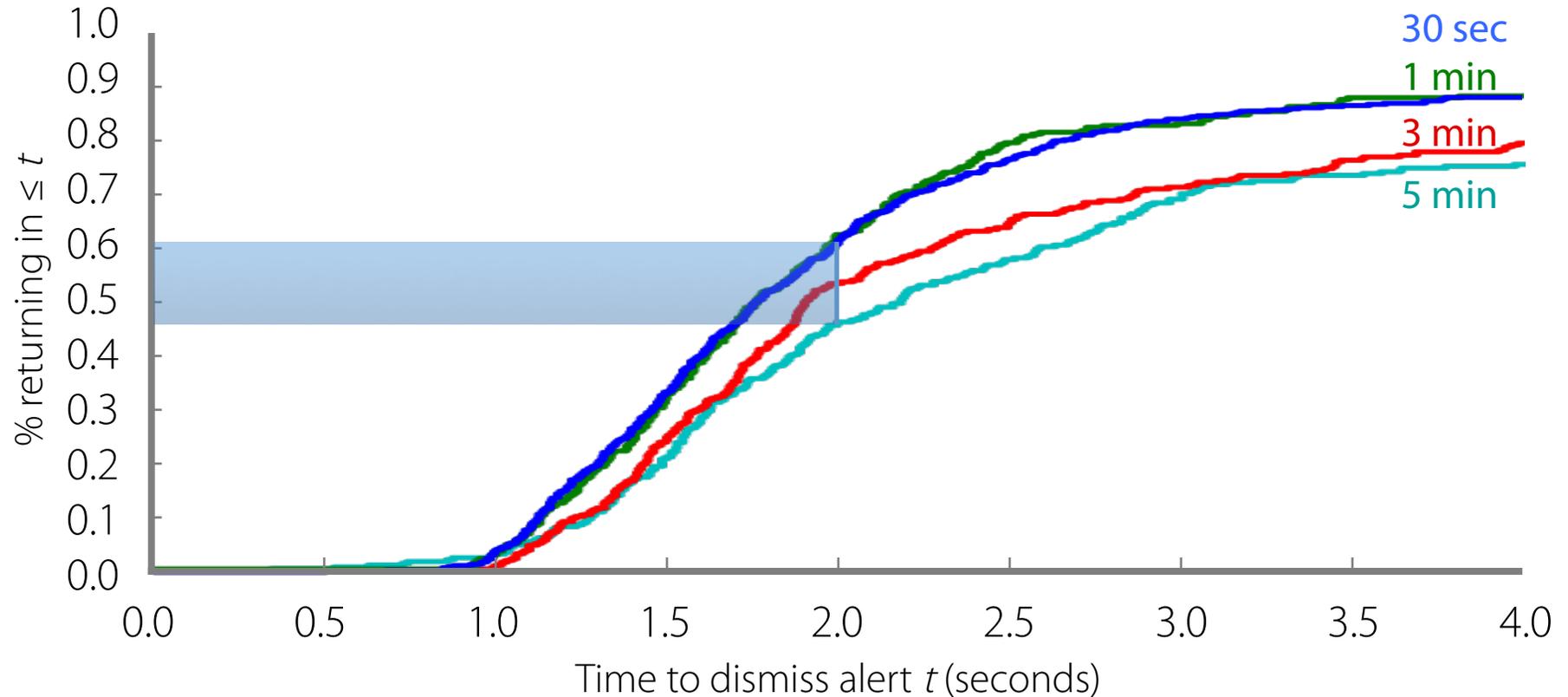
alert()

Start now!

OK

Results: N=1545 tasks

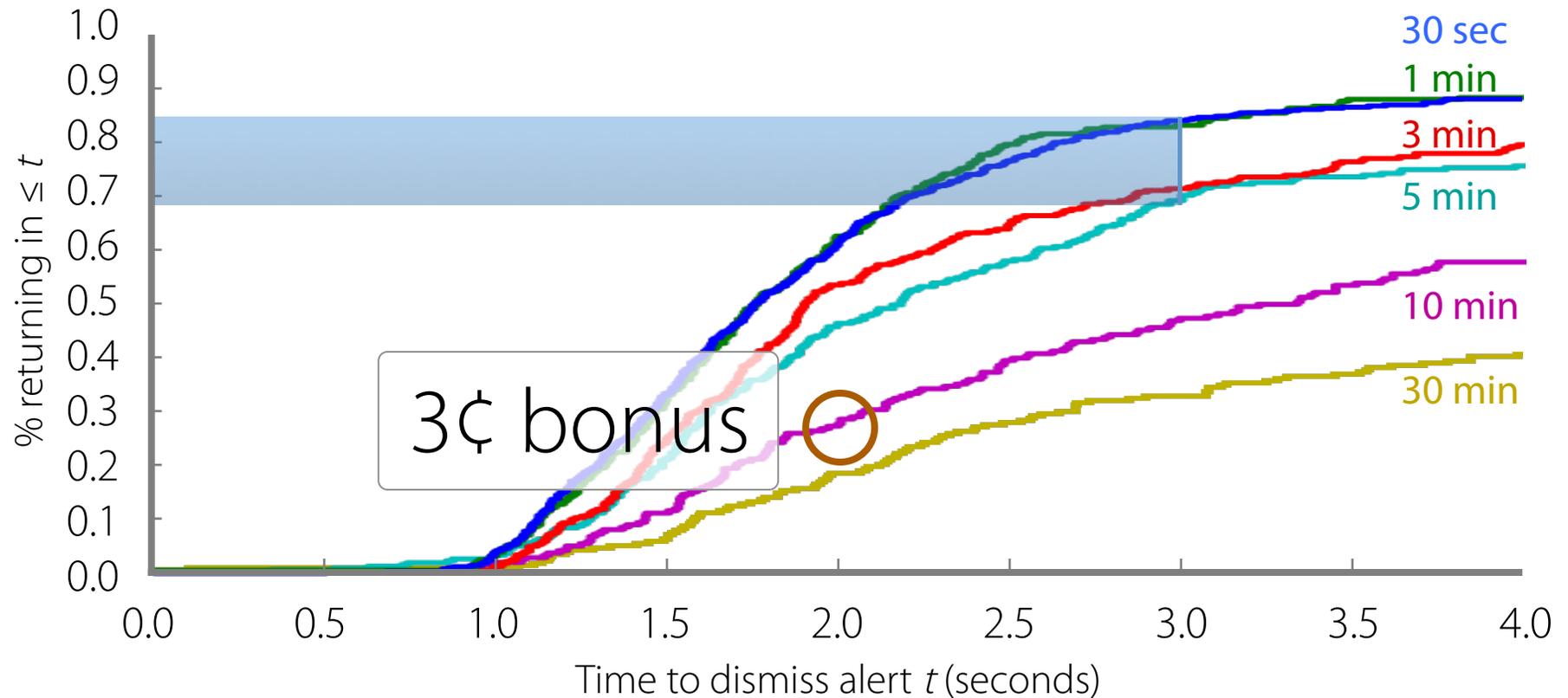
How quickly do retainer workers return?



For retainer times under ten minutes,
46–61% within **2 seconds**.

Results: N=1545 tasks

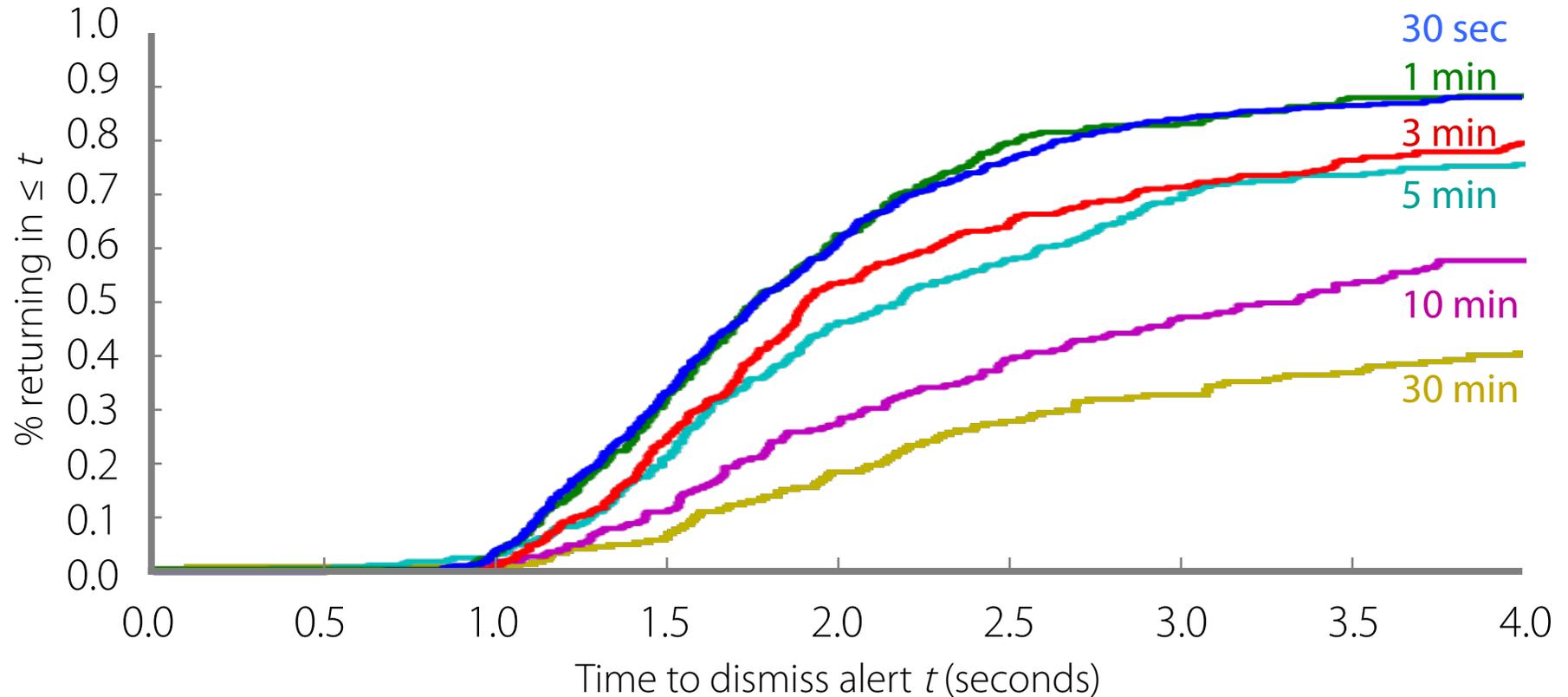
How quickly do retainer workers return?



For retainer times under ten minutes,
69–84% within **3 seconds**.

Results: N=1545 tasks

How quickly do retainer workers return?



One worker on retainer costs **\$0.30 / hour.**

A|B: Instant Votes

Five votes in five seconds:

“Which font should I use?”

“Which tie matches better?”

“Which blog headline is catchier?”



A|B: Instant Votes

Five votes in five seconds:

“Which font should I use?”

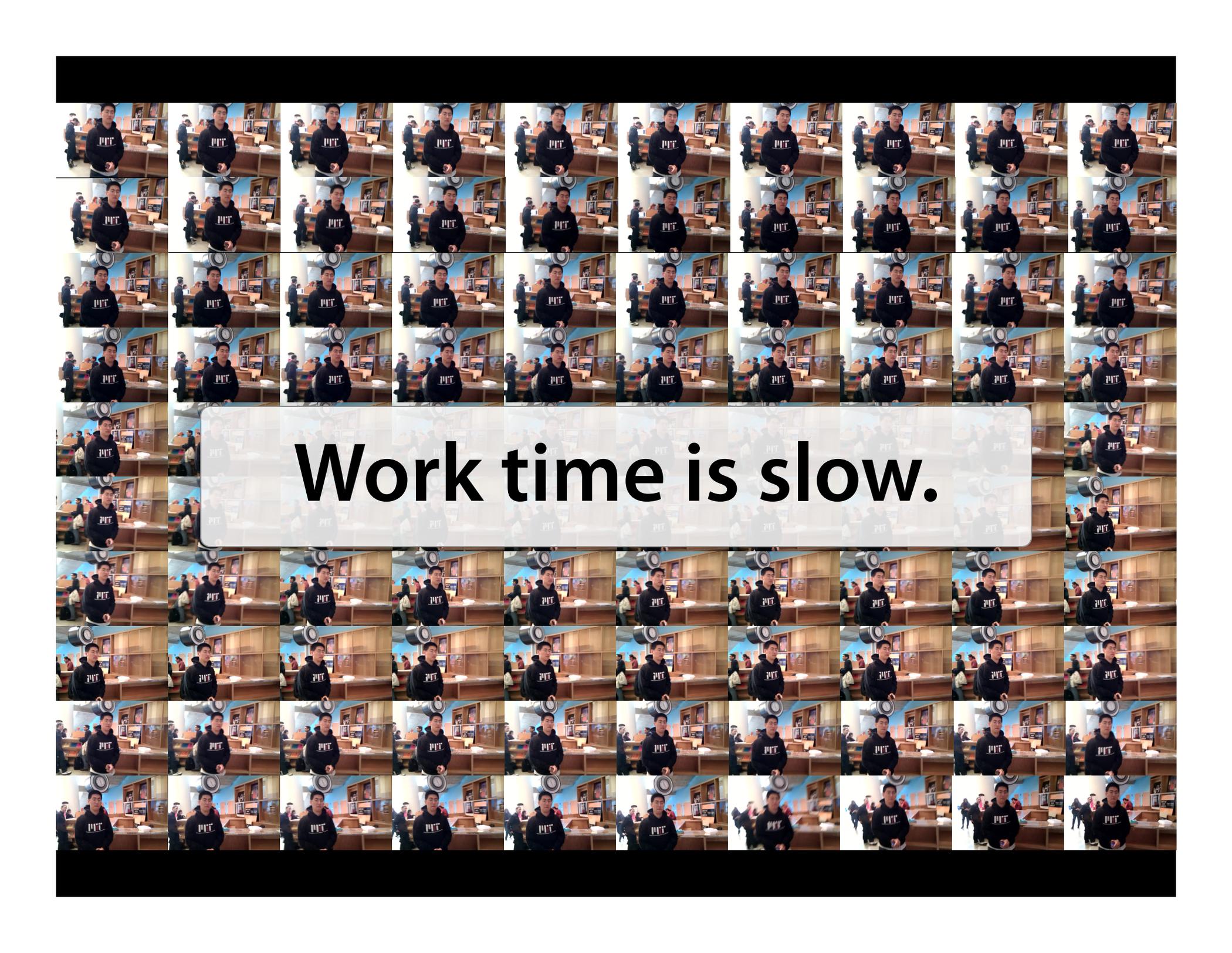
“Which tie matches better?”

“Which blog headline is catchier?”



*The retainer model:
crowds in two seconds
and votes in five seconds.*



The image is a dense grid of small video frames, likely from a time-lapse or a sequence of shots. Each frame shows a person wearing a black hoodie with 'MIT' written on it, sitting at a desk in a library or study area. The person is holding a silver metal bucket on their head. The background shows bookshelves and other people in the distance. In the center of the grid, there is a large, semi-transparent white rectangular area containing the text 'Work time is slow.' in a bold, black, sans-serif font.

Work time is slow.

**How do we overcome
slow work times?**

How do we overcome slow work times?

Synchronous crowds

Crowds can be faster than any individual member

Five black human silhouettes standing in a row, representing a crowd.

Rapid Refinement

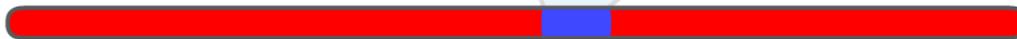
Recognize potential agreement early, then use it to reduce a continuous search space quickly.



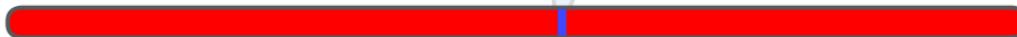
↓ Phase 2



↓ Phase 3



↓ Final Photo



Rapid Refinement



Worker 1



Worker 2



Worker 3

Rapid Refinement



Worker 1

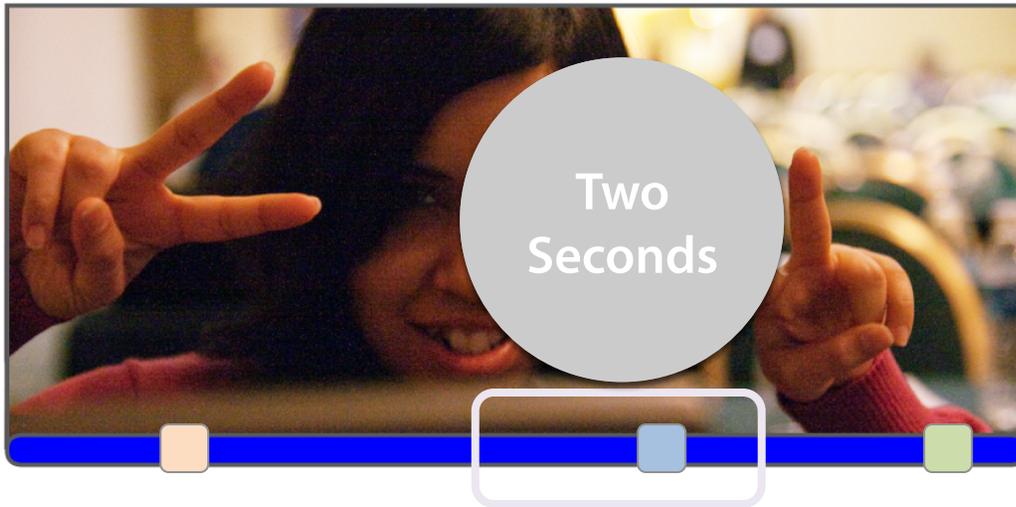


Worker 2



Worker 3

Rapid Refinement



```
while (searchArea.size > 1):  
    a = calculateAgreement(workerPositions,  
                           searchArea,  
                           refinementRatio)
```

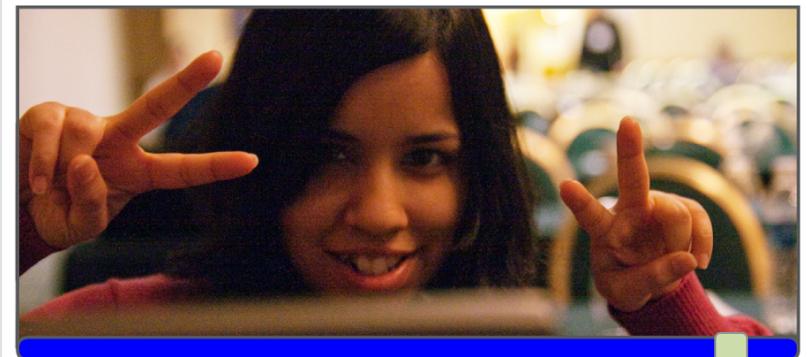
```
    if (a.percent >= 0.66):
```



Worker 1



Worker 2



Worker 3

Rapid Refinement



↓ Phase 2



↓ Phase 3



↓ Final Photo



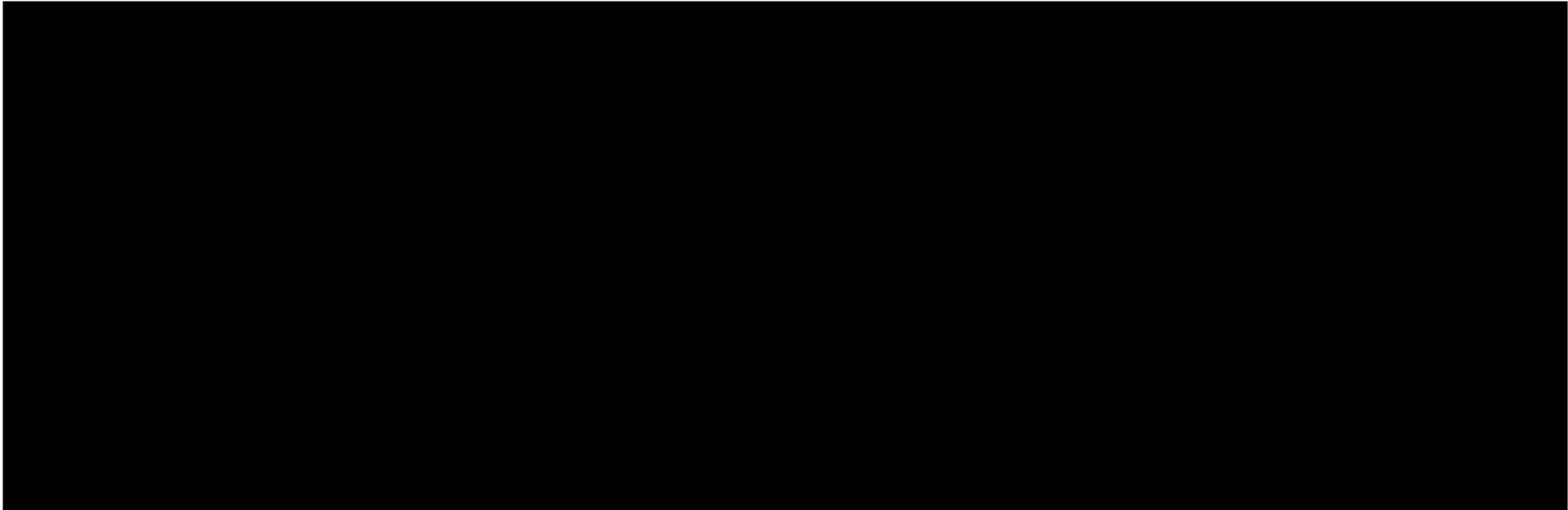
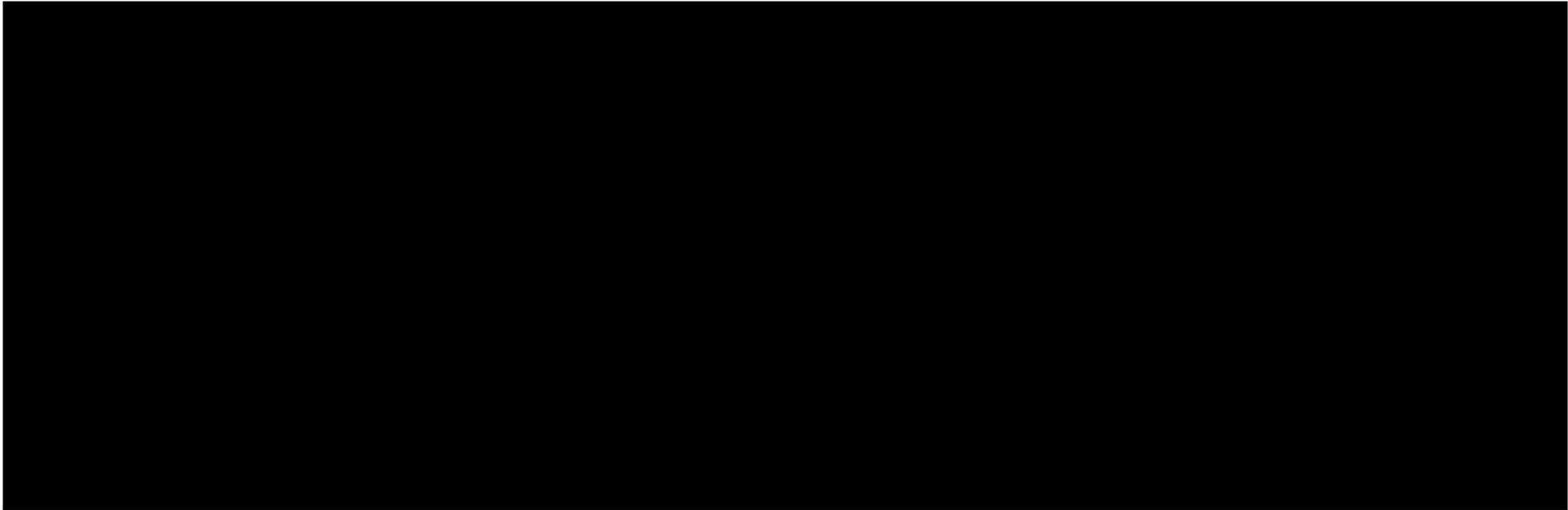
Worker 1



Worker 2



Worker 3



Evaluation

Do the retainer model and rapid refinement produce realtime results?

Crowdsourcing approaches:

- 1** Rapid Refinement
- 2** Generate-and-Vote
- 3** Generate-One

Rapid Refinement

$\mu=5.8, \sigma=2.2$

Computer Vision

$\mu=4.9, \sigma=2.2$

Photographer

$\mu=6.4, \sigma=2.3$



9 point Likert scale on self-rated quality

ANOVA $p < .001$

Rapid Refinement

$\mu=5.8, \sigma=2.2$

Computer Vision

$\mu=4.9, \sigma=2.2$

Photographer

$\mu=6.4, \sigma=2.3$

Good Photos



Typical Photos



Rapid Refinement

$\mu=5.8, \sigma=2.2$

Computer Vision

$\mu=4.9, \sigma=2.2$

Photographer

$\mu=6.4, \sigma=2.3$

Good Photos



Typical Photos



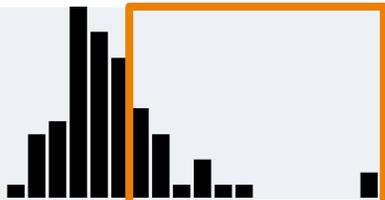
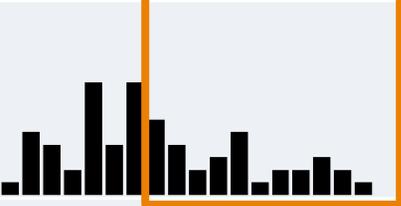
Bad Photos



Agreement misfire

Results

Rapid Refinement Fastest, with Smallest Time Variance

Algorithm	Histogram of Execution Times	N=72 photos
Rapid Refinement		$\mu=12.6$ sec, $\sigma=2.2$ sec 22¢
Generate One		$\mu=16.3$ sec, $\sigma=9.8$ sec 22¢
Generate and Vote		$\mu=45.3$ sec, $\sigma=14.0$ sec 53¢

ANOVA with pairwise posthoc tests $p < .05$

Tradeoffs in Rapid Refinement

Strengths:

- Quick preliminary results (10 sec)
- Combines work and verification

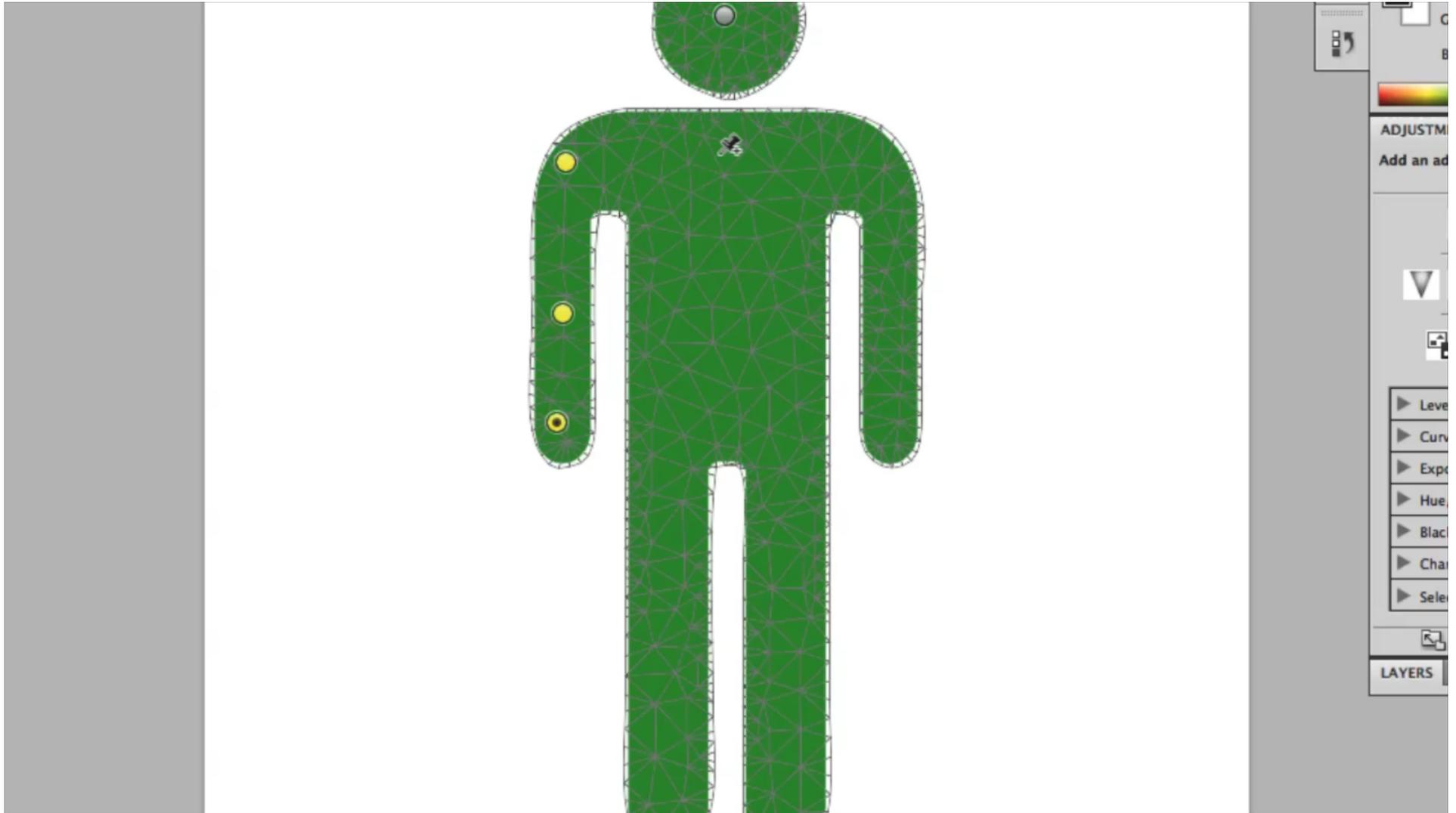
Weaknesses:

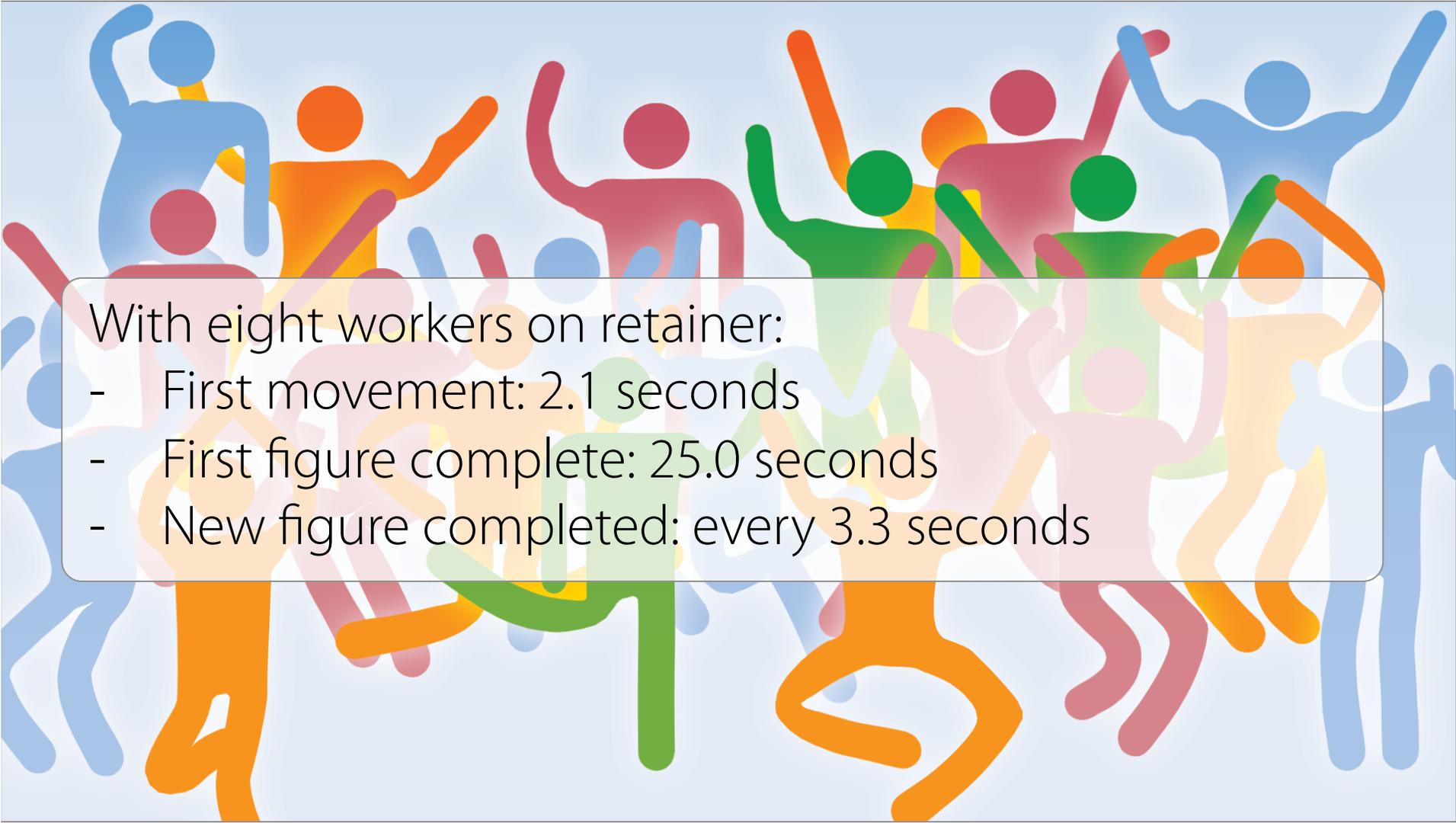
- Sacrifices quality for speed
- Stifles individual creativity

Generalizability:

- Any continuous search space
(e.g., parameter tuning)

*The retainer model
and rapid refinement
execute large searches
in roughly ten seconds.*





With eight workers on retainer:

- First movement: 2.1 seconds
- First figure complete: 25.0 seconds
- New figure completed: every 3.3 seconds

Mathematical modeling to optimize realtime crowdsourcing

M. Bernstein et al. Analytic Methods for Optimizing Realtime Crowdsourcing.
Collective Intelligence 2012.

Queueing Theory Model

Cast the retainer model as an $M/M/c/c$ queue

Formal framework for understanding arrival and service processes with c servers and Poisson arrival rates

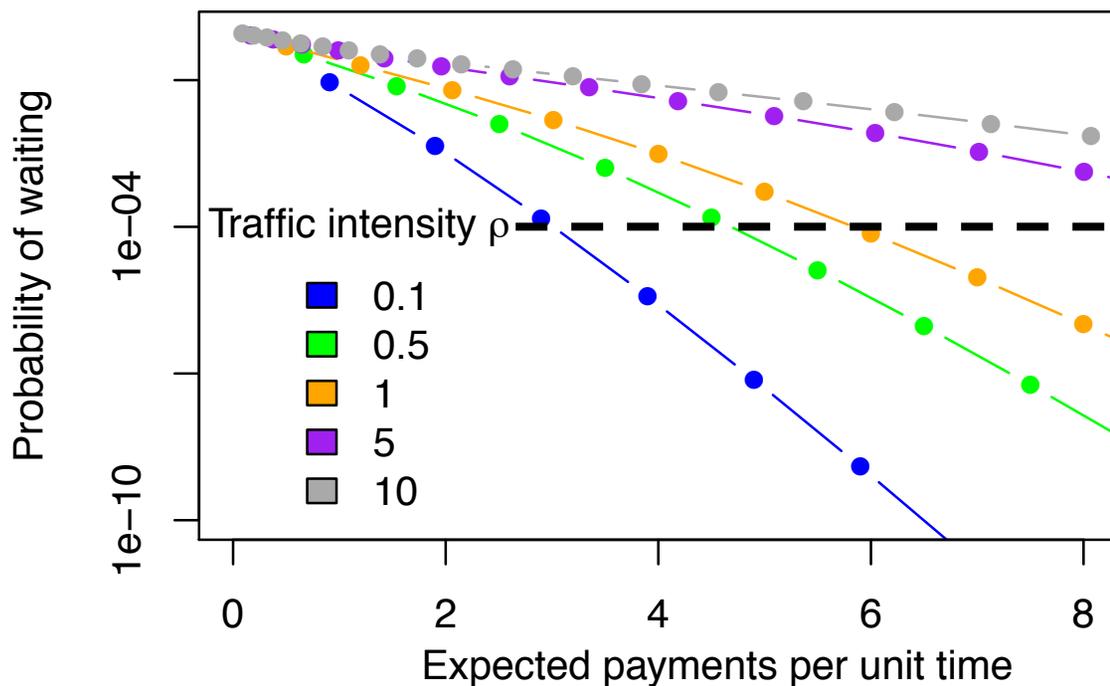
Worker recruitment rate λ , task arrival rate μ ,
traffic intensity $\rho = \lambda / \mu$

Probability of non-realtime service
with c workers on retainer, $\pi(c)$

$$\pi(c) = \frac{\rho^c / c!}{\sum_{i=0}^c \rho^i / i!} \quad \text{Cost} = c - \rho(1 - \pi(c))$$

Queueing Theory

Optimizing Realtime Crowdsourcing



Minimize c such that $\pi(c) \leq p_{max}$

Median feedback in 0.50 seconds (3x improvement)

Other benefits: globally shared retainer pool, task routing, predictive recruitment

Adrenaline



Realtime
crowdsourcing

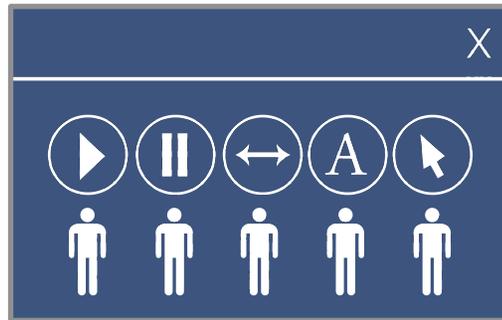
Enabling interactive
crowd-powered systems

Techniques for fast,
synchronous crowds:

- 1** Retainer Model
- 2** Rapid Refinement

Outline

1 Soylent



Word processor
with a crowd inside

2 Adrenaline



Realtime
crowdsourcing

Tail Answers

Direct results for queries in the long tail

molasses substitutes



Substitute for molasses

Replace one cup of molasses with one of the following: 1 cup dark corn syrup, honey or maple syrup; 3/4 cup firmly packed brown sugar or 3/4 cup granulated sugar, plus 1/4 cup water.

Source: <http://frugalliving.about.com/od/makeyour/qt/Molasses/Sub.htm>

[Molasses Substitute Recipe](http://frugalliving.about.com/od/.../qt/Molasses_Sub.htm)

frugalliving.about.com/od/.../qt/Molasses_Sub.htm

Note: These **substitutions** may alter the taste of your recipe a bit. If the **molasses** flavor is vital to the success of your recipe, try the brown sugar **substitute**.

[Molasses Substitutions, Measures, Tips and Cooking Hints](http://homecooking.about.com/od/specificfood/a/molassestips.htm)

homecooking.about.com/od/specificfood/a/molassestips.htm

Green Apple Calories

There are approximately 35 calories in a green apple.

Source: <http://www.livestrong.com/thedailyplate/nutrition->

Inventor of First Light Bulb

The first electric light was made in 1800 by Humphry Davy, an English scientist. He experimented with electricity and invented an electric battery. When he connected wires to his battery and a piece of carbon, the carbon glowed, producing light. This is called an electric arc.

Source: <http://www.enchantedlearning.com/inventors/edison/lightbulb.shtml>

Substitute for molasses

Replace one cup of molasses with one of the following: 1 cup dark corn syrup, honey or maple syrup; 3/4 cup firmly packed brown sugar or 3/4 cup granulated sugar, plus 1/4 cup water.

Source: <http://frugalliving.about.com/od/makeyour/qt/Molasses/Sub.htm>

Disolvable Stitches

It typically takes at minimum one week for the suture to dissolve, i.e. be absorbed by the body.

Source: <http://answers.yahoo.com/question/index?>

How to Mute Audio on Windows Movie Maker

On the Audio or Audio/Music track of the timeline, click the audio clip that you want to mute. To select multiple clips, press and hold down the CTRL key as you click clips. Click Clip, point to Audio, and then click Mute.

Source: <http://windows.microsoft.com/en-US/windows-vista/Adjusting-audio->

IRS Milage

The IRS allows reimbursement for business miles driven at a rate of for 51 cents per mile.

Source: <http://www.irs.gov/newsroom/article/0,,id=232017,00.html>

How to Turn Up Volume on Your Computer

Start>All Programs>Accessories>Entertainment>Volume Control>Wave Setting. Inceze it and the Volume should go higher.

Source: <http://answers.yahoo.com/question/index?>

Fish Frying Temperature

350 degrees for 3 minutes is the ticket! Also, make sure to put just enough fillets in the basket to cover the bottom of it.

Source: <http://www.walleycentral.com/forums/showthread.php?t=146552>

Area Code 407

Area code 407 is the area code for the Orlando metro area including all of Orange, Osceola, and Seminole counties, as well as small portions of Volusia and Lake counties.

Source: http://en.wikipedia.org/wiki/Area_code_407

Ireland Currency

Euro (EUR)

Source: <http://www.greenwichmeantime.com/time-zone/europe/european->

New York City Sales Tax 2010

New York City sales tax rate is 8.875%

Source: <http://ny.rand.org/stats/govtfin/salestax.html>

Integrate social and crowd intelligence as core parts of interaction, software, and computation.

The Future of Crowd Work

Cyber-Taylorism and the crowd worker as API call

Embed human-human contract ethics

Expected wages and living wages

Future of education, reputation, and promotion

Would you be proud of your own child if they decided to do full-time crowd work?

How would you design a crowd work platform?

N. Kittur, J. Nickerson, M. Bernstein, et al. The Future of Crowd Work.
CSCW 2013.

Crowd-powered systems enable experiences that neither crowd nor machine intelligence can support alone.

Computation will be critical to the wisdom of crowds.

Collaborators

Rob Miller and David Karger, Björn Hartmann, Desney Tan, Eric Horvitz, Greg Little, Joel Brandt, Katrina Panovich, Mark Ackerman, Mary Czerwinski

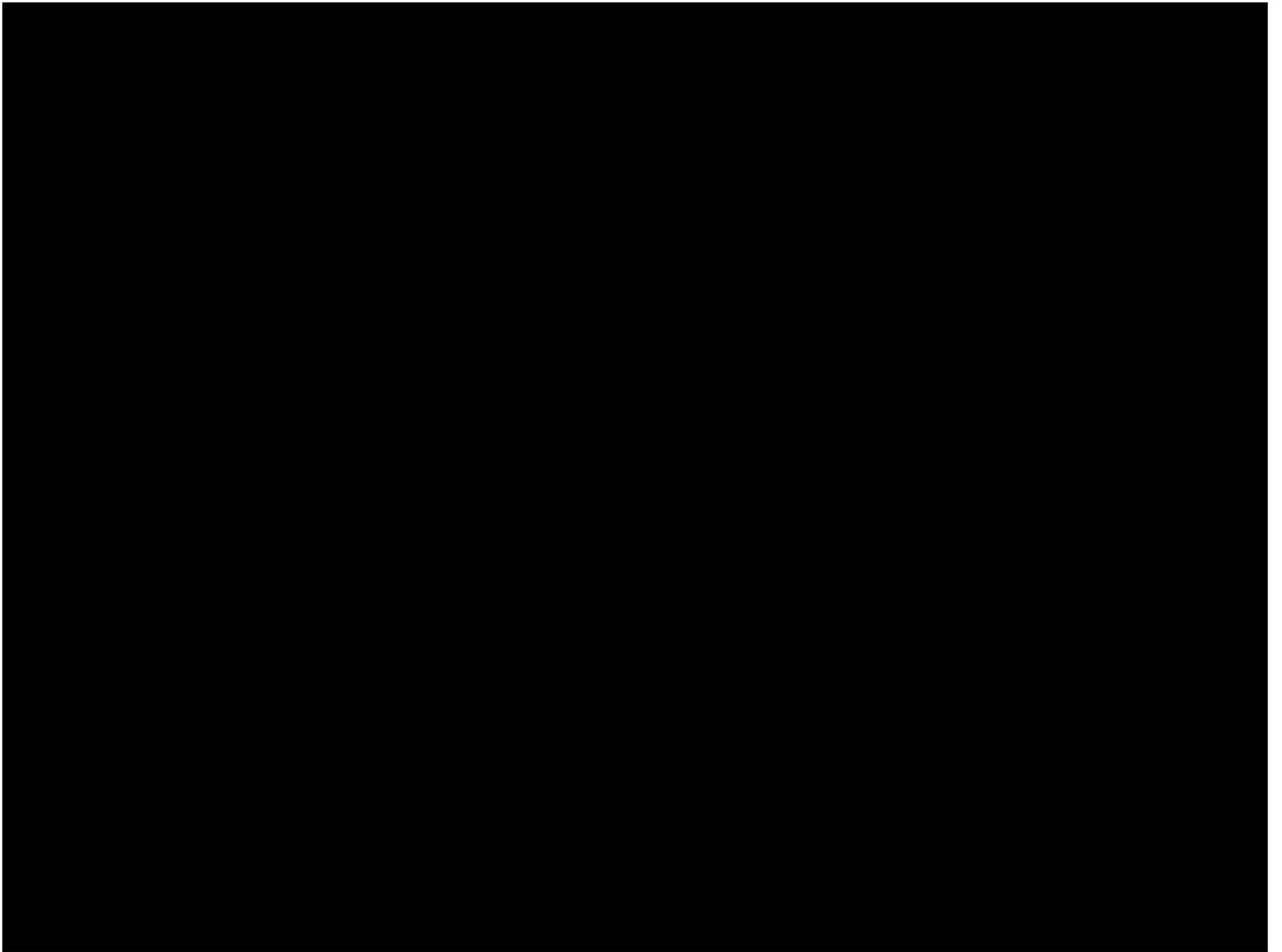
Justin Cheng, Ethan Fast, Joy Kim

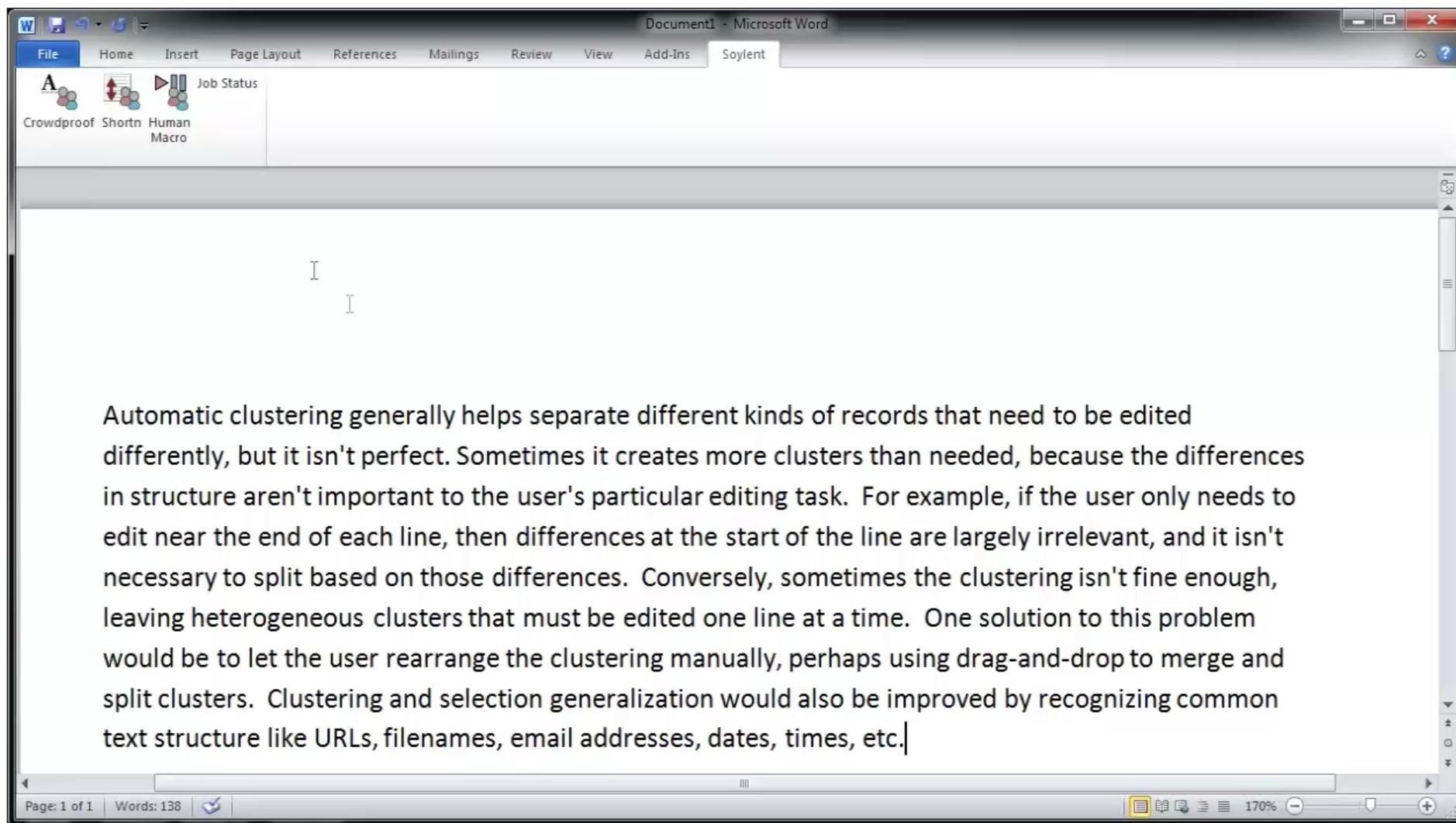
Image credit: hristine Daniloff (MIT News Office), Creative Commons: auntiep, jeffwilcox, jmpk, ebriel, jwl, takuhitosotome, d!zzy

Crowd-Powered Systems

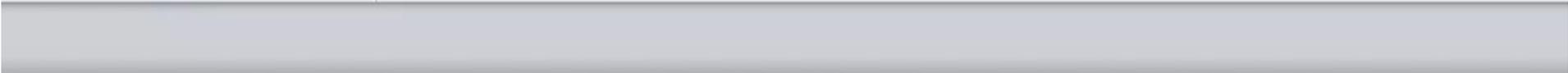


Michael Bernstein
<http://hci.stanford.edu/msb>

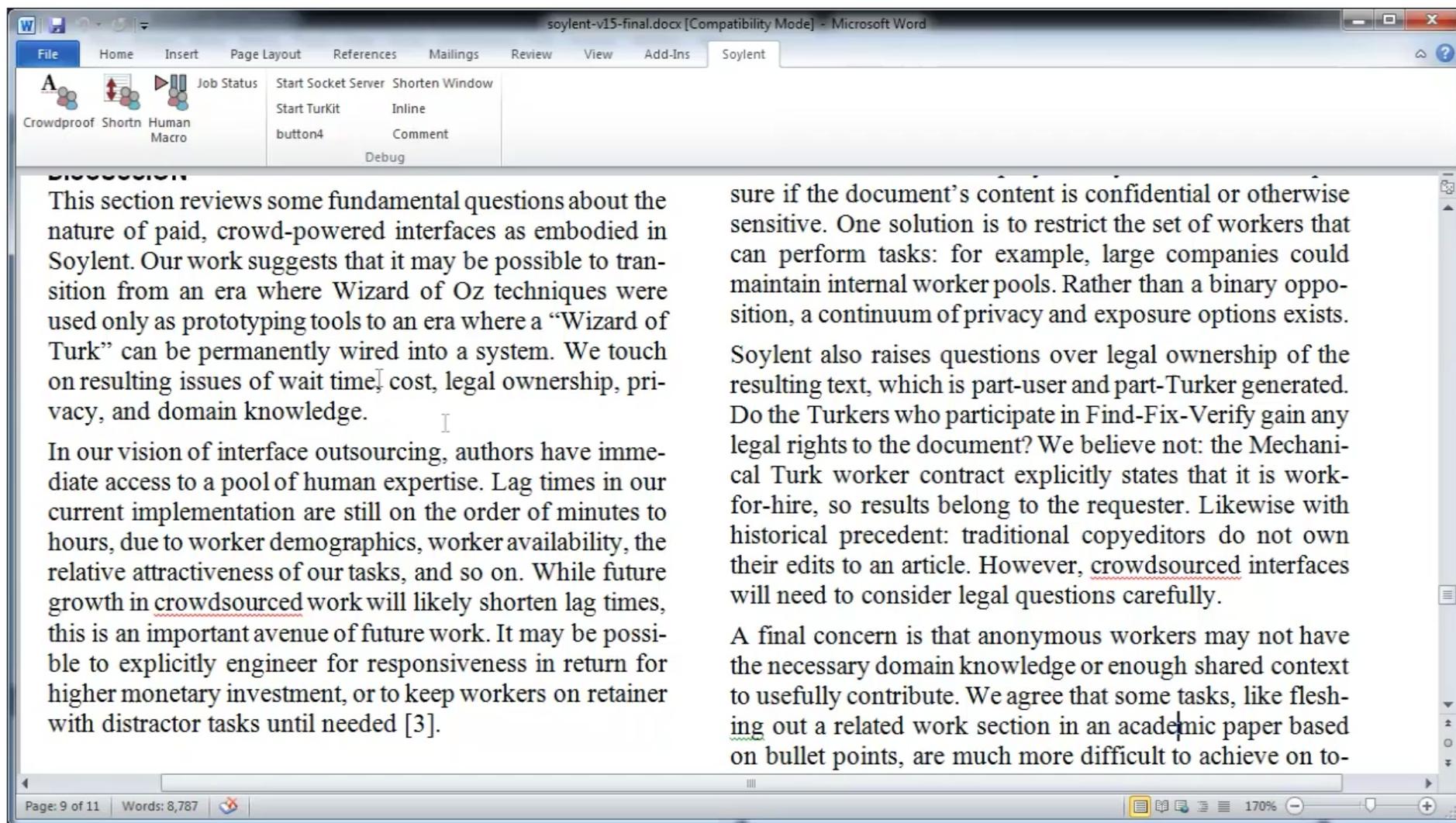




Automatic clustering generally helps separate different kinds of records that need to be edited differently, but it isn't perfect. Sometimes it creates more clusters than needed, because the differences in structure aren't important to the user's particular editing task. For example, if the user only needs to edit near the end of each line, then differences at the start of the line are largely irrelevant, and it isn't necessary to split based on those differences. Conversely, sometimes the clustering isn't fine enough, leaving heterogeneous clusters that must be edited one line at a time. One solution to this problem would be to let the user rearrange the clustering manually, perhaps using drag-and-drop to merge and split clusters. Clustering and selection generalization would also be improved by recognizing common text structure like URLs, filenames, email addresses, dates, times, etc.

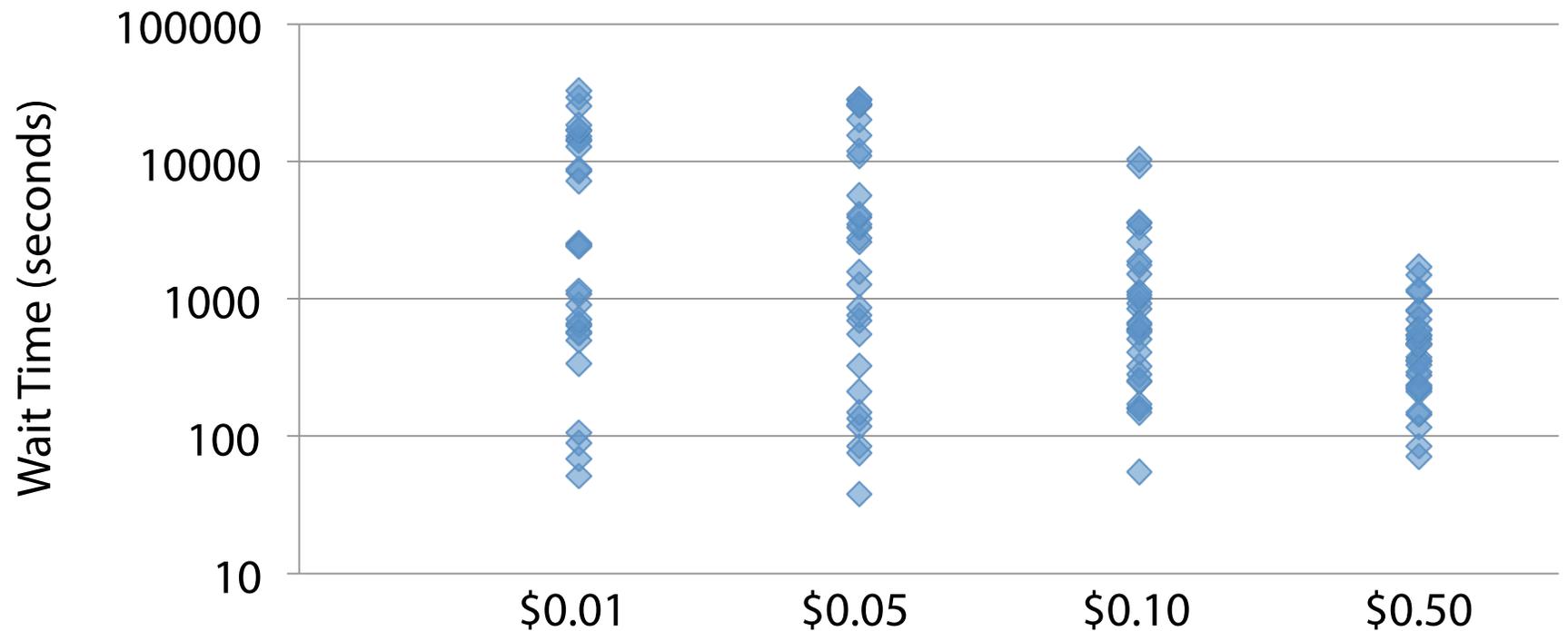


GUIs made using computers be more intuitive and easier to learn, it didn't let computers efficiently. Masses only can use the software unless they know how to write programs. In other words, if one who knows nothing needs to click through 100 buttons to complete her job everyday, the only thing is to click through those buttons by hand every time. But if she happens to be a coder, there is a little chance that she can write a program to automate everything. Why? In fact, each GUI application is a big black box, which usually have no out



Effect of Price on Wait Time

Paying more had no effect on early arrivals, but sped up the latecomers



Results: Cost

\$0.08 per Find, \$0.05 per Fix, and \$0.04 per Verify

Average paragraph cost \$1.41 to Shortn:

\$0.55 to Find an average of two patches

\$0.48 to Fix each patch

\$0.38 to Verify each patch

Lower bound with \$0.01 per task:

\$0.30 per paragraph

Retainer Experiment

1545 tasks from 280 workers

He leapt the fence and dashed toward the door.

Manipulate retainer time:

{0.5, 1, 2, 5, 10, 30} minutes

Sample wait time from [0, retainer time]

Measure: time to dismiss the alert

alert()

Start now!

OK

Retainer Design Experiment

Four designs:

1 Baseline
(no alert)



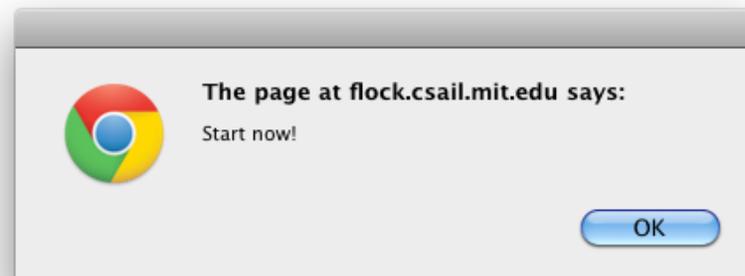
Retainer Design Experiment

Four designs:

1 Baseline
(no alert)

2 Alert

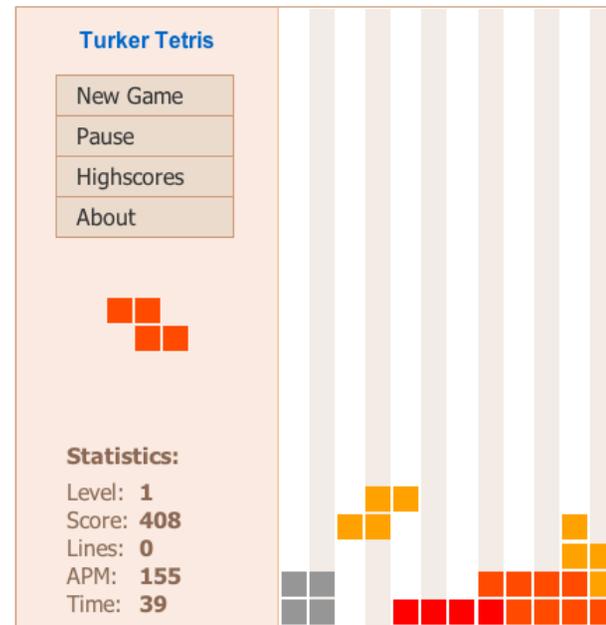
```
playAudio("alert_chime.mp3");  
alert("Start now!");
```



Retainer Design Experiment

Four designs:

- 1 Baseline
(no alert)
- 2 Alert
- 3 Game



Retainer Design Experiment

Four designs:

1 Baseline
(no alert)

2 Alert

3 Game

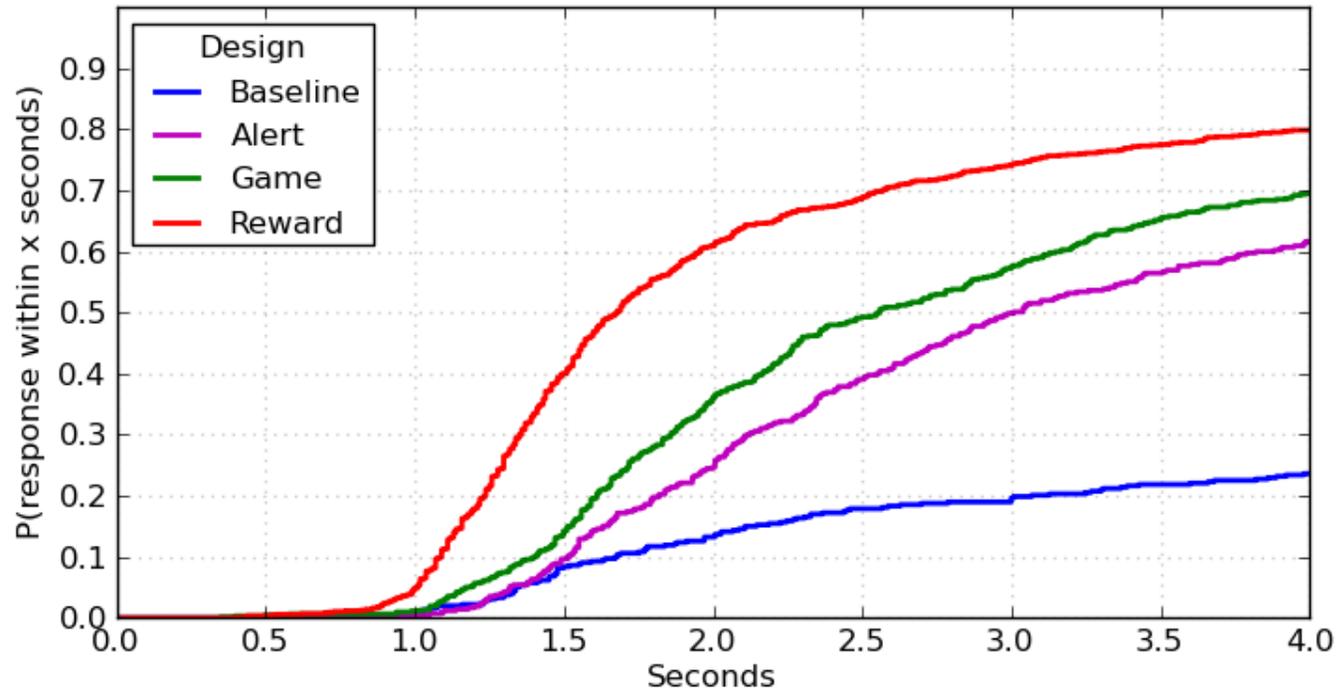
4 Reward

3¢ bonus
for dismissing
the alert
within 2 seconds



Between subjects, N=1913 tasks

Retainer Time Results



Results: Quality

Rapid Refinement had lower variance than Generate-One.

($\sigma=2.2$ vs. $\sigma=2.6$ on a 9-point Likert scale)

Generate-and-Vote matches the professional photographer.

($\mu=6.6$ vs. $\mu=6.4$)

Cost:

- Rapid Refinement and Generate-One: 22¢
- Generate-and-Vote: 53¢

Results: Delay

