

Crowdsourcing and HCI 2: Privacy and Latency

Crowdsourcing and Human Computation

Instructor: Chris Callison-Burch

Website: crowdsourcing-class.org

Privacy

Would you let crowd workers read your email?

The screenshot displays the GmailValet interface, which integrates email and task management. On the left, the **Inbox** contains 52 conversations, each with a subject, timestamp, and a star icon. The conversations listed are:

- Bud Newton: IMPORTANT: Had a really good time. Nice one, buddy! Let's see whether... (8:05 pm)
- Target Inc.: We miss you at Target! If this email can't be displayed correctly... (8:04 pm)
- Ludwig, me: How to change SQL queries on. I thought this might be interesting for you: ... (8:03 pm) - 4 starred
- Martha: Revised England Vacation. Yes! There's a great opportunity there! Let's book our ... (2:02 pm) - 1 gold star
- Claude: Sorry, Tuesday won't work! Let's reschedule by next week... (Nov 30) - 1 blue star

On the right, the **Task Stream** lists 10 tasks, each with a checkbox and a due date. The tasks are:

- Update Ludwig on TestFlight details
- Schedule phone chat with TA for CS 189
- Stay in touch on unconference and other events with Martha
- Send a few dates for Marvin, so that he can plan the bash
- Meet Bud for brunch Saturday in 4 days
- Respond to Elen about CC of choice
- Get back to Bud Newton about CS Masters program at Stanford
- Prepare meeting with Claude
- Reschedule meeting with Claude in 6 days
- Book flight tickets to England in 2 days

Problems with email as a task management tool

- Never-ending stream of incoming requests
- New messages push important requests out of view
- Some important requests can be unintentionally missed
- People spend a lot of time carefully processing their inboxes or triaging to select important messages

Email Valet

- Targeted at people who receive a large volume of email
- Tries to stop bad practice of using tricks like marking an email as unread to flag that it has something actionable, since those techniques are unreliable

Email Valet

- Recruits personal assistants for you from oDesk
- Your personal assistant reads your email and creates todo items for you
- Goal is to create an actionable task list so that things don't get lost in large steam of email
- Combine advantages of PAs with the scalability and affordability of crowds workers

Crowdsourced Personal Assistants

- oDesk is “expert” crowdsourcing platform
- Assistants are shared across multiple people
- Increases employment for assistants, reduces costs for individual users

Executive Assistants

- Microsoft Outlooks allow users to delegate limited inbox access to assistants
- Focusing their boss's attention on important messages
- Autonomously handle simple tasks
- Crowdsourcing bring assistants to new class of people – not just executives

Initial interviews

- People are of two minds about recruiting remote assistants for managing personal information
- People want the help
- But they have concerns about giving strangers unfettered access

How people use email now

- 77% send email reminders to themselves
- 47% use their inbox as a to-do list
- 41% would be willing to use an online service helps with email task management

Privacy concerns

- 38% were unwilling to share anything
- 35% were only willing to share a few messages manually
- 26% were fine with automatic rules
- 4% were ready to share their entire inbox

Email Valet



GmailValet

Forum Preferences

Inbox 52 Conversations

2 Bud Newton
IMPORTANT: Had a really good time
Nice one, buddy! Let's see whether...

Target Inc.
We miss you at Target!
If this email can't be displayed correctly...

Ludwig, me
How to change SQL queries on
I thought this might be interesting for you: ...

✓ Martha
Revised England Vacation
Yes! There's a great opportunity there! Let's book our

2 Claude
Sorry, Tuesday won't work!
Let's reschedule by next week...



Task Stream 10 Tasks

- Update Ludwig on TestFlight details Your assist
- Schedule phone chat with TA for CS 189 Your assist
- Stay in touch on unconference and other events with Martha
- Send a few dates for Marvin, so that he can plan the bash
- Meet Bud for brunch Saturday in 4 days
- Respond to Elen about CC of choice
- Get back to Bud Newton about CS Masters program at Stanford
- Prepare meeting with Claude
- Reschedule meeting with Claude in 6 days
- Book flight tickets to England in 2 days

Privacy protections

- You can create a whitelist of messages that your assistant can see (starred, labeled “assistant”, messages you send to yourself)
- You can create a blacklist to block your assistant from seeing messages from certain people, or with certain keywords
- You can limit the assistant to only viewing your most-recent messages (default: 100)

Restricting access

Emails my assistant should help me with:

My assistant can **see** (👤) the 100 most recent conversations in my inbox.

Restrict to conversations which are any of the following:

- marked Important
- starred
- labeled 'Assistant'
- emails I sent to myself

My assistant will be **blocked** (🔒) from seeing emails containing the terms:

private.gmailvalet.com ?

monica@gmail.com x

bank x

password x

One term at a time





Inbox 52 Conversations



2 Bud Newton

IMPORTANT: Had a really good one, buddy! Let's see whether...

8:05 pm



Visible to assistant

Target!nc.

We miss you at Target!

If this email can't be displayed correctly...

8:04 pm



Visible only to you

Ludwig, me

How to change SQL queries on

I thought this might be interesting for you: ...

8:03 pm

4



MVisible to my assistant because

Revised **it is starred**

Yes! There's a great opportunity there! Let's book our

2:02 pm



Reason why this is visible to assistant

2 Claude

Sorry, Tuesday won't work!

Let's reschedule by next week...

Nov 30



Stanford Federal Credit Union

Your password was reset as requested

Nov 28



CONFIDENTIAL INFORMATION BELOW

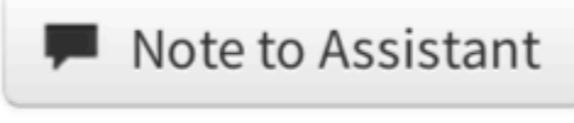
Handing over control

- You control what actions your assistant is allowed to do:
 - Create new task
 - Delete emails
 - Reply to emails

User's view of tasks

Forum Preferences alan.turing@gmail.com ▾

Task Stream 10 Tasks

Note to Assistant 

 Calendar  Task stream

<input type="checkbox"/> Update Ludwig on TestFlight details	<i>Your assistant created this task</i>	 Accept  Decline
<input type="checkbox"/> Schedule phone chat with TA for CS 189	<i>Your assistant created this task</i>	 Accept  Decline
<input type="checkbox"/> Stay in touch on unconference and other events with Martha		
<input type="checkbox"/> Send a few dates for Marvin, so that he can plan the bash		
<input type="checkbox"/> Meet Bud for brunch Saturday	 in 4 days	
<input type="checkbox"/> Respond to Elen about CC of choice		
<input type="checkbox"/> Get back to Bud Newton about CS Masters program at Stanford		
<input type="checkbox"/> Prepare meeting with Claude		

Assistant's view of tasks

- Book flight tickets to England**  **in 2 days** by client
- Schedule call with Ada** accepted
- Check out cheap online deals**
Feedback from client: *I'm not interested in ads.* declined
- Submit conference paper** undecided

Other feedback

- Users can leave notes for new assistants
- Ask assistant to prioritize certain senders
- Or add labels to tasks (“put [Event] in front of every event”)
- Assistants and users can also open a chat window to clarify any confusion

Accountability

- EmailValet displays a log of all of the actions that your assistant took, for each of the emails that they processed
- Does not prevent abuse but leaves “fingerprints” that reveal it
- May act as a deterrent

Accountability

IMPORTANT: Had a really good time at Stanford

From Bud Newton to Alan Turing

From

Bud Newton <bud.newton@gmail.com>

To

Alan Turing <alan.turing@gmail.com>

Received

Aug 23, 2012 at 4:19 AM

Assistant Log

- Anne-Marie read message
- Anne-Marie added a task

scallion rutabaga okra tigernut winter purslane tenun konirabi watercress soko punya nuts beet

Endive maize rutabaga fennel kakadu plum potato celery jícama leek salad turnip greens bur

Accountability

Action	Task	Message	Assistant	Time
Read message	—	[#VWJ-824-43394]: Explanation from Stanford U.	Ludwig	6 min ago
Created task	Set up unconference meeting with Martha	MyGSB Digest for the week of from MyGSB Digest	Ludwig	6 min ago
Read message	—	MyGSB Digest for the week of from MyGSB Digest	Ludwig	16 min ago

Study

- Do you think that having an assistant would increase your productivity?
- How would you measure that?

Weeklong Study

Control group	Couldn't see assistant-created tasks and couldn't create their own tasks
DIY group	Couldn't see assistant-created tasks, but could create their own tasks
Assisted group	Saw assistant-created tasks and create their own tasks. Could give feedback to their assistant.

Participants rotated through each of the 3 conditions for 2 days at a time, after 1 day warm-up.

Study participants, Assistants

- 28 university students (6 MBAs, 22 tech)
- Participants were paid \$50 gift certificate
- 3 online assistants hired through oDesk
- Paid \$8 per hour to process all shared emails during the study

What was measured

- How many tasks that the assistant created were accepted by user
 - In control and DIY groups, the user marked the hidden tasks at the end of the 2-day period to create ground truth
- How many tasks were completed during the 2 day period
 - Manually merged the DIY tasks and the assistant tasks at the end

Precision

- 72% of assistant-created tasks were accepted by users
- Precision increased over time from 62% on first day to 85% on the last day

“it has become easier to extract good and accurate tasks from my clients’ emails over time. I feel I have gotten to know my clients better and understand the conversations better” –assistant

Recall

- How many of the tasks were created by the assistant? *How many of the user-created tasks did the assistant miss?*
- Only measured on assisted-condition when users could add tasks in real time
- 69% recall. However, sometimes the user logged in before the assistant, so potential recall may be higher.

Free-form survey

- Were the assistants' tasks relevant, or just busywork?
- 67%: valuable tasks worth completing
- Some said assistants were overeager, e.g. creating todos from mailing lists
- Still felt that it was easier to delete than create tasks

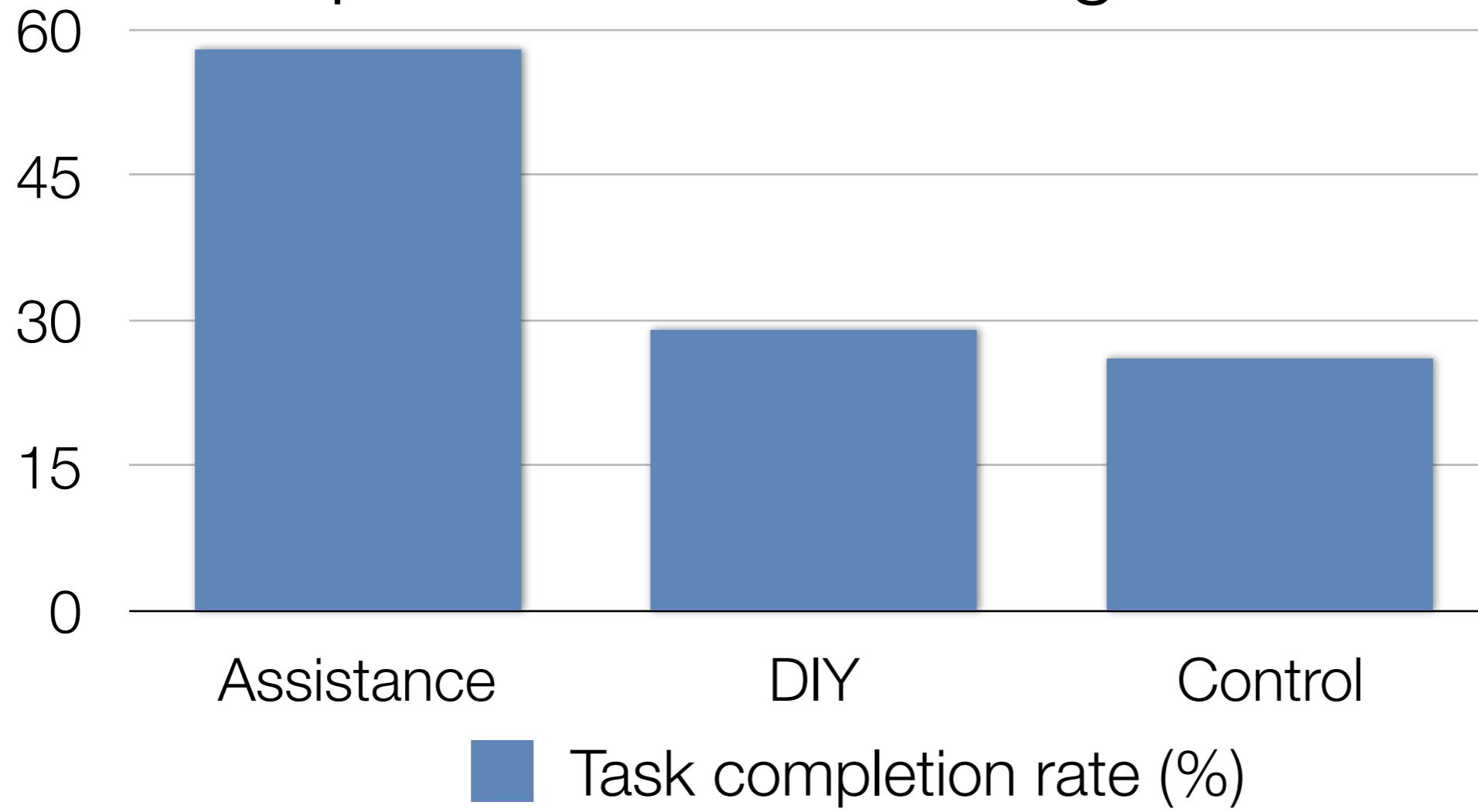
Free-form survey

- Were users confident that their assistants would not miss important tasks?
- 61% felt they could fully or almost fully rely on their assistant
- Most common cause of missing tasks was lack of contextual knowledge

“Many important tasks (that are not obvious) are not extracted.”

Did EmailValet increase productivity?

- Users found the assistants to be generally accurate, but did the system help those users manage their tasks?



Enthusiasm

- “any help in making sure everything gets done is greatly appreciated.”
- “What I need is an extra pair of eyes.”
- Assistant’s tasks were “like magic”: “very convenient and much easier than doing it myself.”

Contributions of EmailValet

- Crowdsourced expert assistants to support personal information management
- An email task management system with integrated feedback structure
- Empirical results indicate that assistants manage information accurately, enabling users to accomplish more

Limited Access in a Transparent Fashion

- Give assistants only as much access as they actually need
- Interface access boundaries transparent so users have an accurate model of what the assistant can and cannot do
- Audit log creates fingerprints of any possible transgressions

Economics of shared assistants

- Assistants worked for 70 hours total
- Processed 12k messages (~3/minute)
- Created 780 tasks (~7 per 100 emails)

Economics of shared assistants

- Each assistant could do ~1,400 messages per day if working full time
- Each user got about 40 messages per day
- Could support 36 users simultaneously
- Cost to users would be \$1.78 per day

Possible extensions

- Support other delegated tasks
- Summarize messages
- Negotiate meeting times
- Draft/send replies

Would you let crowd workers read your email?

The screenshot displays the GmailValet interface, which integrates email and task management. On the left, the **Inbox** contains 52 conversations, each with a subject, timestamp, and a star icon. The conversations listed are:

- Bud Newton: IMPORTANT: Had a really good time. Nice one, buddy! Let's see whether... (8:05 pm)
- Target Inc.: We miss you at Target! If this email can't be displayed correctly... (8:04 pm)
- Ludwig, me: How to change SQL queries on. I thought this might be interesting for you: ... (8:03 pm) - 4 starred
- Martha: Revised England Vacation. Yes! There's a great opportunity there! Let's book our ... (2:02 pm) - 1 gold star
- Claude: Sorry, Tuesday won't work! Let's reschedule by next week... (Nov 30) - 1 blue star

On the right, the **Task Stream** lists 10 tasks, each with a checkbox and a due date. The tasks are:

- Update Ludwig on TestFlight details
- Schedule phone chat with TA for CS 189
- Stay in touch on unconference and other events with Martha
- Send a few dates for Marvin, so that he can plan the bash
- Meet Bud for brunch Saturday in 4 days
- Respond to Elen about CC of choice
- Get back to Bud Newton about CS Masters program at Stanford
- Prepare meeting with Claude
- Reschedule meeting with Claude in 6 days
- Book flight tickets to England in 2 days

Latency

Crowds in the interface

- Tasks like email are reasonably asynchronous, so some delay is acceptable
- For other tasks, like Word Processing, we would like a rapid response
- Soylent and TurKit both suffered from a problem of latency

Latency in HCI is disastrous

- Users are not used to waiting, and will abandon interfaces that are slow to react
- Search engine usage decreases linearly as delays grow
- Ten seconds is the maximum delay before a user loses focus on an interaction

How can we solve the
problem of latency?

VizWiz: Nearly Real-time Answers to Visual Questions

What denomination is this bill?



(24s) 20
(29s) 20

Do you see picnic tables across the parking lot?



(13s) no
(46s) no

What temperature is my oven set to?



(69s) it looks like 425 degrees but the image is difficult to see.
(84s) 400
(122s) 450

Can you please tell me what this can is?



(183s) chickpeas.
(514s) beans
(552s) Goya Beans

What kind of drink does this can hold?



(91s) Energy
(99s) no can in the picture
(247s) energy drink

Pre-recruit workers

- VizWiz tried to reduce latency by pre-recruiting workers
- Workers complete a series of assignments in on HIT
- The user's request with the least responses gets put at the head of the queue

Know when work is imminent

61 seconds	Start app, take picture
71 seconds	Record the question
78 seconds	Press send
221 seconds	Wait for response

Start
recruiting
workers

Maintain a work pool

- TurKit also experimented with maintaining a group of workers, even when there was no work
- Created dummy assignments from past assignments, to ensure work
- When a new request arrived a dummy was replaced with the real request
- Can be costly to constantly maintain a pool

Retainer model

- Alternate to maintaining worker pool with dummy tasks
- Hire crowd workers in advance, and pay them a small amount to wait for work to come online
- All them to pursue other work while waiting
- Alert them when our task is ready with a popup box, and pay them for that work too

Goals of Retainer Model

1. Guarantee a fast response time
2. Be cheap enough to scale
3. Maintain response time after a long wait

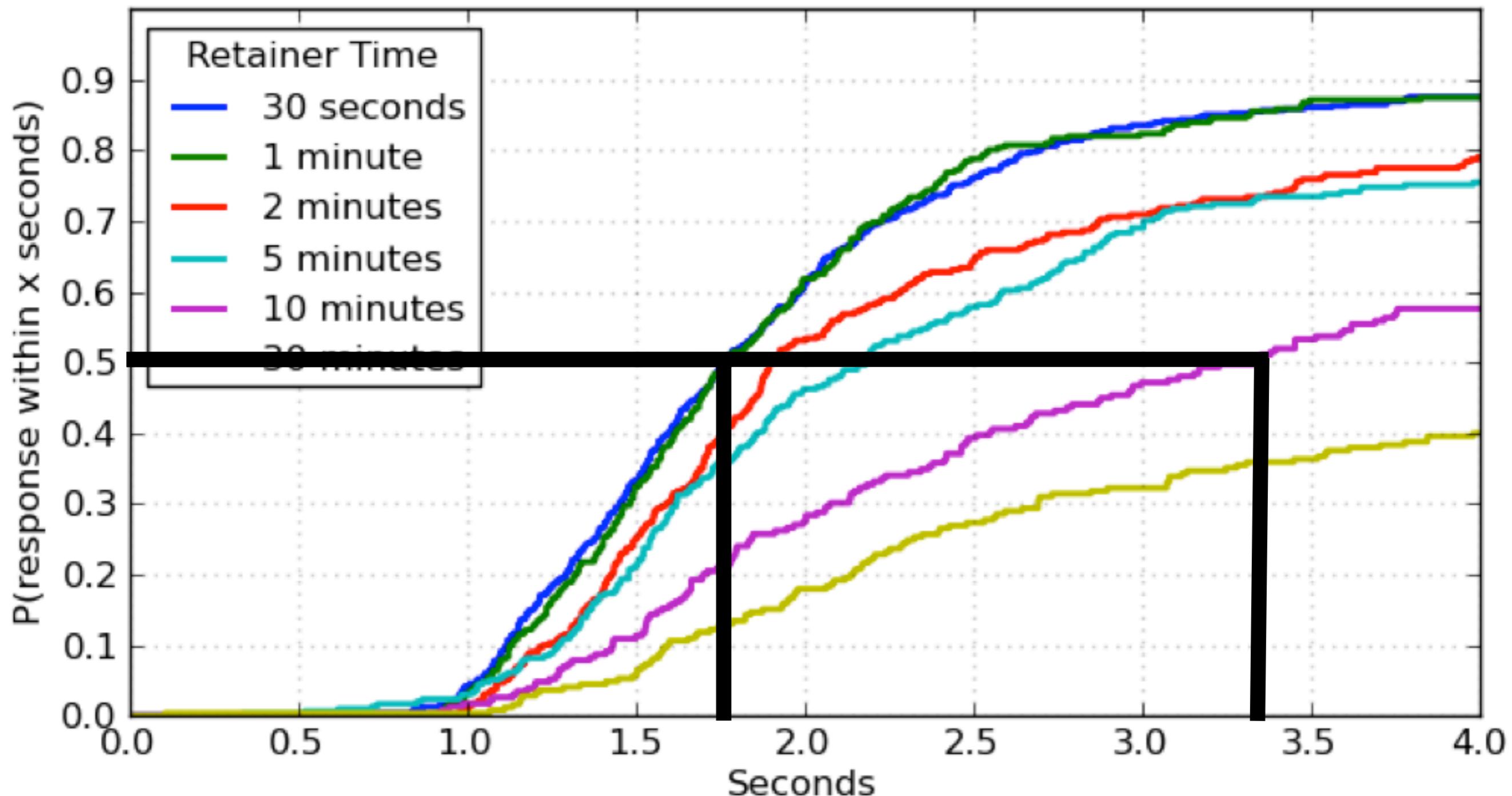
Getting paid to wait

- Turkers were \$0.005 – \$0.01 per minute, scaled based on expected wait time
- Asked them to keep the tab open and told them that they were free to do other tasks while waiting
- Javascript alert when work was ready
- Optionally, offer a small bonus to reward quick responses
- If no work is ready at end of retention period, given them an old task to complete

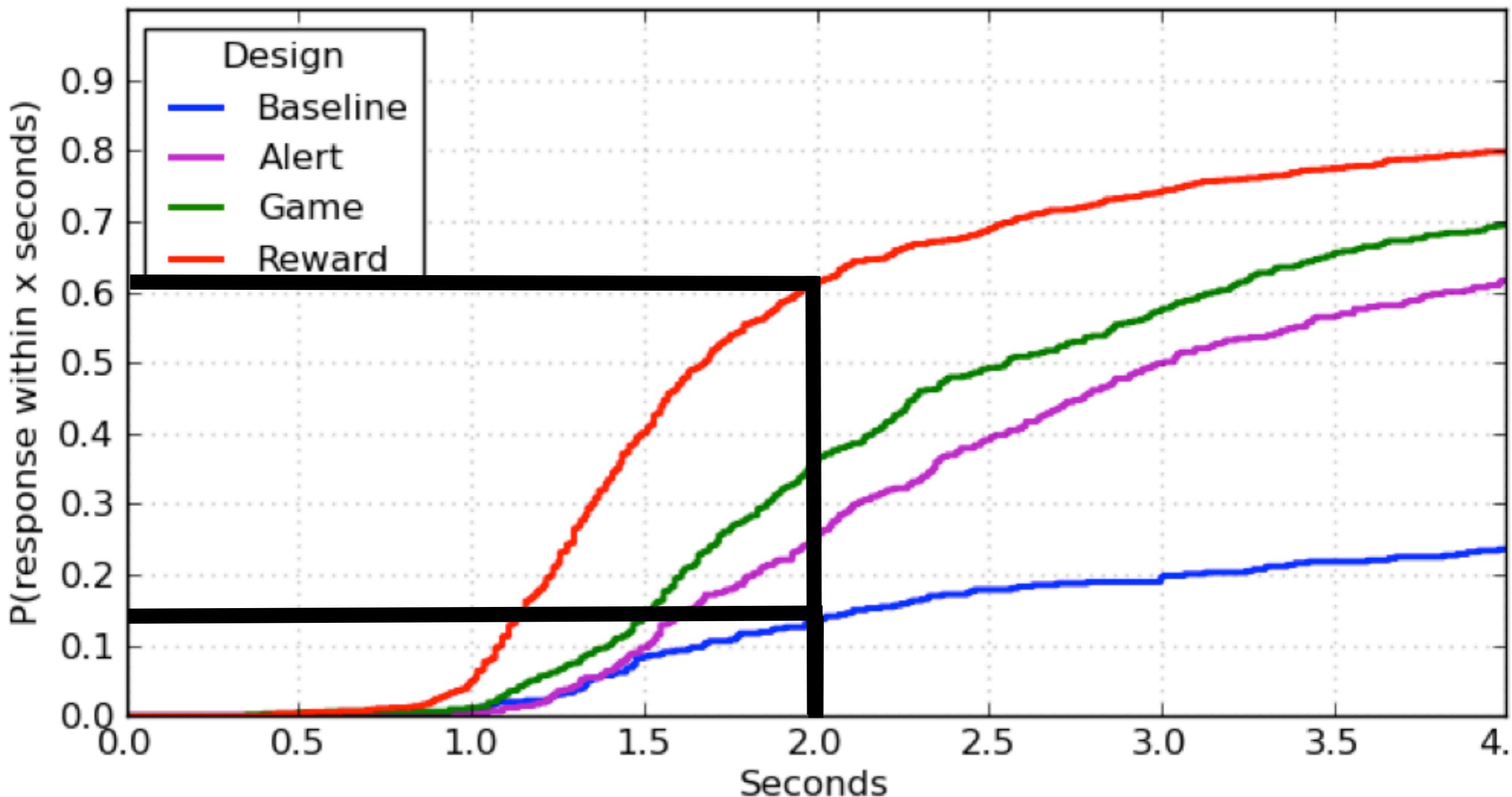
Super-quantifiable HCI experiment

- Vary retainer time between 0.5, 1, 2, 5, 10, and 30 minutes
- Pay workers: 2¢, 3¢, 4¢, 7¢, 12¢, 32¢
- Measure time from Javascript alert appearing until they dismiss it

Response time



Improving 10 minute retainer response time



Dramatic speedups

- No longer wait for minutes or hours
- Nearly zeros out wait time
- Approaches human limits on the cognitive recognize-act cycle and motor reaction times

Cost of retainer

- Cost of the retainer model is attractive because it pays workers a small amount to wait, rather than spending money to repeat old tasks
- Cost depends on the desired arrival time, and the empirical arrival distribution, and the desired number of workers

Instant-on crowd

- What becomes possible if we can have access to workers in ≤ 2 seconds?

Synchronous Crowds

- With the retainer model, we have guarantees about the arrival time for workers
- This applies not just to individual workers, but for groups of workers
- We can do tasks that require multiple workers interacting, or that composite results from multiple workers to get the task done even faster

Novel Applications with Synchronous Crowds

- Adrenaline - a camera application that selects a photo from an action video
- Puppeteer - a way of manipulating lots of movable digital puppets to create a scene
- A|B - a quick voting system for A/B testing (which font is the best?)

Video

Reminder: Part 1 of your final
project (videos + questions)
due on Wednesday

We'll do in-class
presentations on Wednesday
and Monday