Much Ado About Time: Exhaustive Annotation of Temporal Data

Gunnar A. Sigurdsson (CMU), Olga Russakovsky (CMU), Ali Farhadi (UW, AI2), Ivan Laptev (INRIA), Abhinav Gupta (CMU)

http://allenai.org/plato/charades/   *Published in HCOMP'16*

Goal: explore cost-effective strategies for annotating temporal data (e.g., videos)

☑ Cat?   ☑ Cat walking?
☑ Sandwich?   ☑ Eating a sandwich?
☒ Phone?   ☑ Playing on phone?

150 ms   **30 sec**
+ 2 sec per question   + 2 sec per question

| Image | Video/Audio/Text |

How many questions about a video can we ask simultaneously?

Few questions: questions can be kept in worker's short-term memory, increasing annotation recall

Many questions: video watching time can be amortized, reducing annotation cost

Instructions

Below is a link to a video of one or two people, please watch each video and answer the questions.

- This HIT contains multiple videos, each followed by few questions. *The number of videos and questions is balanced such that the task should take 3 minutes.*
- Make sure you *fully and carefully watch each* video so you *do not miss anything. This is important.*
- It is possible that many of the actions in this HIT do not match. It is important to verify an action is indeed *not* present in the video.
- **Check all that apply!** If there is any doubt, check it anyway for good measure.
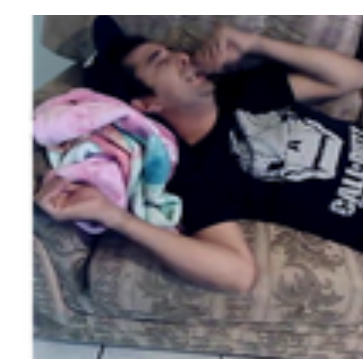- **Read each and every question carefully. Do not take shortcuts, it will cause you to miss something.**

☑ Check here if someone is *Taking a picture of something* in the video
☑ Check here if someone is *interacting with cup/glass/bottle* in the video
If checked, how? **(Select that apply. Use ctrl or cmd to select multiple):**

Drinking from a cup/glass/bottle
Holding a cup/glass/bottle of something
Pouring something into a cup/glass/bottle
Putting a cup/glass/bottle somewhere
Taking a cup/glass/bottle from somewhere
Washing a cup/glass/bottle
Other

☐ Check here if someone is *interacting with laptop* in the video
☐ Check here if someone is *interacting with doorknob* in the video
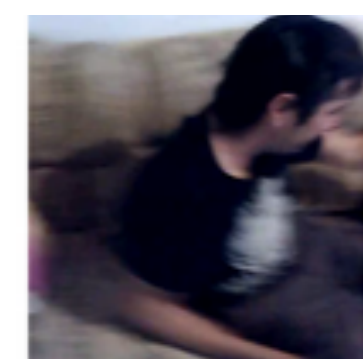☐ Check here if someone is *interacting with table* in the video

---

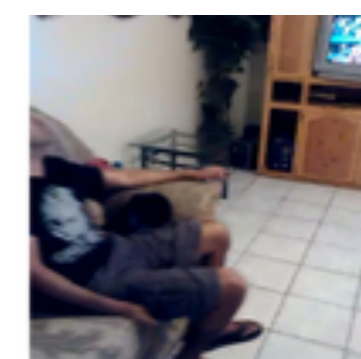**Dataset:** Charades video dataset of [Sigurdsson ECCV16]

Watching/Looking outside a window
Sitting at a table
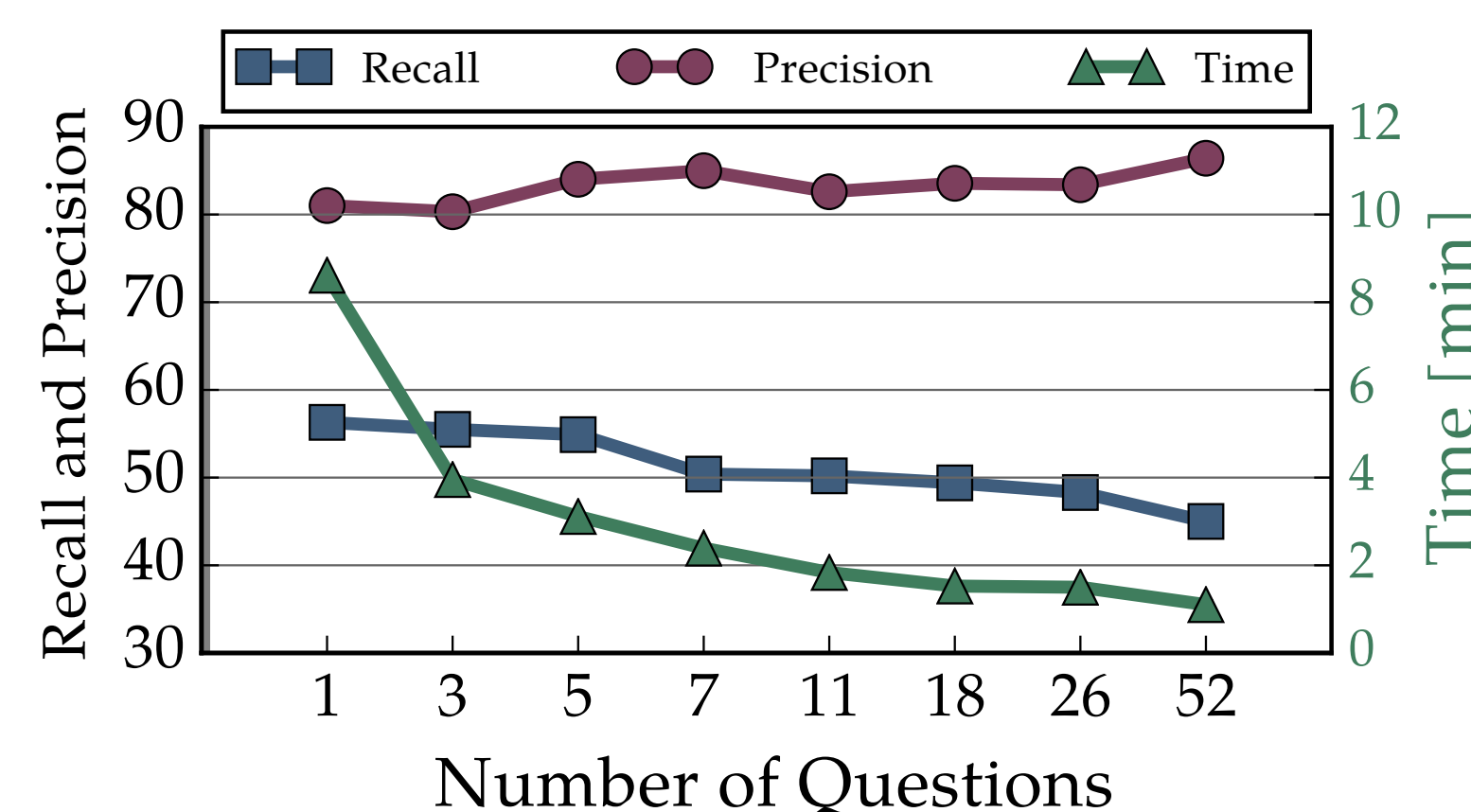Lying on a sofa/couch
Sitting on a sofa/couch
Watching television

- Household activities
- Videos recorded by AMT workers
- 9,848 videos, averaging 30 secs
- 157 activity classes (52 hierarchically)
- Some activities pre-annotated but not exhaustively

Experiment #1: Run one round of annotation

Annotate 157 actions in 140 videos on Amazon Mechanical Turk

- **Recall** is *lower* with more questions, as expected
- **Precision** turns out to be *higher* with more questions
- **Time/cost** is *lower* with more questions, as expected

Analyze worker behavior during annotation

- Worker enjoyment doesn't depend on number of questions
- Workers **interact with the video** more when asked more questions
- Workers **annotate fewer actions** when asked more questions

---
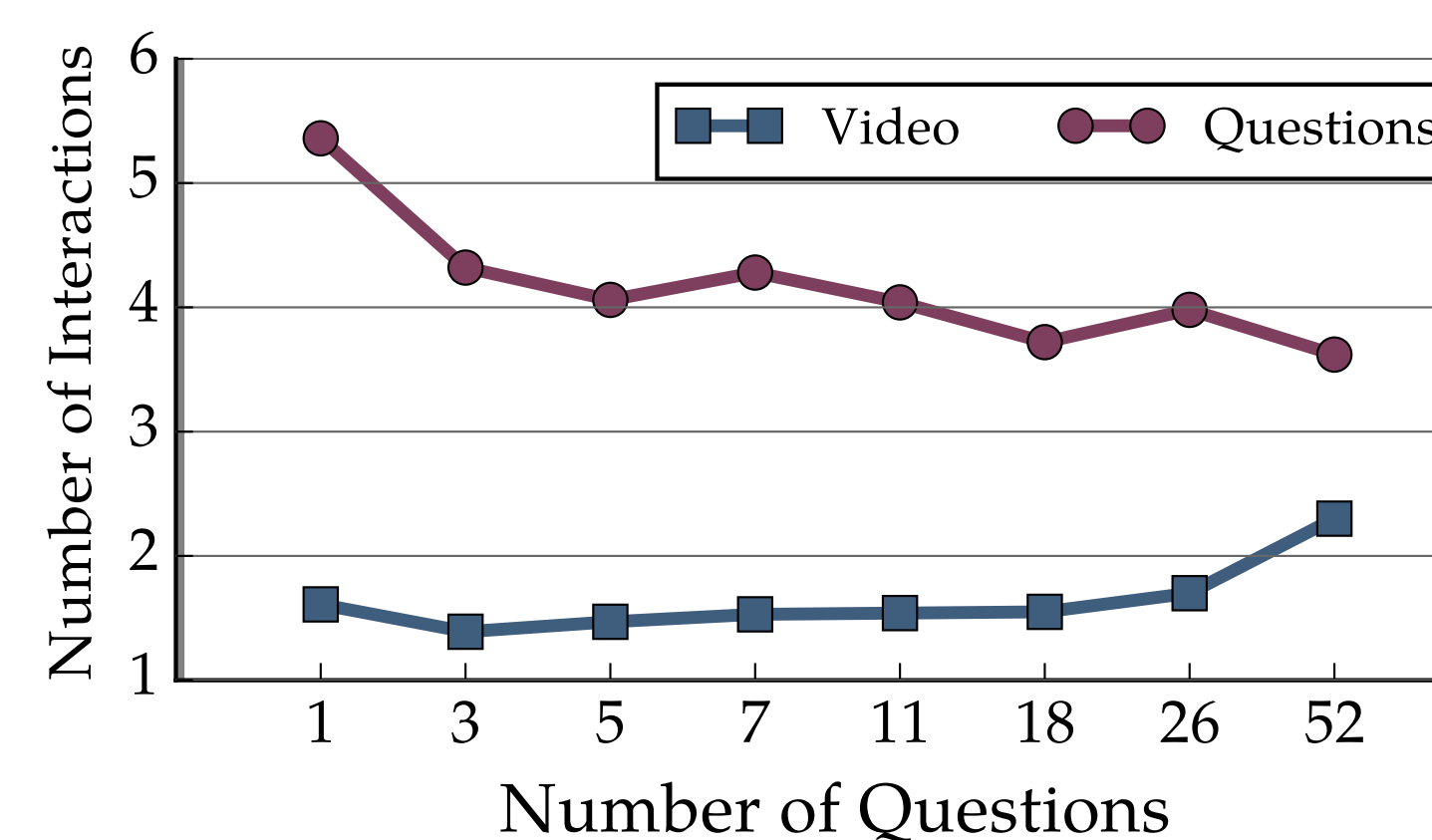
Experiment #2: Improve and re-evaluate the UI

- **Improving time/cost:** Keep questions consistent within a HIT
  - 13.6% reduction in annotation time
- **Improving precision:** Use the available annotations to inject more positive questions into the HIT
  - 2.3% improvement in precision
- **Improving recall try #1:** First request a 20-word video summary
  - No effect on recall but 40% slower annotation
- **Improving recall try #2:** Force yes/no response to every question
  - Recall actually drops (annoys workers?)
- **Improving recall** for real: Run multiple rounds of annotation
  - Recall improves from 58.0% to 83.3% after 3 rounds

Experiment #3: Bringing it all together

52 Questions
26 Questions
5 Questions
1 Question
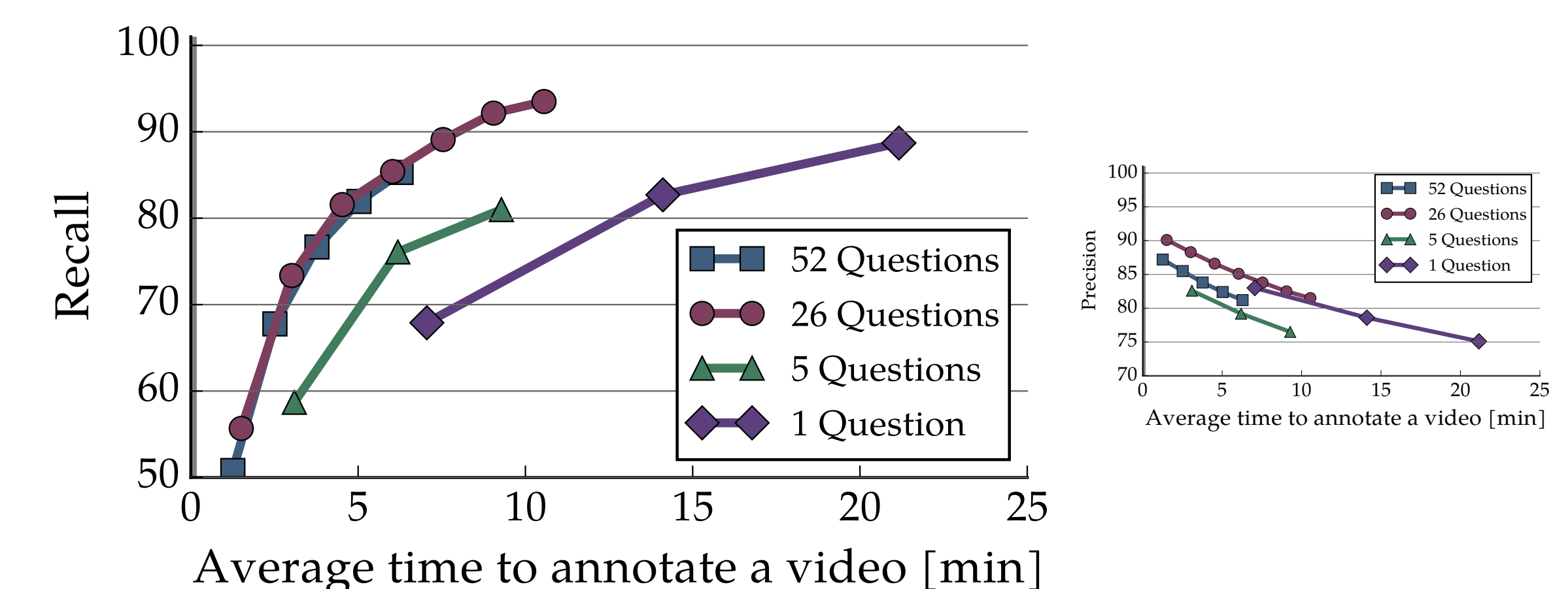
Conclusion: Asking *many questions* per video is a much more effective use of worker time than asking just a *few questions*

Annotated data: 9,848 videos exhaustively annotated

- 66,963 action instances from 1,310,014 collected annotations
- Increased density of annotation from 3.7 to 9.0 labels per video
- Additionally collected temporal annotations
- All publicly available at *http://allenai.org/plato/charades/*