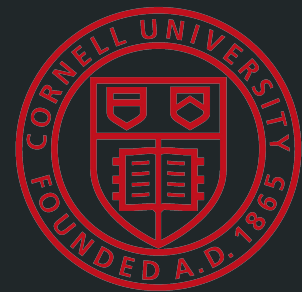


Consensus Agreement Games

Paul Upchurch, Daniel Sedra, Andrew Mullen, Haym Hirsh and Kavita Bala
Cornell University
HCOMP 2016





A group of people are seated around a dining table in a restaurant. The table is set with plates of food, glasses, water bottles, and a large bowl of chips. A man in an orange and white shirt is looking towards the right. A woman with long blonde hair is looking towards the left. A man in a blue and white shirt is in the foreground, looking towards the right. A woman in a green shirt is on the right side of the table. A man in a blue shirt is in the foreground on the left. A yellow label 'Lamp' points to a lamp in the background. Two yellow labels 'Plate' point to plates on the table. The background shows a restaurant interior with other tables and a window looking out at night.

Lamp

Plate

Plate

A group of people are seated around a dining table in a restaurant. The table is set with plates of food, glasses, water bottles, and a centerpiece of flowers. A man in an orange and white shirt with 'CASA SOL' on it is visible. A man in a blue and white shirt is in the foreground. A yellow callout box labeled 'Glass' points to a glass on the table. Another yellow callout box labeled 'Food' points to a bowl of food. A third yellow callout box labeled 'Fabric' points to a placemat on the table. The background shows a restaurant interior with other tables and a window looking out at night.

Glass

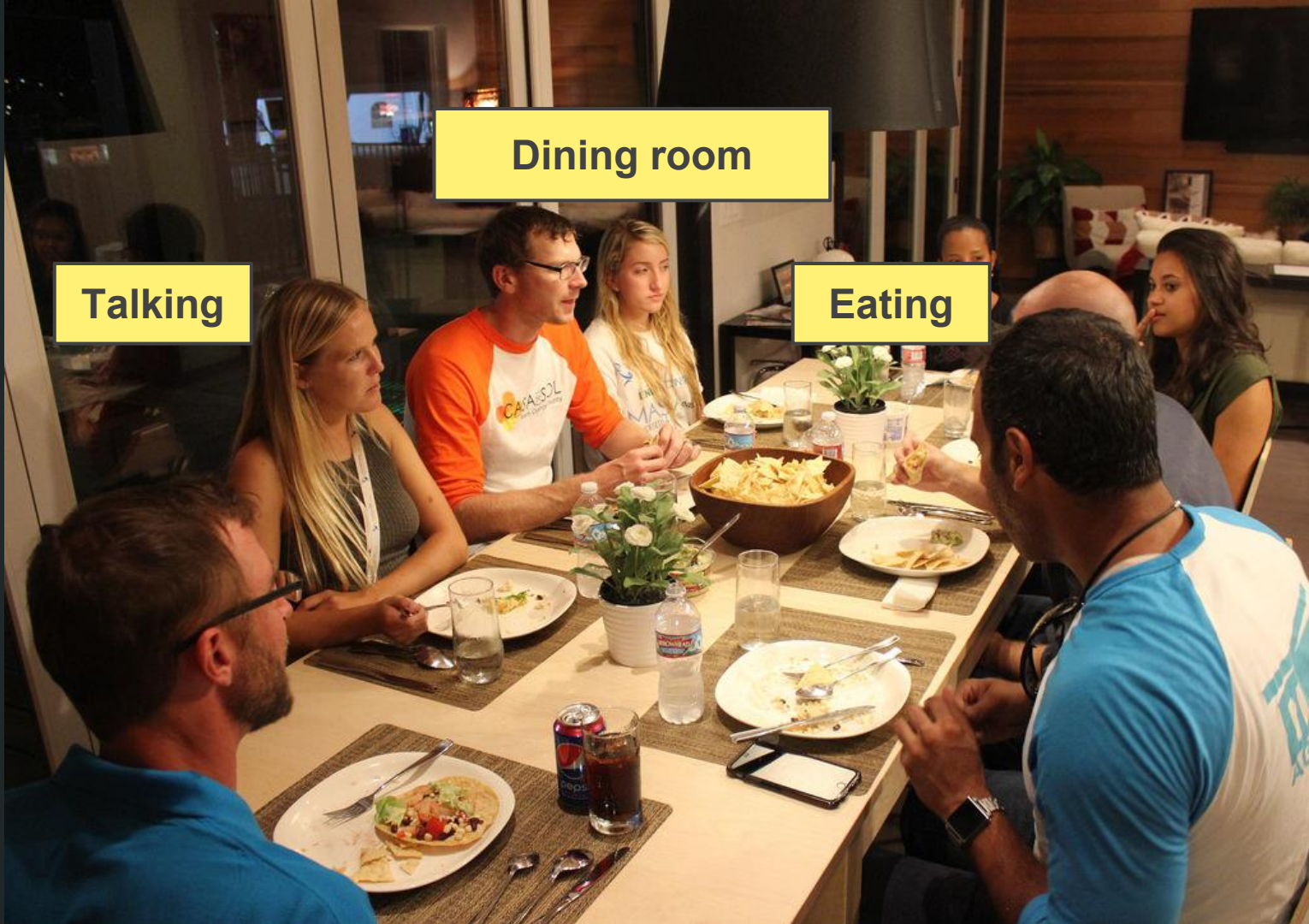
Food

Fabric

Dining room

Talking

Eating





OpenSurfaces, hard-to-label examples



Ceramic or
glass?

OpenSurfaces, hard-to-label examples



Ceramic or
glass?



Painted or
wood?

OpenSurfaces, hard-to-label examples



Ceramic or
glass?



Painted or
wood?



Ceramic, plastic
or foliage?

OpenSurfaces, hard-to-label examples



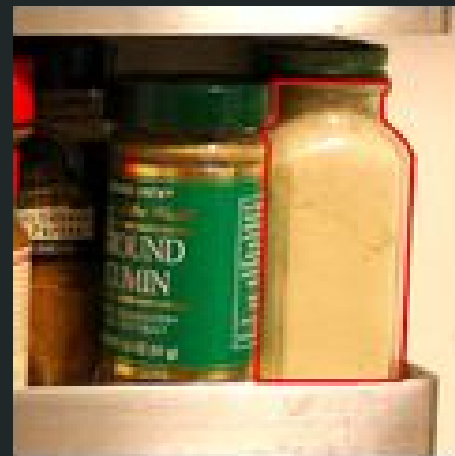
Ceramic or
glass?



Painted or
wood?



Ceramic, plastic
or foliage?



Plastic or
glass?

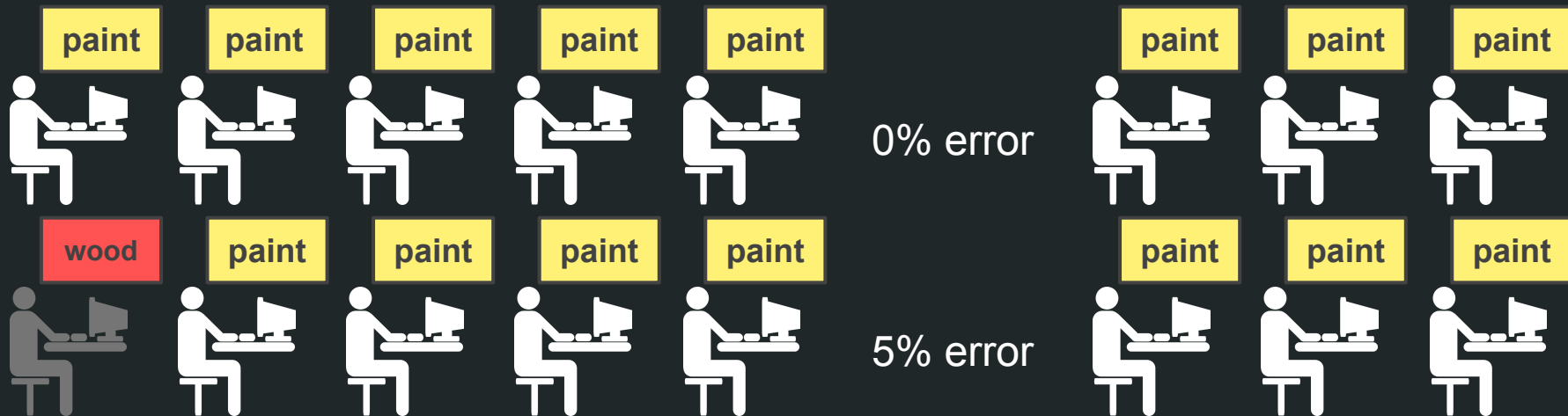
Worker Error Rate by Agreement



MTurk workers

Experts

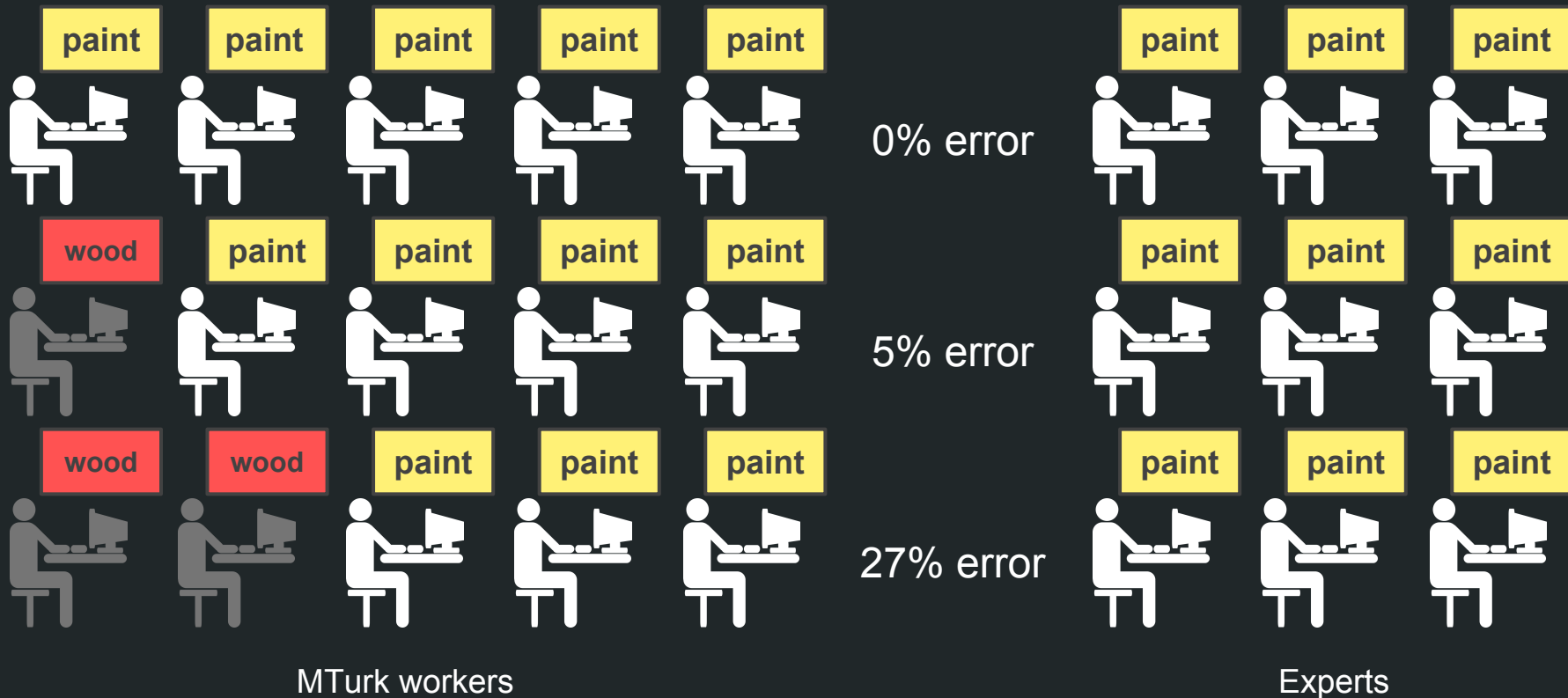
Worker Error Rate by Agreement



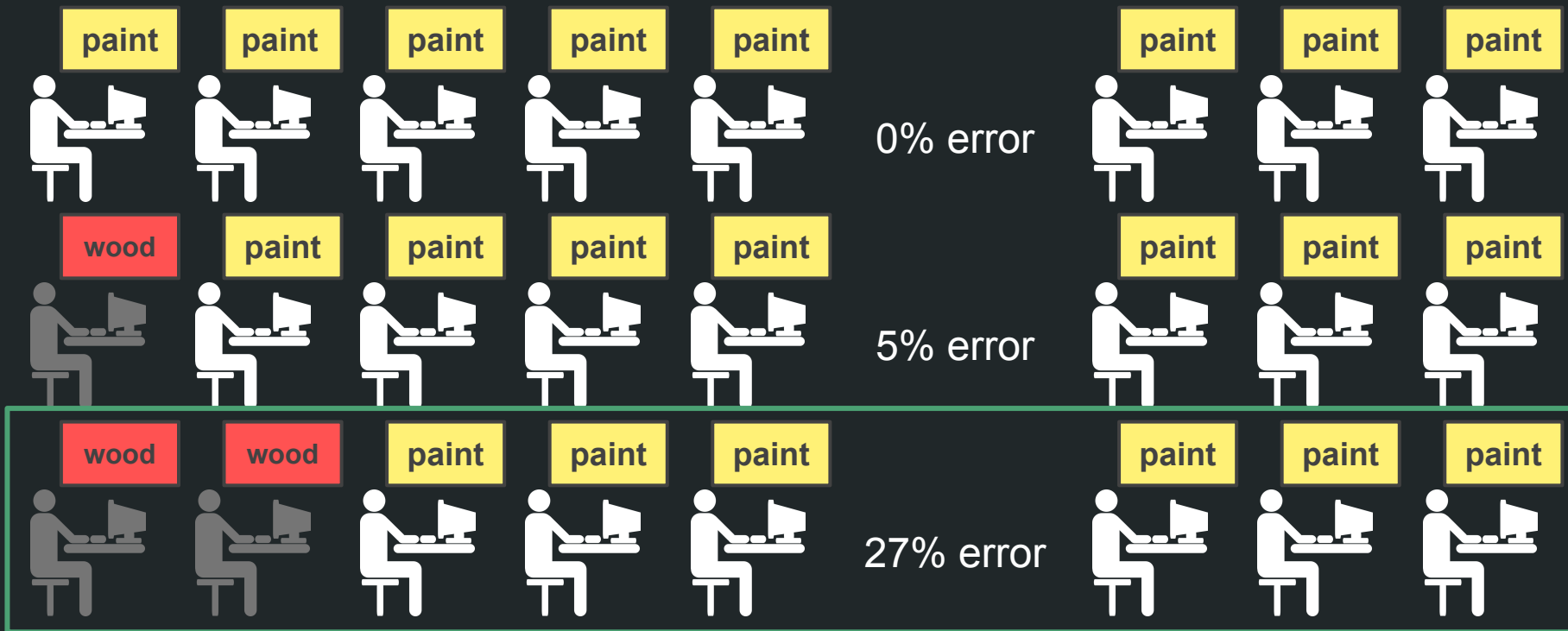
MTurk workers

Experts

Worker Error Rate by Agreement



Worker Error Rate by Agreement



MTurk workers

Experts

Get More Worker Opinions?



7 votes
19.4% error

Get More Worker Opinions?



7 votes
19.4% error



11 votes
19.0% error

Get More Worker Opinions?



7 votes
19.4% error

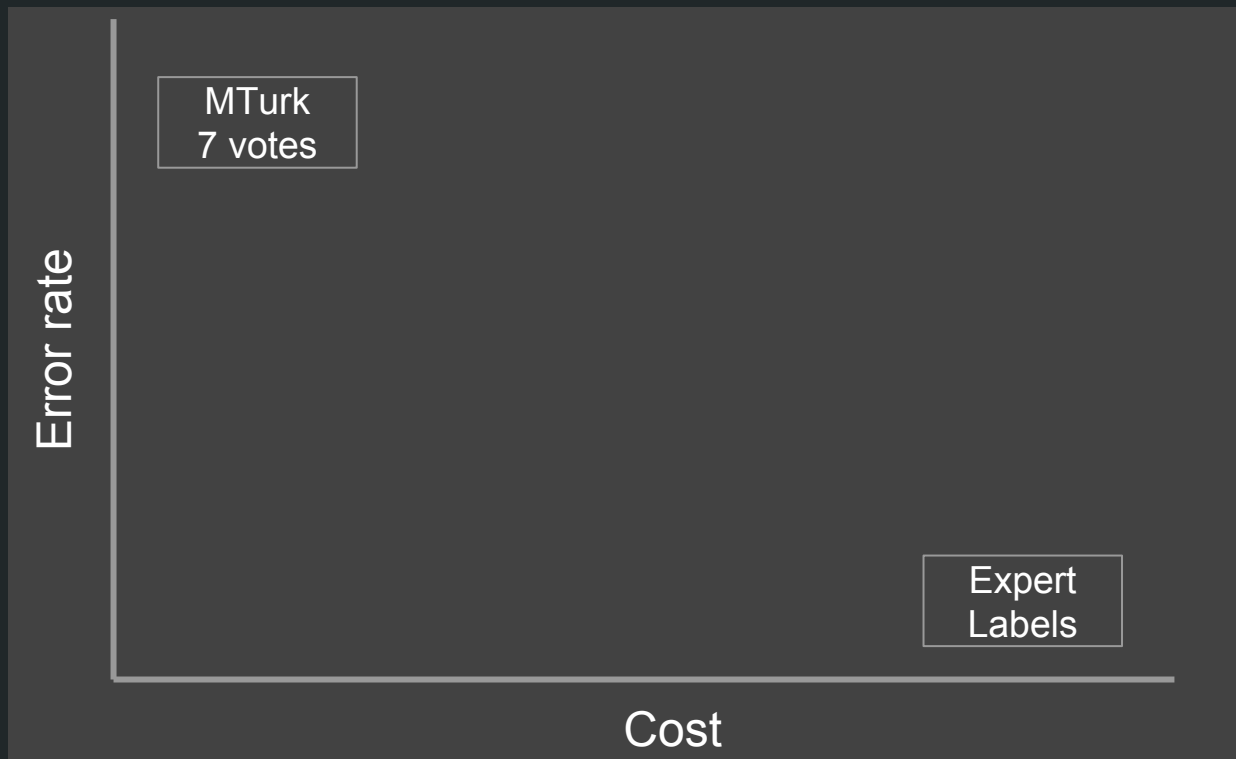


11 votes
19.0% error

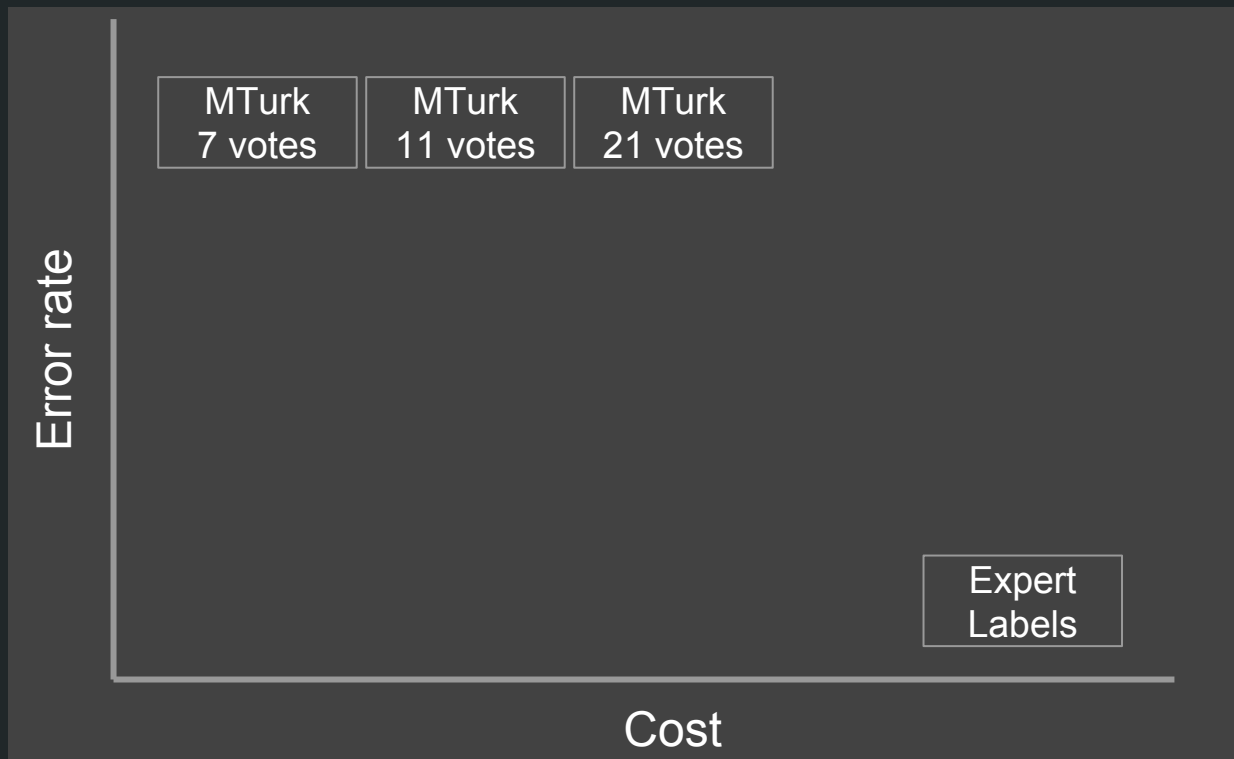


21 votes
19.6% error

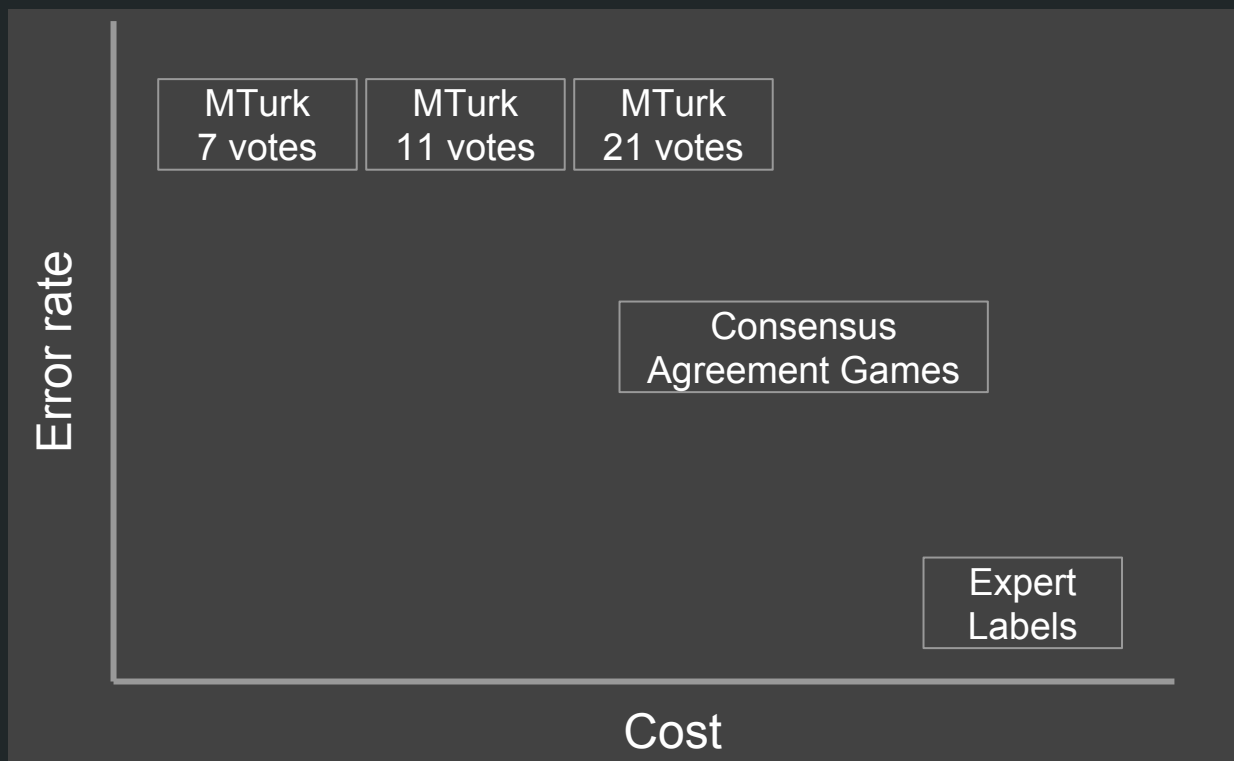
Cost vs. Quality on 3 out of 5 images



Cost vs. Quality



Cost vs. Quality



Outline

The problem

Related Work

Our method

Experimental results

Related Work

Image labeling games.

The ESP Game. Luis von Ahn and Laura Dabbish. (2004)



Related Work

Image labeling games.

The ESP Game. Luis von Ahn and Laura Dabbish. (2004)

Consensus on graphs.

Network graph coloring. Stephen Judd, Michael Kearns and Yevgeniy Vorobeychik. (2010)



Related Work

Image labeling games.

The ESP Game. Luis von Ahn and Laura Dabbish. (2004)

Consensus on graphs.

Network graph coloring. Stephen Judd, Michael Kearns and Yevgeniy Vorobeychik. (2010)

Peer agreement methods.

Peer Truth Serum. Boi Faltings, Pearl Pu, Bao Duy Tran and Radu Jurca. (2014)

Outline

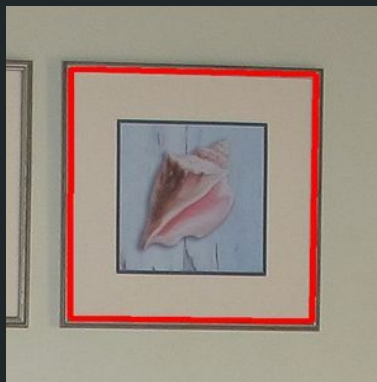
The problem

Related work

Our method

Experimental results

Idea



Idea



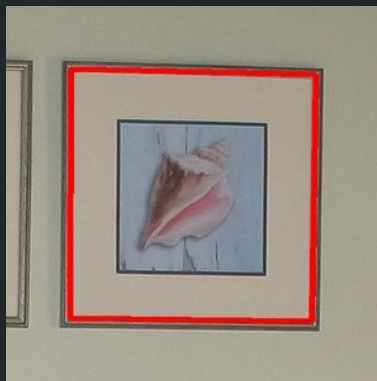
paper



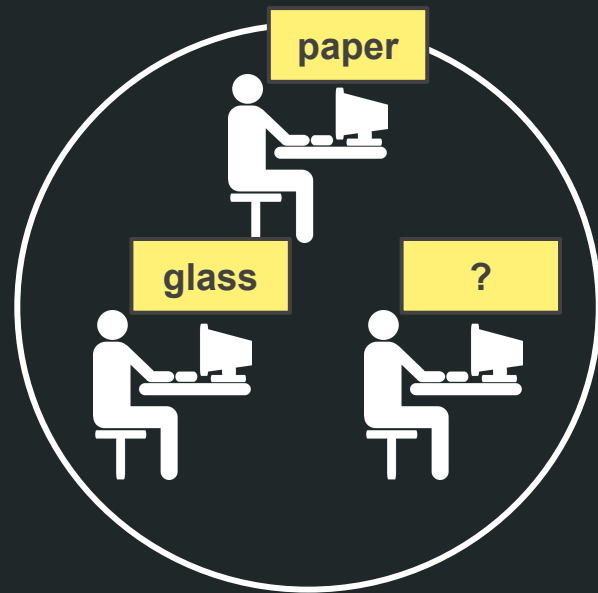
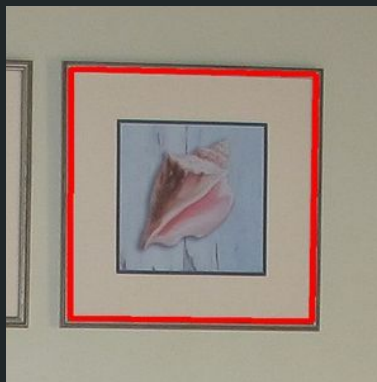
glass



?



Idea



Consensus Agreement Games

Clique 1

Workers see all clique 1 votes



Clique 2

Workers see all clique 2 votes



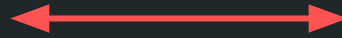
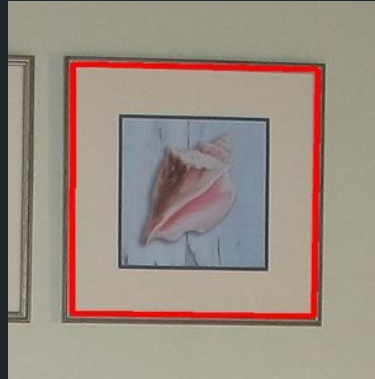
Consensus Agreement Games

Clique 1

Workers see all clique 1 votes



Shared image




No communication
between cliques

Clique 2

Workers see all clique 2 votes



Worker Experience



Player A :

Your Vote :

Player C :

Time remaining:
120 seconds

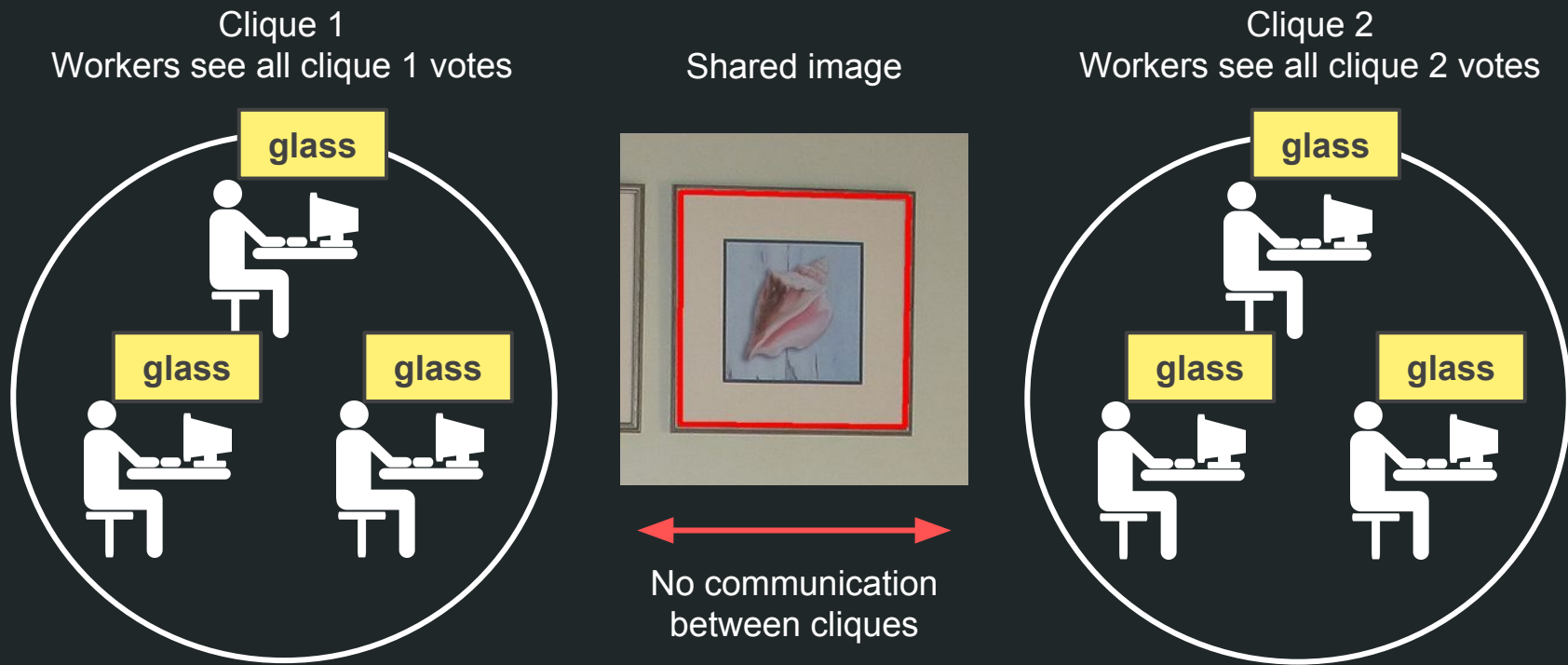
Payments:

1/8 cent

+ 1/2 cent consensus

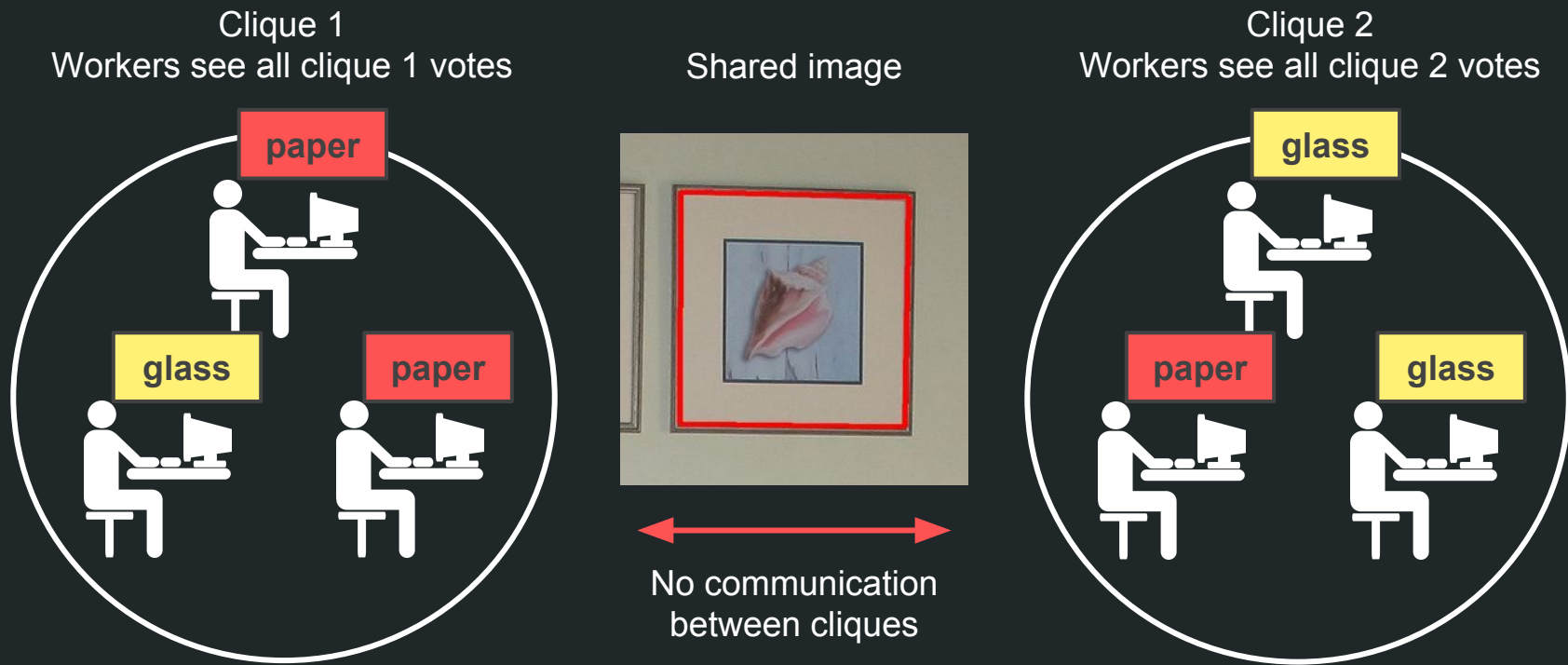
+ 1 1/2 cents agreement

Consensus Agreement Games



Case 1: consensus agreement

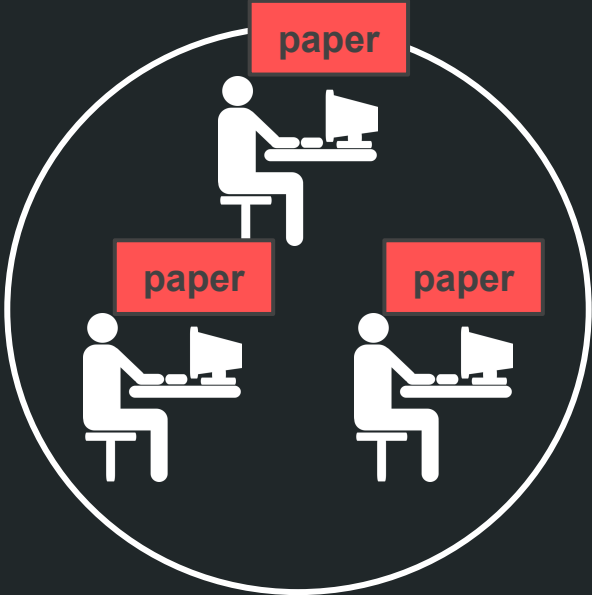
Consensus Agreement Games



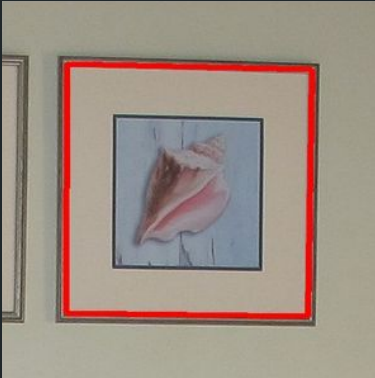
Case 2: no consensus

Consensus Agreement Games

Clique 1
Workers see all clique 1 votes

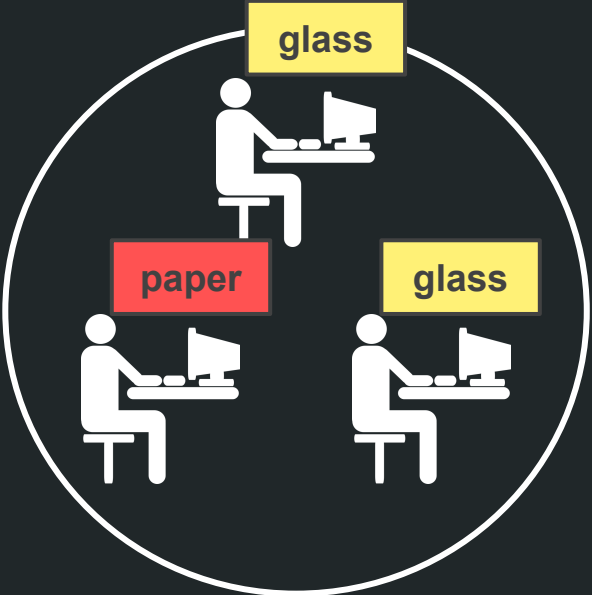


Shared image



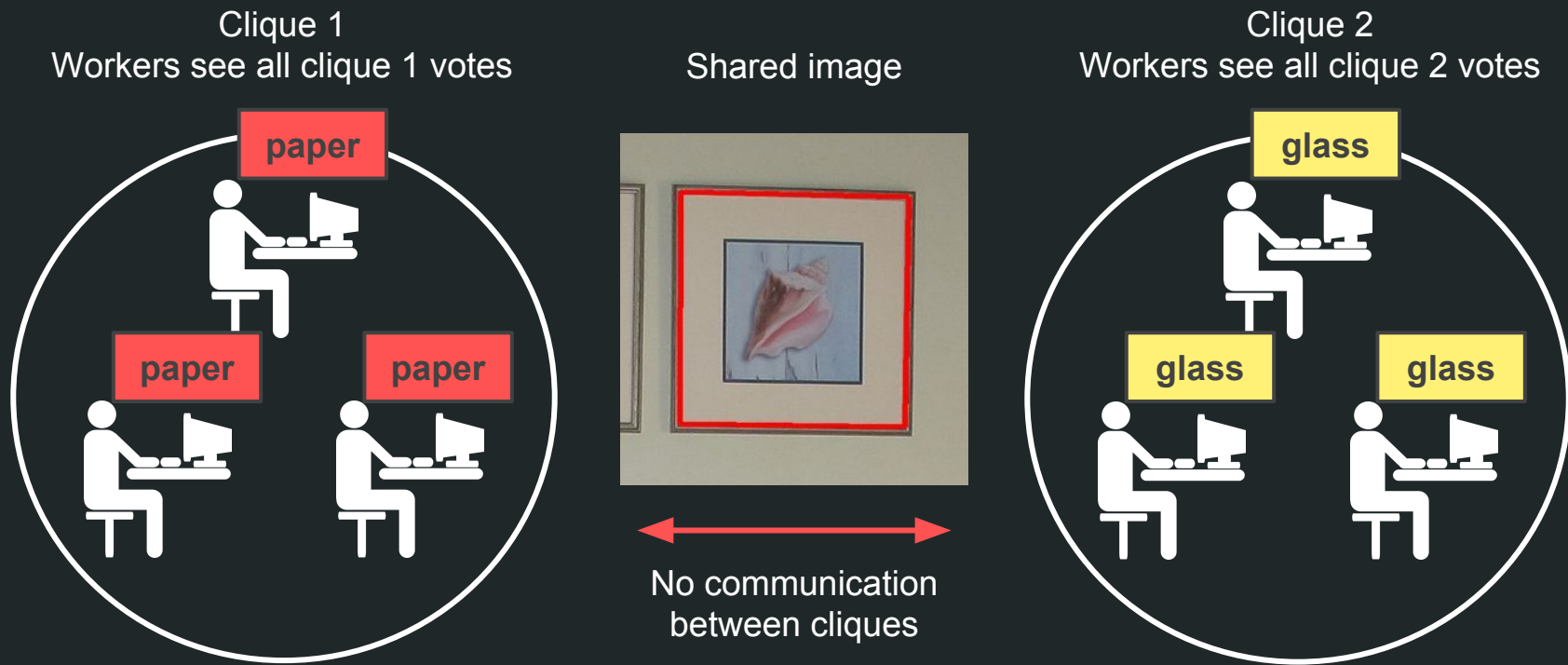
No communication
between cliques

Clique 2
Workers see all clique 2 votes



Case 3: single consensus

Consensus Agreement Games



Case 4: consensus disagreement

Outline

The problem

Related work

Our method

Experimental results

OpenSurfaces, material recognition



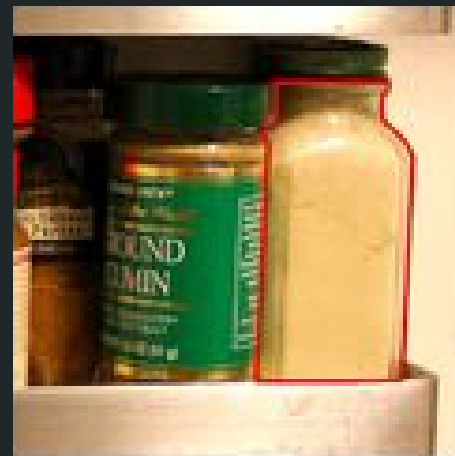
Ceramic or
glass?



Painted or
wood?



Ceramic, plastic
or foliage?



Plastic or
glass?

Places, scene recognition



Bedroom, hotel
room or nursery?



River, ocean
or coast?

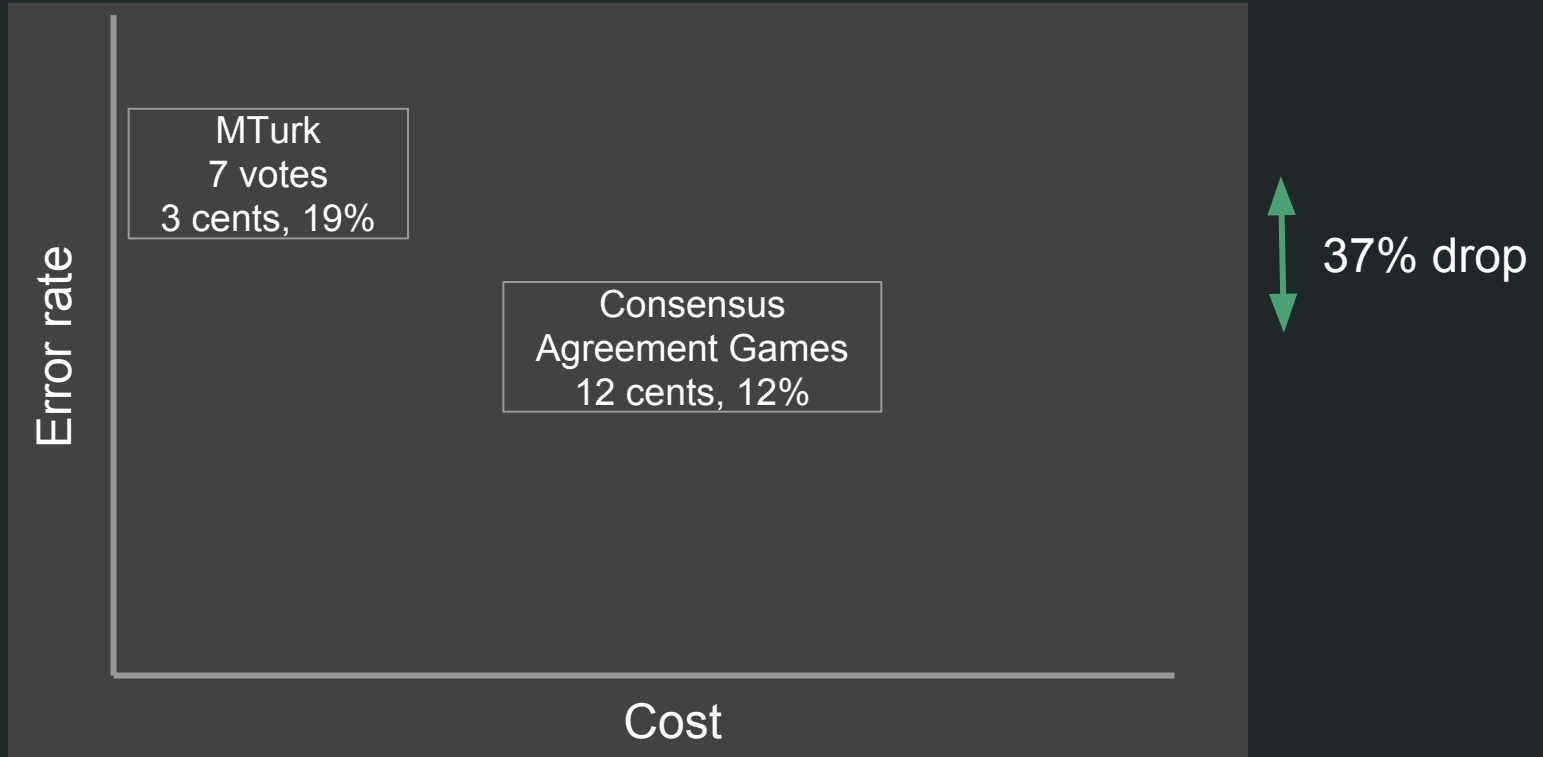
Clique size?



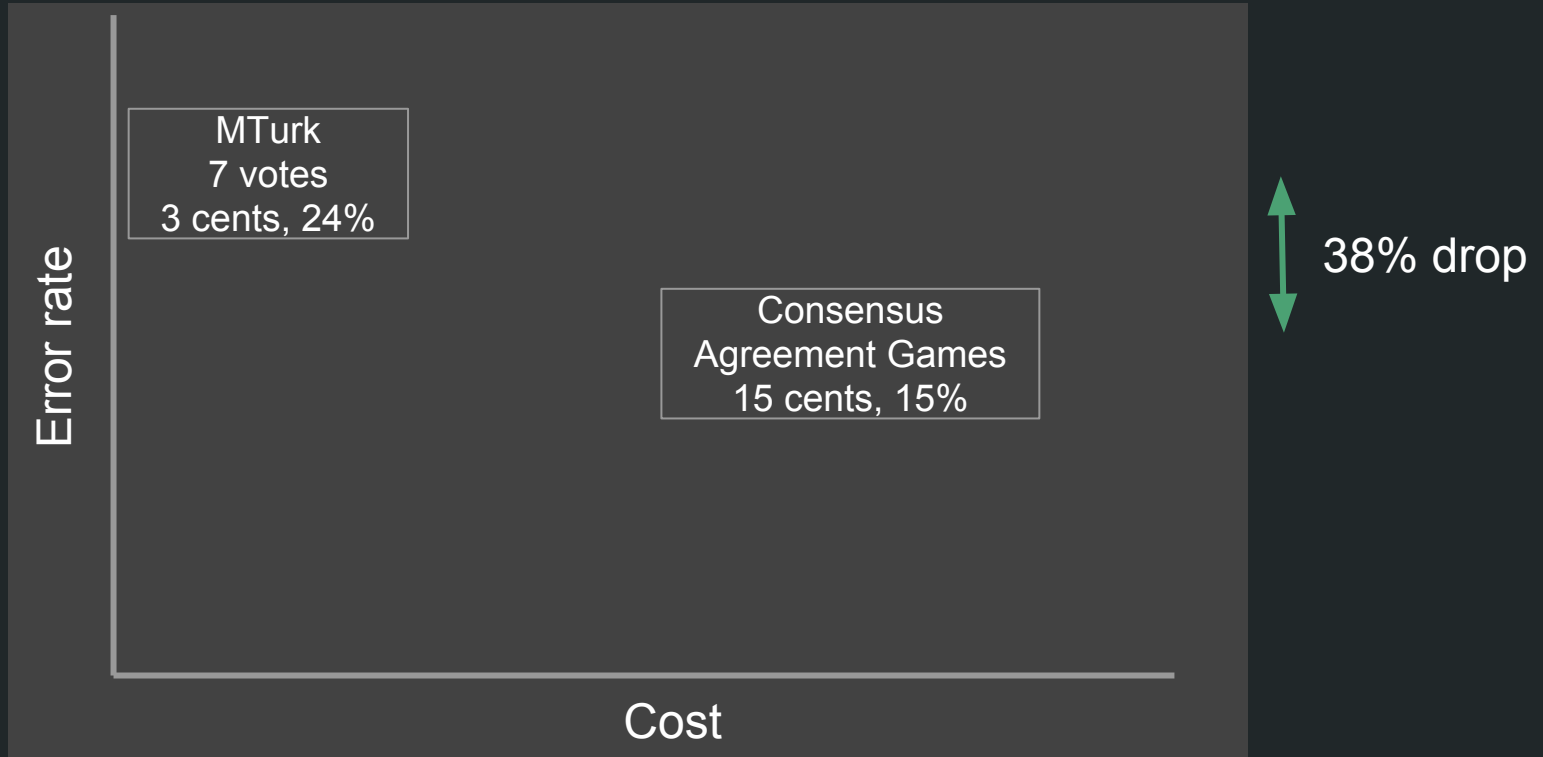
versus



OpenSurfaces, cliques of 2



Places, cliques of 3

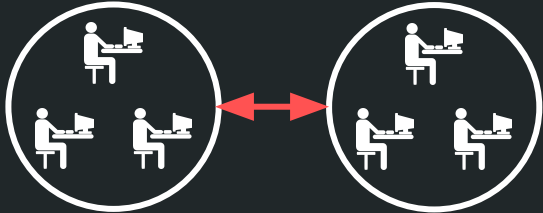


Consensus Agreement Games

Consensus



Agreement



Peer
Incentives

Payments:
1/8 cent
+ 1/2 cent consensus
+ 1 1/2 cents agreement

Interactive
Game

Player A :	glass	painted	paper	metal
Your Vote :	glass	painted	paper	metal
Player C :	glass	painted	paper	metal

Questions?

Thanks for listening!