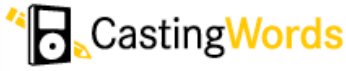
Evaluating and Improving the Usability of Mechanical Turk for Low-Income Workers in India

Shashank Khanna, IIT Bombay
Aishwarya Ratan, Microsoft Research India
James Davis, UC Santa Cruz
Bill Thies, Microsoft Research India

The Rise of Paid Crowdsourcing

- In the last decade, over 1 million workers have earned $1-2 billion via crowdsourced work*



- Opportunity for workers in developing regions?
  - Eliminates need for co-location and formal contracts
  - Flexible hours – can work in “free time”

** B. Frei. Paid Crowdsourcing: Current State & Progress towards Mainstream Business Use. Smartsheet White Paper, Sep 2009*

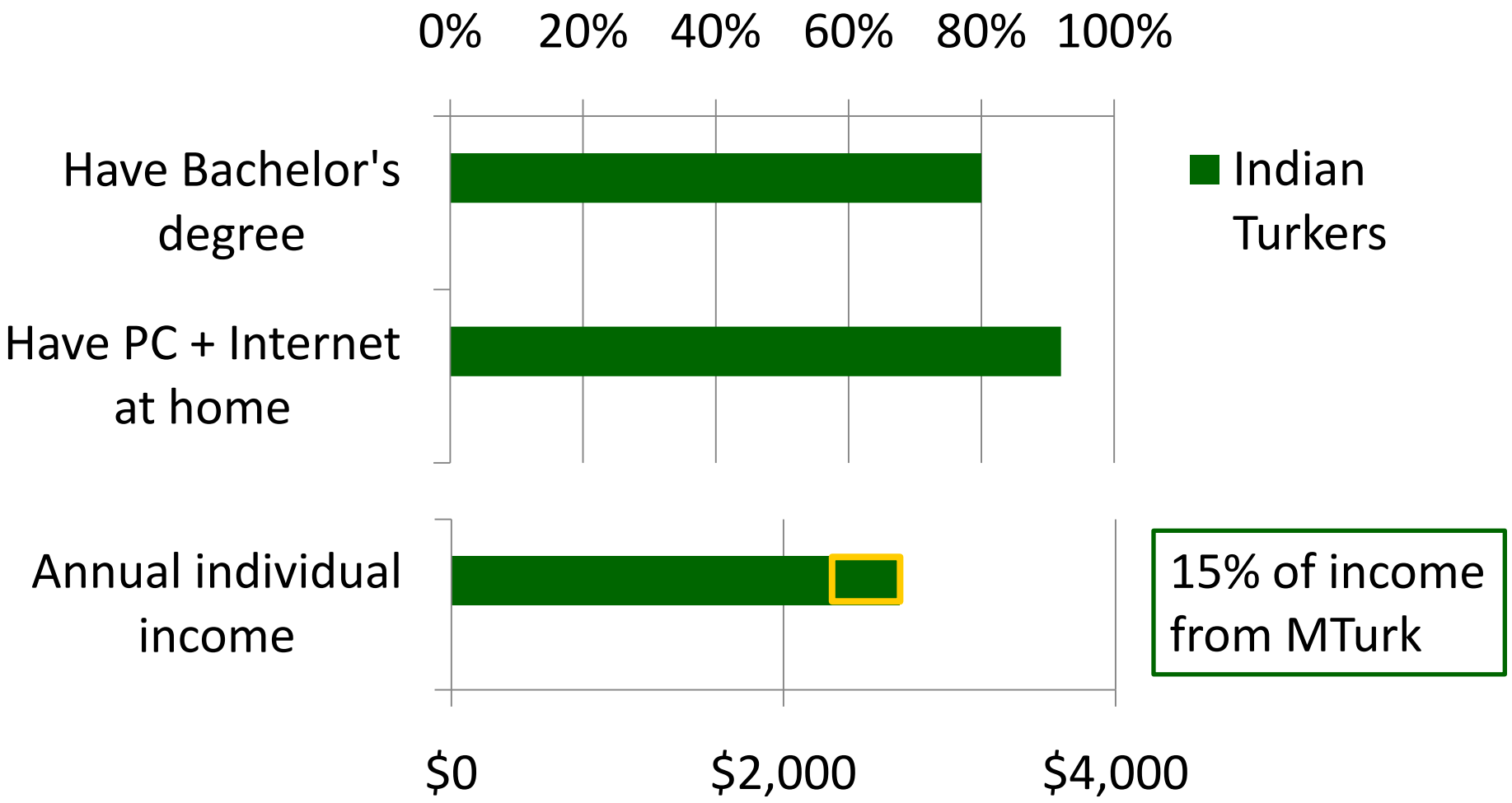
Mechanical Turk Changes Lives in India

- 36% of MTurk workers are in India [Ross'10]
- From our survey of 200 Indian Turkers (July 2010):

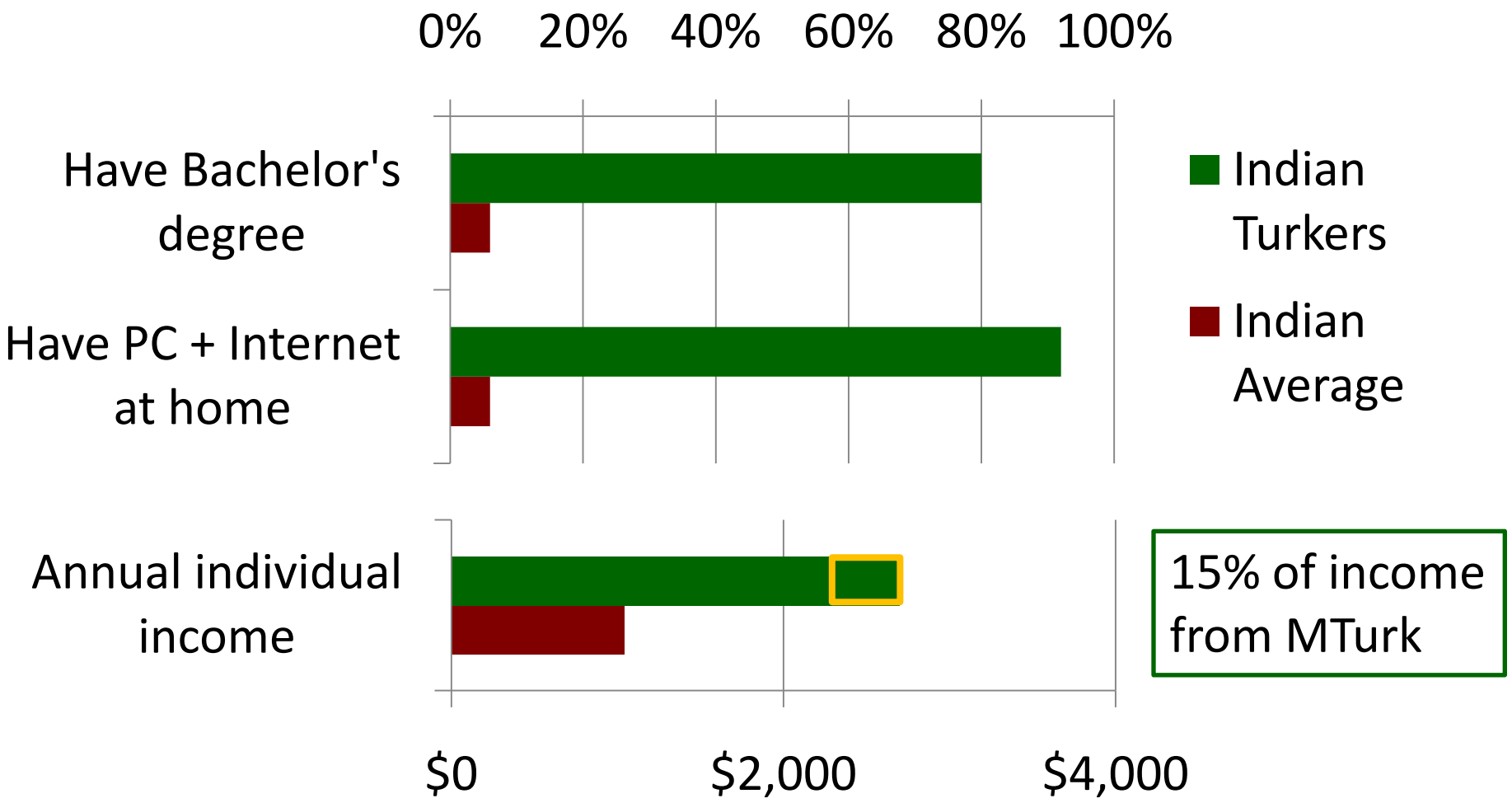
> "I'm from a middle class family. After completing my degree I looked for job everywhere but failed. But when I found MTurk, it changed my life. It helped me a lot."
>
> *— 26-year old college graduate from Kolkata. Earns $1860 / year on Turk.*

> "MTurk [is] really an advantage to me, it helps me to pay my college fees myself. It made me to feel I'm on my own. I got the respect while studying by this reasonable income."
>
> *— Respondent from Trichy. Earns $1600 / year on Turk.*

But Most Users are in High-Income Group



But Most Users are in High-Income Group



Our Study: Evaluating and Improving MTurk for Low-Income Workers in India

- Methods:
  - Observe 7 users attempting various tasks on MTurk
  - Pick a single task (bounding box), iteratively refine UI
  - Evaluate 5 variations of user interface across 49 users
- Results:
  - The UI is a bottleneck for low-income users on MTurk
  - Language localization is necessary, but not sufficient
  - Simplified interfaces and task instructions can boost completion of bounding box task from 0% to 66%

Closely Related Work

- Samasource
- txteagle
- CrowdFlower
- Prior studies of MTurk [Ross'10] [Ipeirotis'10]

In This Talk

- Usability Barriers
- Iterative Design
- Earning Potential

Focus: Lower-Income Urban Users

- Participants from two locations:
  - Office support staff: security guards, housekeeping, maintenance staff, etc.
  - Nonprofit IT training center: members with and without jobs, many students
- Median education: 12 years
- Median income: $1330 / year
  - 2nd quintile (20-40%) for urban India
- Went to local-language school, but know basic English
- Have basic digital literacy, but no exposure to MTurk

Outside the IT training center

Initial Observations

	Verify Address	Test New CAPTCHA	Label Image
Input Method	Text	Graphical	Graphical
Output Method	Text	Text	Graphical

- With each of 7 participants:
  - Participant registers on MTurk and attempts 1-2 tasks
  - Hour-long 1-on-1 session, providing help if needed

Initial Observations

	Verify Address	Test New CAPTCHA	Label Image
Input Method	Text	Graphical	Graphical
Output Method	Text	Text	Graphical
Inherent Barriers to Completing Task			

- With each of 7 participants:
  - Hour-long 1-on-1 session, providing help if needed
  - Participant registers on MTurk and attempts 1-2 tasks

Initial Observations

	Verify Address	Test New CAPTCHA	Label Image
Input Method	Text	Graphical	Graphical
Output Method	Text	Text	Graphical
Inherent Barriers to Completing Task	• Evaluating trust on Web • Nuanced use of language		

- With each of 7 participants:
  - Hour-long 1-on-1 session, providing help if needed
  - Participant registers on MTurk and attempts 1-2 tasks

Initial Observations

	Verify Address	Test New CAPTCHA	Label Image
Input Method	Text	Graphical	Graphical
Output Method	Text	Text	Graphical
Inherent Barriers to Completing Task	• Evaluating trust on Web • Nuanced use of language	• Ignoring truly illegible letters • Converting to unformatted text	

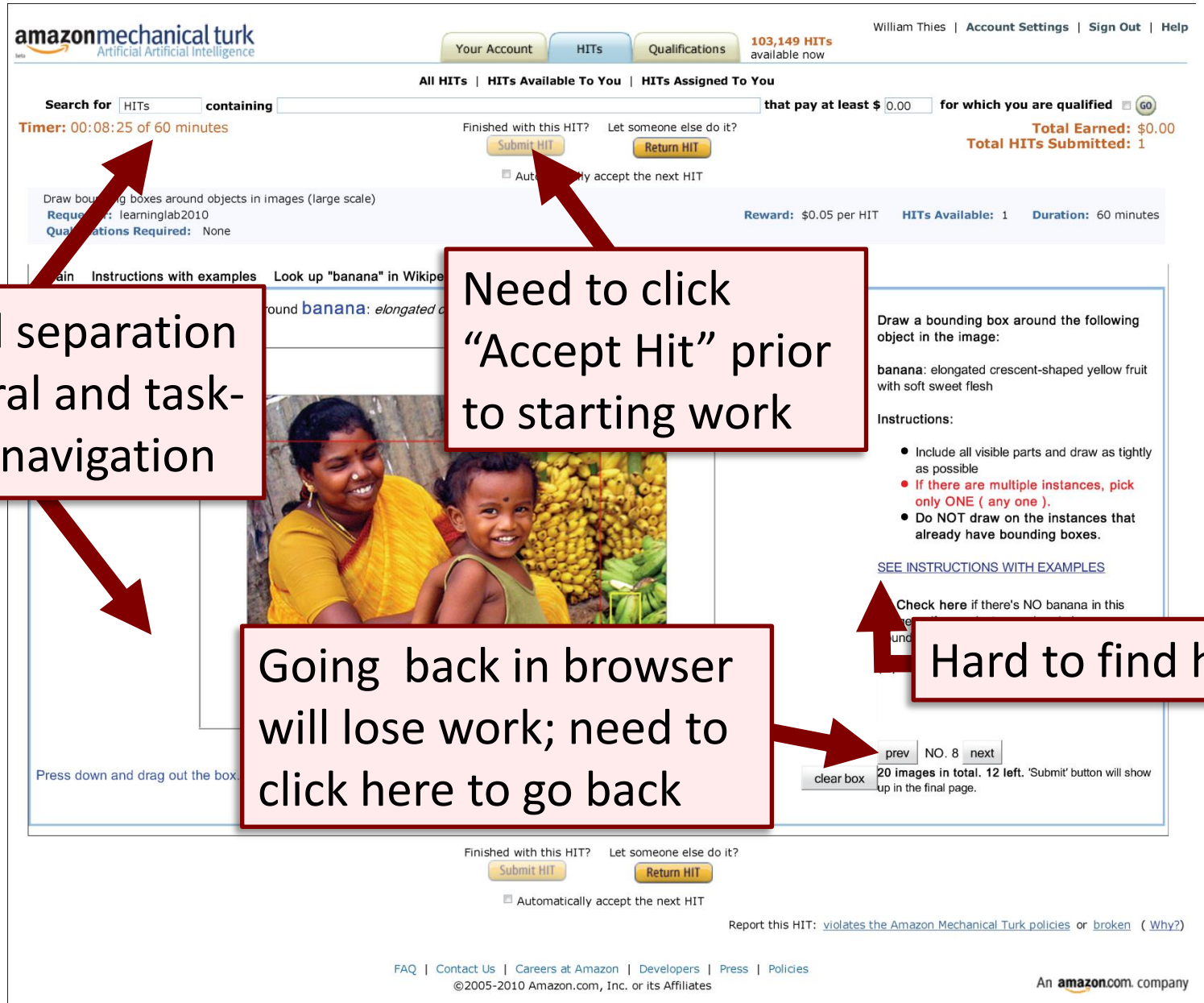
- With each of 7 participants:
  - Hour-long 1-on-1 session, providing help if needed
  - Participant registers on MTurk and attempts 1-2 tasks

Initial Observations

	Verify Address	Test New CAPTCHA	Label Image
Input Method	Text	Graphical	Graphical
Output Method	Text	Text	Graphical
Inherent Barriers to Completing Task	• Evaluating trust on Web • Nuanced use of language	• Ignoring truly illegible letters • Converting to unformatted text	(Unfamiliar with using click-and-drag interaction)

- With each of 7 participants:
  - Hour-long 1-on-1 session, providing help if needed
  - Participant registers on MTurk and attempts 1-2 tasks

Usability Barriers Across Tasks



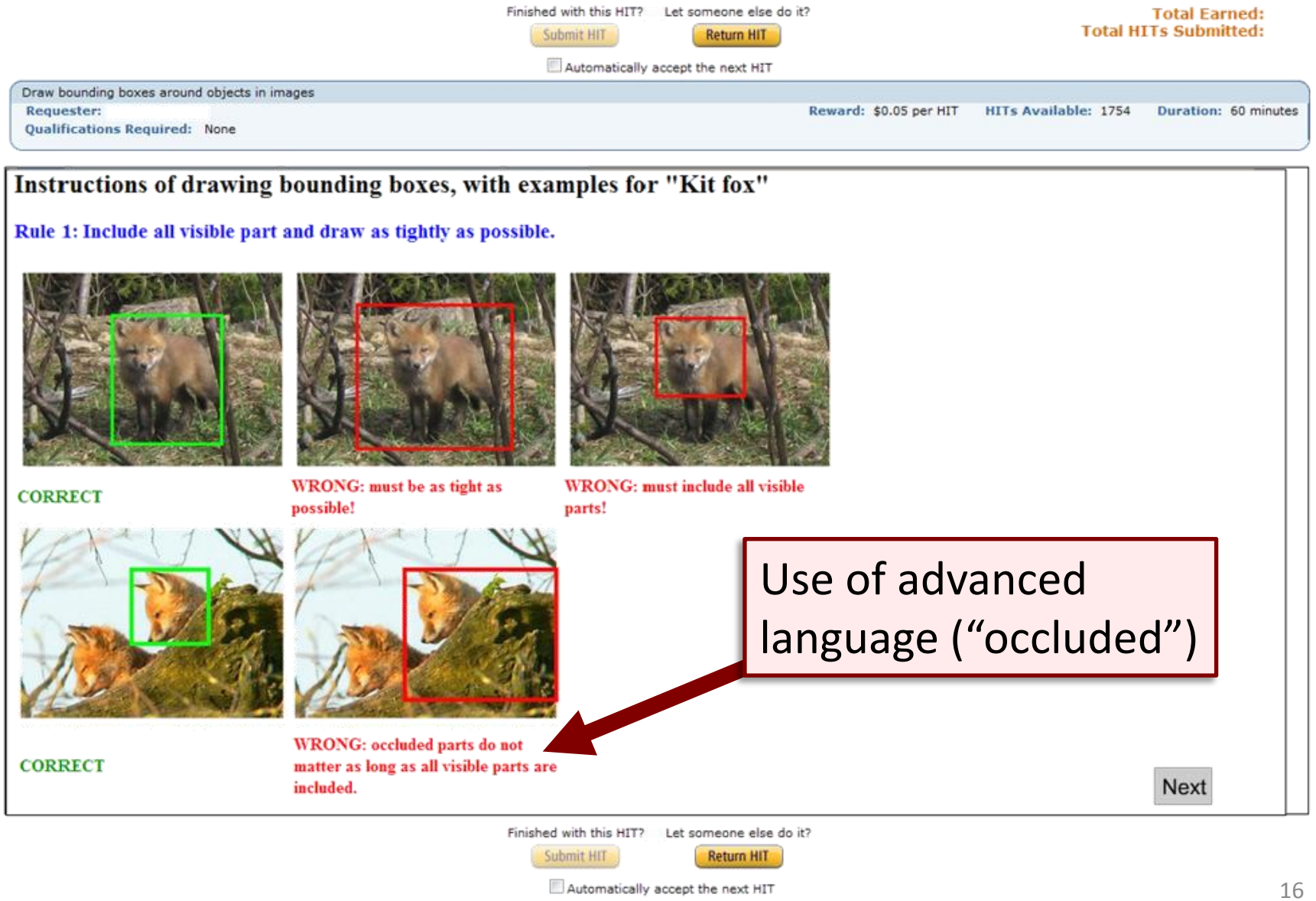
Minimal separation of general and task-specific navigation

Need to click "Accept Hit" prior to starting work

Going back in browser will lose work; need to click here to go back

Hard to find help

Difficulty Understanding the Instructions



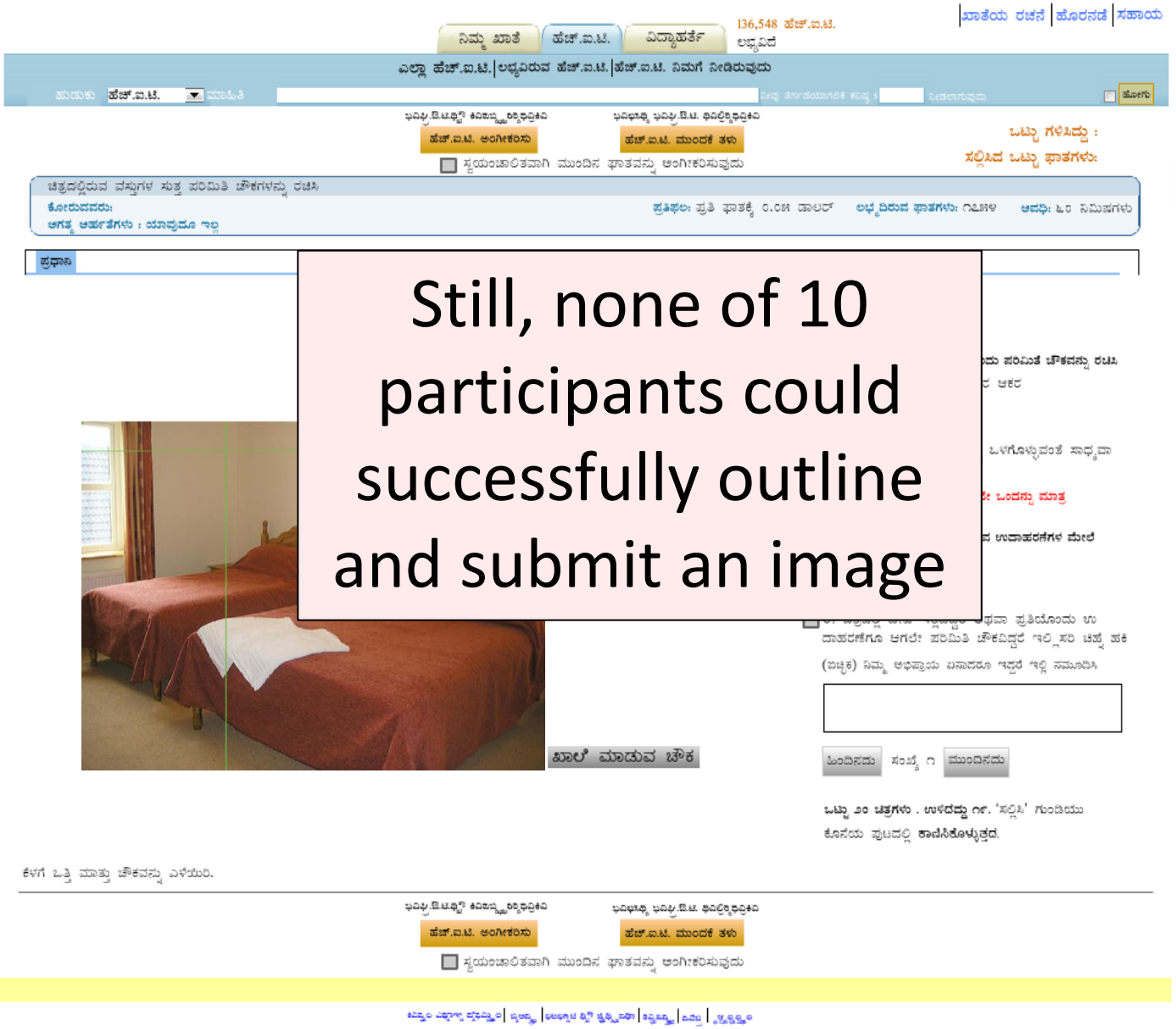
Difficulty Understanding the Instructions

System is Unusable Without Assistance

- None of 9 users could label an image in 30 min
- Methodology used in this talk:
  - Task: outline an object (lamp) in each of 20 images
    - Or indicate that no lamp is present
    - Maximum time: 30 minutes
  - Users receive an overview of MTurk
  - But NO assistance is offered in understanding or doing the task

Iterative Design and Evaluation

Design 1: Translation to Local Language



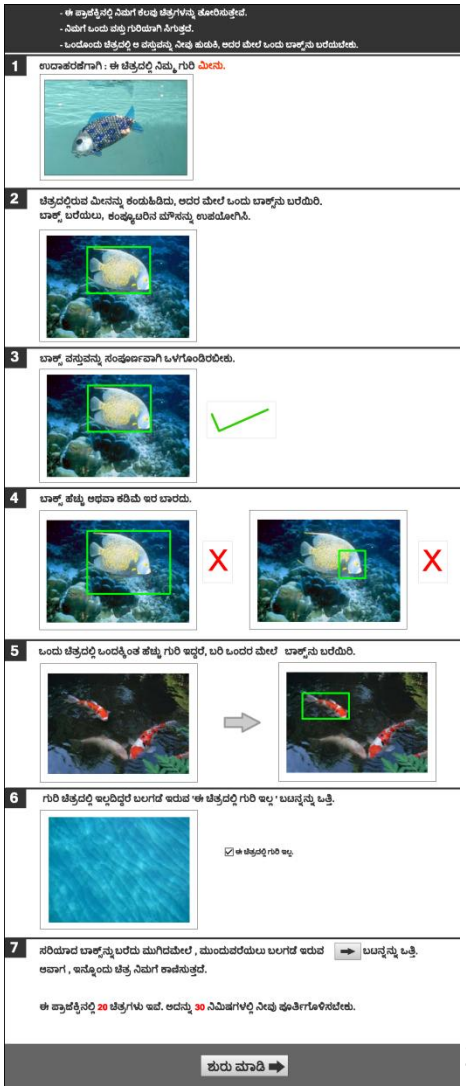
Design 2: New Instructions and Interface

Design 2: New Instructions and Interface

Original Instructions

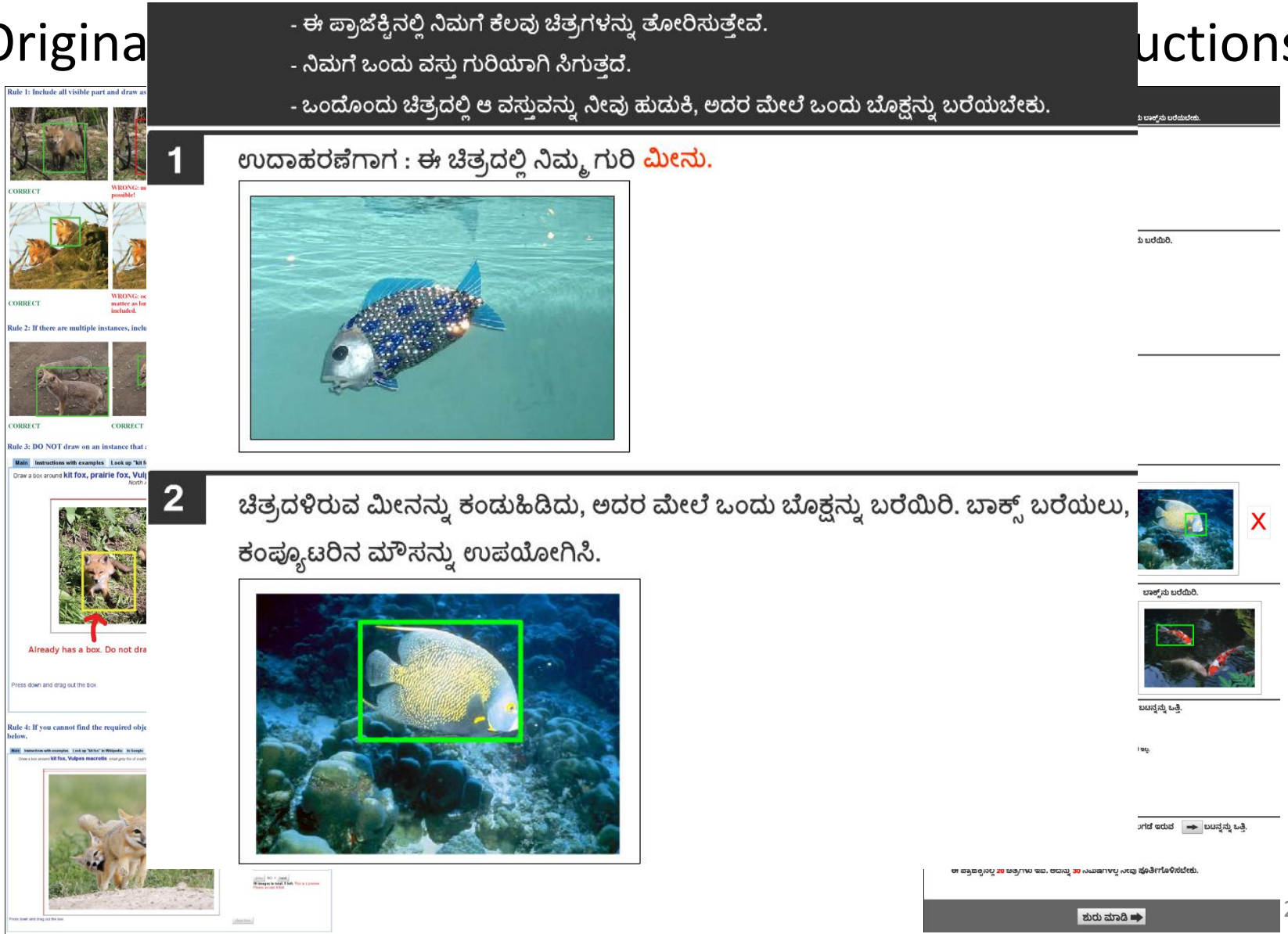
Add Structure
Simplify Language
Improve Illustrations

New Instructions



Design 2: New Instructions and Interface

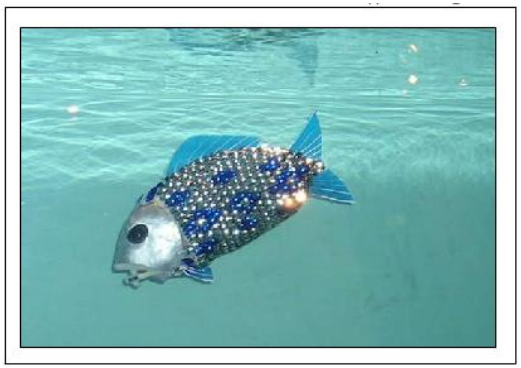
Origina uctions



Design 2: New Instructions and Interface

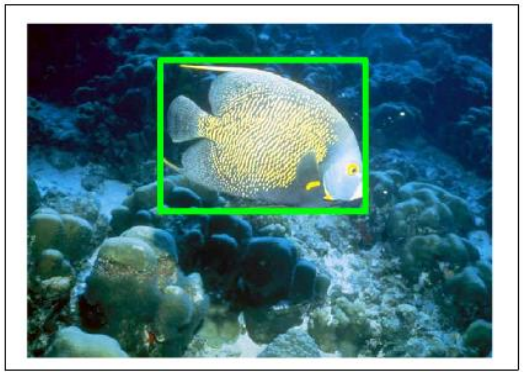
- In this project we will show you some pictures.
- You will get a target object.
- In each picture, you should search for that object and draw a box around it.

1



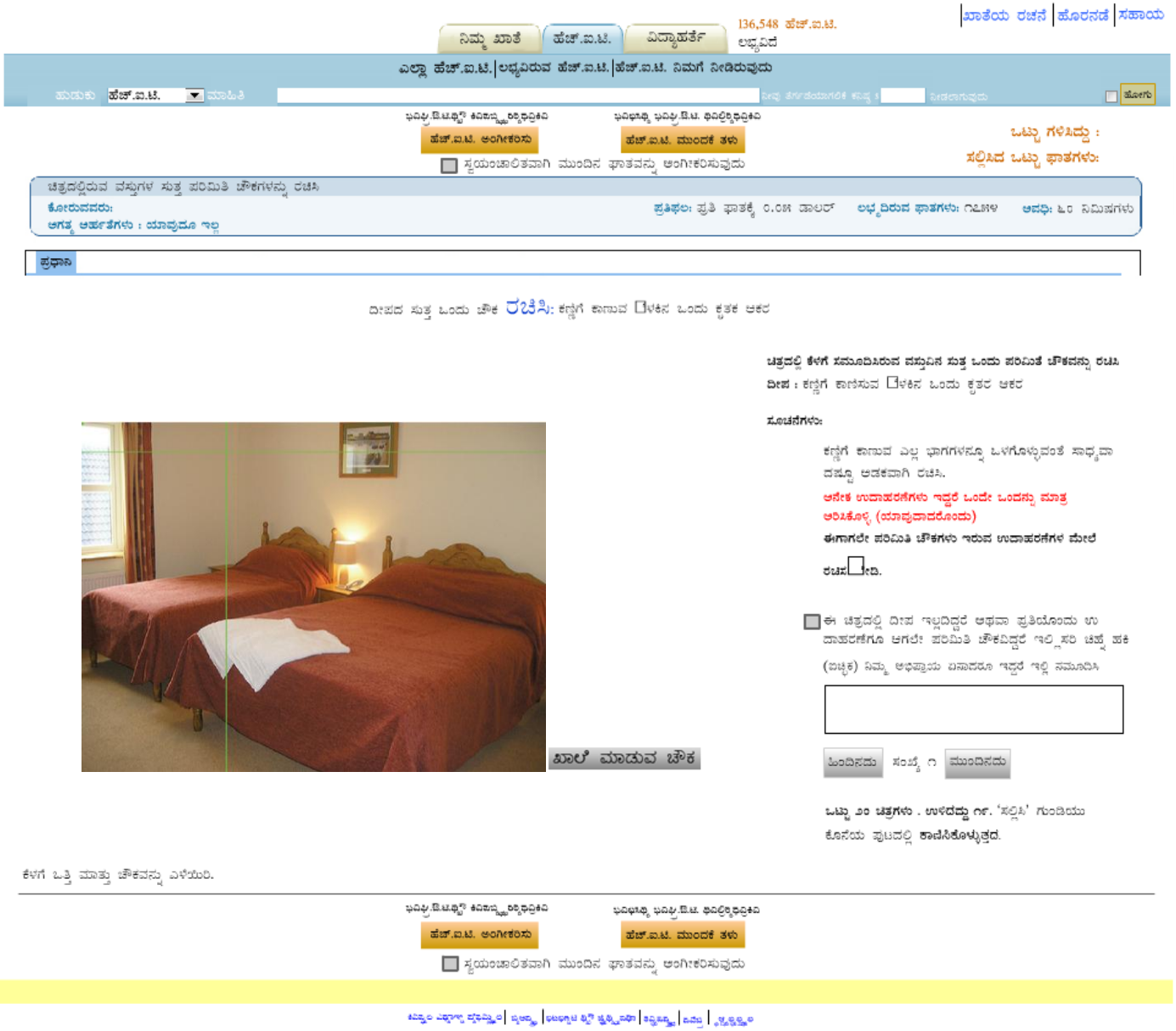
For example: In this picture, your target is fish.

2



Search and find the fish in the picture, and then draw a box around it. To draw the box, use the computer's mouse.

Design 2: New Instructions and Interface



Design 2: New Instructions and Interface

ಈ ಚಿತ್ರಗಳಲ್ಲಿ ನಿಮ್ಮ ಗುರಿ: ಲ್ಯಾಮ್ಪ್ (ದೀಪ).

ಒಂದೊಂದು ಚಿತ್ರದಲ್ಲಿರುವ ದೀಪವನ್ನು ಹುಡುಕಿ, ಆದರ ಮೇಲೆ ಬಾಕ್ಸ್ ಬರೆಯಿರಿ.

☐ ಈ ಚಿತ್ರದಲ್ಲಿ ಗುರಿ ಇಲ್ಲ.

◄ 3/20 ►

Design 2: New Instructions and Interface

- In this picture, your target is: lamp.
- Look for the lamp in each picture and draw a box over it.

☐ The target is not present in this picture.

3/20

Evaluation

Design	Images Annotated Correctly
0. Original MTurk (English)	0
1. Original MTurk (Kannada)	0

Evaluation

Design	Images Annotated Correctly
0. Original MTurk (English)	0
1. Original MTurk (Kannada)	0
2. New Instructions, New Interface (Kannada)	**66%**

Evaluation

Design	Images Annotated Correctly
0. Original MTurk (English)	0
1. Original MTurk (Kannada)	0
2. New Instructions, New Interface (Kannada)	66%

Evaluation

Design	Images Annotated Correctly
0. Original MTurk (English)	0
1. Original MTurk (Kannada)	0
2. New Instructions, New Interface (Kannada)	66%
3. Video Instructions, New Interface (Kannada)	

Evaluation

Design	Images Annotated Correctly
0. Original MTurk (English)	0
1. Original MTurk (Kannada)	0
2. New Instructions, New Interface (Kannada)	66%
3. Video Instructions, New Interface (Kannada)	**63%**

Evaluation

Design	Images Annotated Correctly
0. Original MTurk (English)	0
1. Original MTurk (Kannada)	0
2. New Instructions, New Interface (Kannada)	66%
3. Video Instructions, New Interface (Kannada)	63%

Evaluation

Design	Images Annotated Correctly
0. Original MTurk (English)	0
1. Original MTurk (Kannada)	0
2. New Instructions, New Interface (Kannada)	66%
3. Video Instructions, New Interface (Kannada)	63%
4. Video Instructions (Kannada), Original Interface (English)	

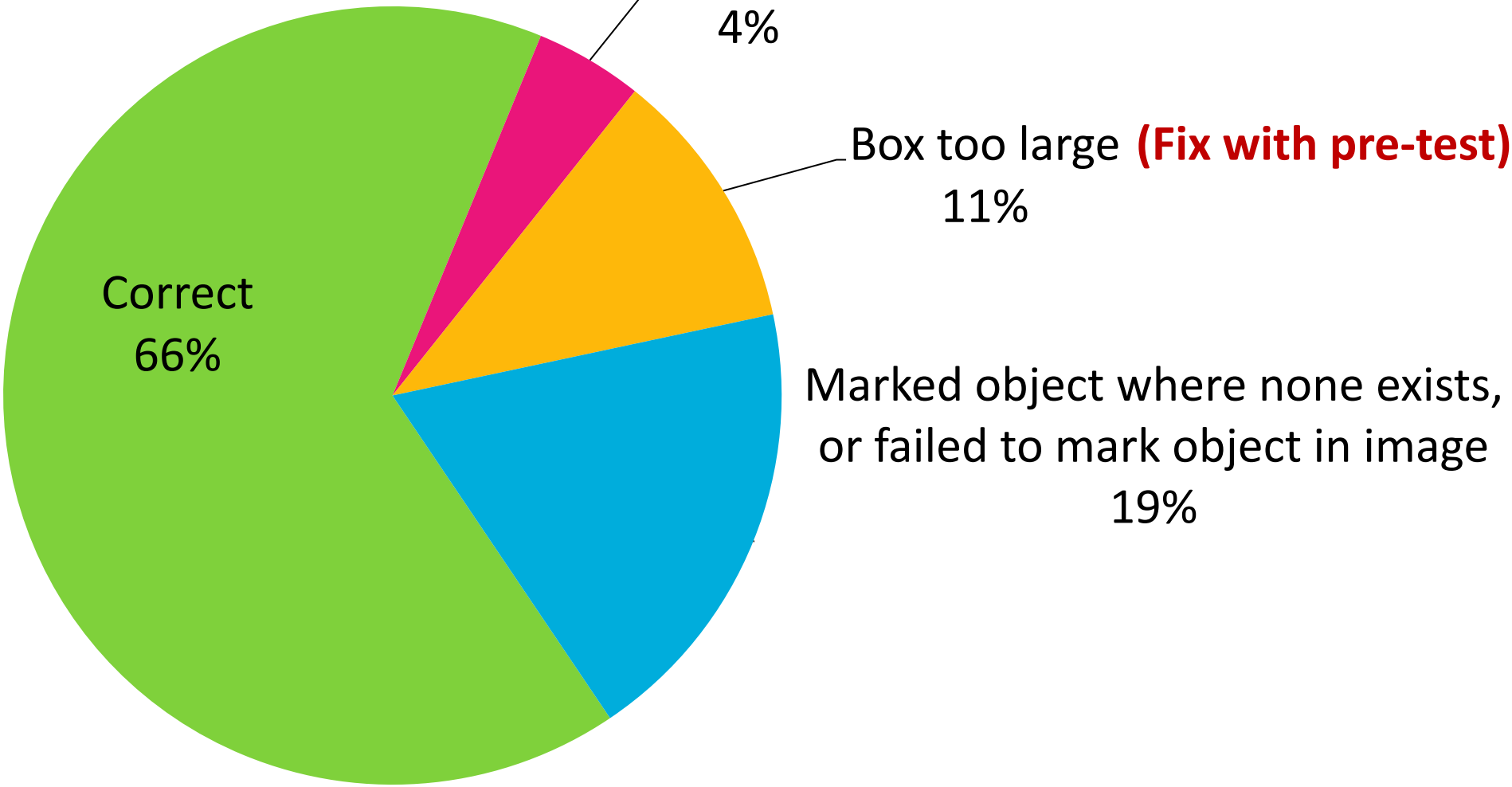
Evaluation

Design	Images Annotated Correctly
0. Original MTurk (English)	0
1. Original MTurk (Kannada)	0
2. New Instructions, New Interface (Kannada)	66%
3. Video Instructions, New Interface (Kannada)	63%
4. Video Instructions (Kannada), Original Interface (English)	**40%**

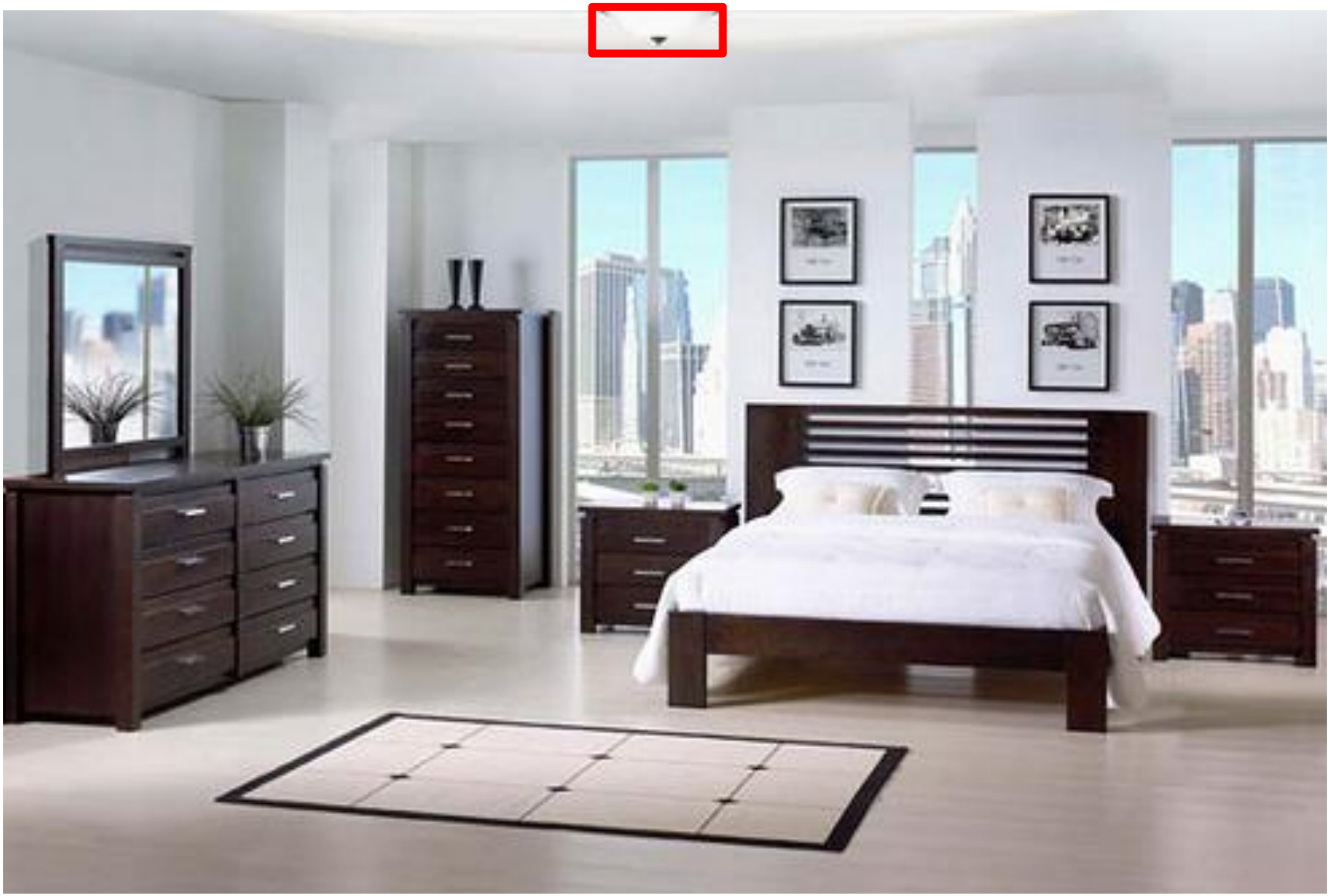
Evaluation

Design	Images Annotated Correctly
0. Original MTurk (English)	0
1. Original MTurk (Kannada)	0
2. New Instructions, New Interface (Kannada)	66% ✓
3. Video Instructions, New Interface (Kannada)	63% ✓
4. Video Instructions (Kannada), Original Interface (English)	40%

Sources of Error



Errors Due to Cultural Context?



Errors Due to Cultural Context?

Errors Due to Intrinsic Difficulty of Task

Disagreement among authors:

Participant found lamp that we did not:

Workers’ Earning Potential

Workers' Earnings Potential

- Bounding box tasks pays $0.05 for 20 images
  - Accuracy requirements unknown (we assume 75%)

	Time to Submit 20 Images	Gross Payment
Median participant	7m 20s	$0.41 / hr

- Baseline wage for median participant is $0.83 / hr

Workers' Earnings Potential

- Bounding box tasks pays $0.05 for 20 images
  - Accuracy requirements unknown (we assume 75%)

	Time to Submit 20 Images	Gross Payment
Fastest participant	1m 32s	$1.96 /hr
Median participant	7m 20s	$0.41 / hr
Slowest participant	23m 49s	$0.13 / hr

- Baseline wage for median participant is $0.83 / hr

Workers' Earnings Potential

- Bounding box tasks pays $0.05 for 20 images
  - Accuracy requirements unknown (we assume 75%)

	Time to Submit 20 Images	Gross Payment	Net Earnings (paying $0.30 / hr for Internet)
Fastest participant	1m 32s	$1.96 /hr	$1.52 / hr
Median participant	7m 20s	$0.41 / hr	$0.11 / hr
Slowest participant	23m 49s	$0.13 / hr	-$0.17 / hr

- Baseline wage for median participant is $0.83 / hr

Conclusions

- MTurk has yet to reach low-income workers in India
- We expose new barriers to usage by this group
  - Textual tasks difficult, but graphical tasks within reach
  - Current instructions and interfaces are a bottleneck
- We demonstrate that new designs can overcome barriers, improving image labeling from 0 to 66%
- Additional research needed to improve earnings
  - Increasing speed of task completion
  - Reducing cost of computer access
  - Making it easier to author usable tasks

Extra Slides

Design Recommendations

How to Design Microtasking Sites for Low-Income Workers?

- Improved instructions and interfaces are needed
  - Use simple, clear illustrations for each task
  - Minimize visual complexity
  - Streamline navigation
  - Anticipate sequencing of steps
- Language localization is necessary but not sufficient
- Video instructions work comparably to simplified text instructions, and thus are unlikely to be worth it

MTurk and Professional Development

- Microtasking can pose hazards to workers [Zittrain'08]
  - No affiliation with a team
  - Inability to understand moral implications of work
  - No working regulations, e.g., on wages or hours
- Is not necessarily limited to menial tasks
  - Creative tasks: design logos, taglines, graphics, etc.
  - Skilled tasks: writing, copyediting, programming, etc.
  - Thus could be a pathway to higher-level employment
- Might be more suitable for supplemental income
  - Offers extreme flexibility relative to other employment