



Crowdsourcing Civil Engineering Tasks



Comparing the Effectiveness of Amazon Mechanical Turk Master Workers versus Engineering Graduate Students on Basic Virtual Wind Tunnel Analytics

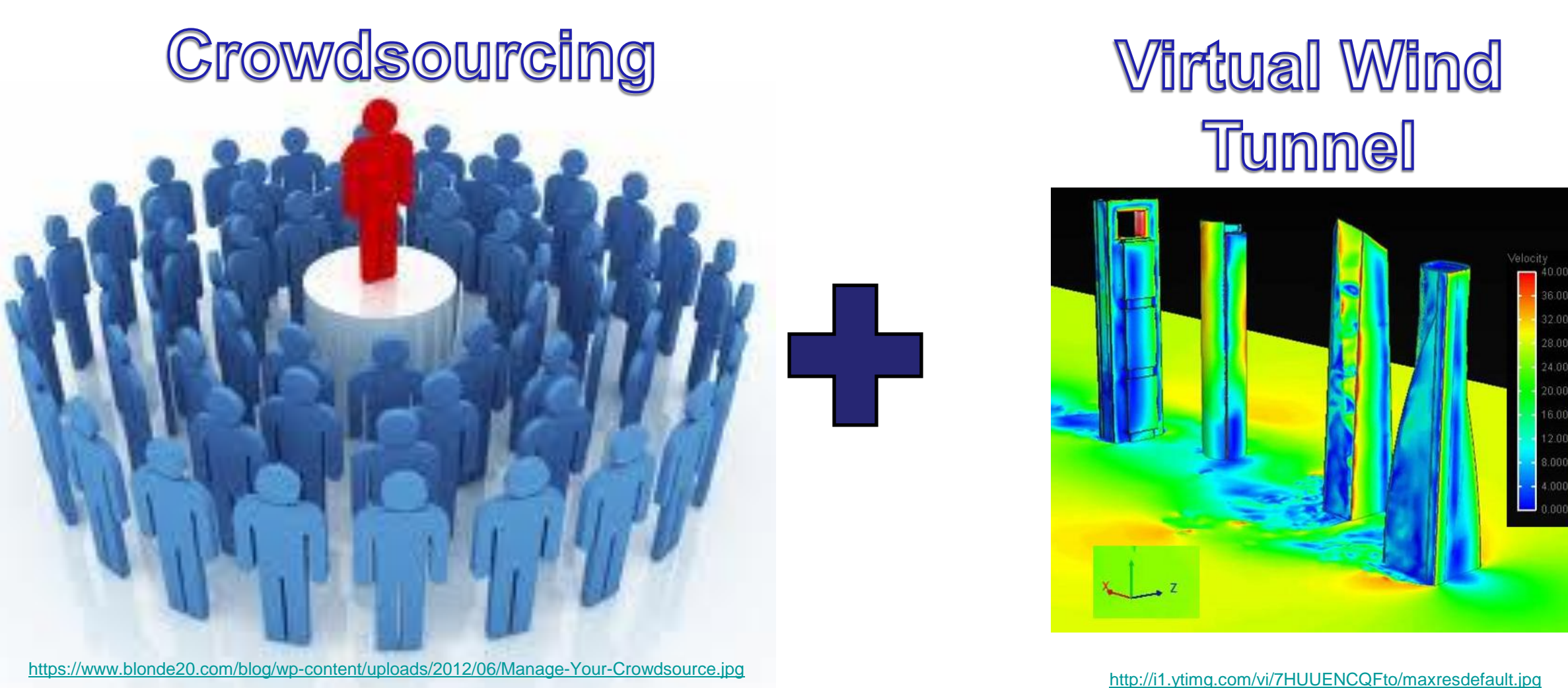
Presenter: Matthew Staffelbach

Team: Dr. Tracy Kijewski – Correa¹, Dr. Greg Madey², Dr. Zhi Zhai², Dr. David Hachen³, Peter Sempolinski, Dr. Daniel Wei¹, Dr. Ahsan Kareem¹, and Matthew Staffelbach²
Department of Computer Science and Engineering², Department of Sociology³, Department of Civil and Environmental Engineering and Earth Sciences¹, University of Notre Dame

Background

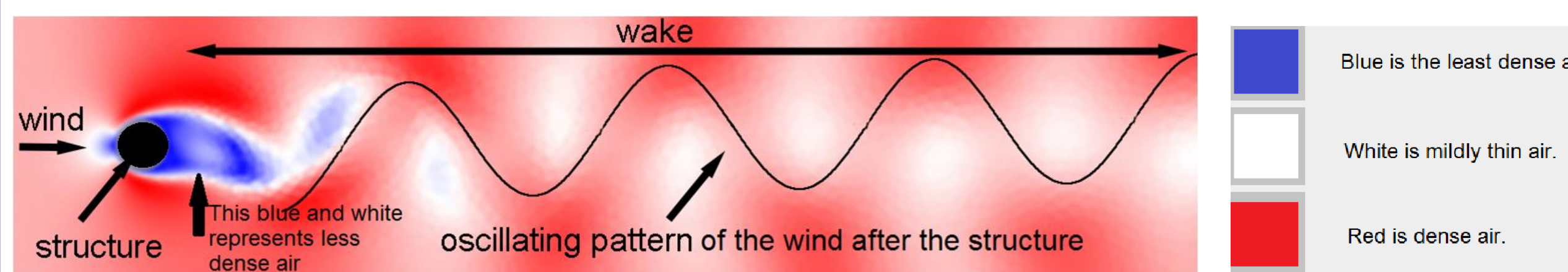
Crowdsourcing

Crowdsourcing is the practice of attaining needed ideas, services, or content by requesting contributions from a large group of people. Amazon Mechanical Turk (Mturk) is a web-marketplace for crowdsourcing microtasks, such as answering surveys, image tagging, and audio transcription. However we used Mturk for a more complicated task: Virtual Wind Tunneling Analytics. This meant that we had to engage in more training than employers who simply ask the turkers to fill out a personal survey.



Virtual Wind Tunnel Simulations

We used OpenFOAM which is a free, open source CFD software package. OpenFOAM is a program that can run wind tunneling experiments without costly wind tunnels or models. This allows engineers to run rapidly many simulations with differing objects (planes, skyscrapers, or missiles) and under differing wind conditions



Cool Fact: In general wind pulls buildings down more than it pushes them down.

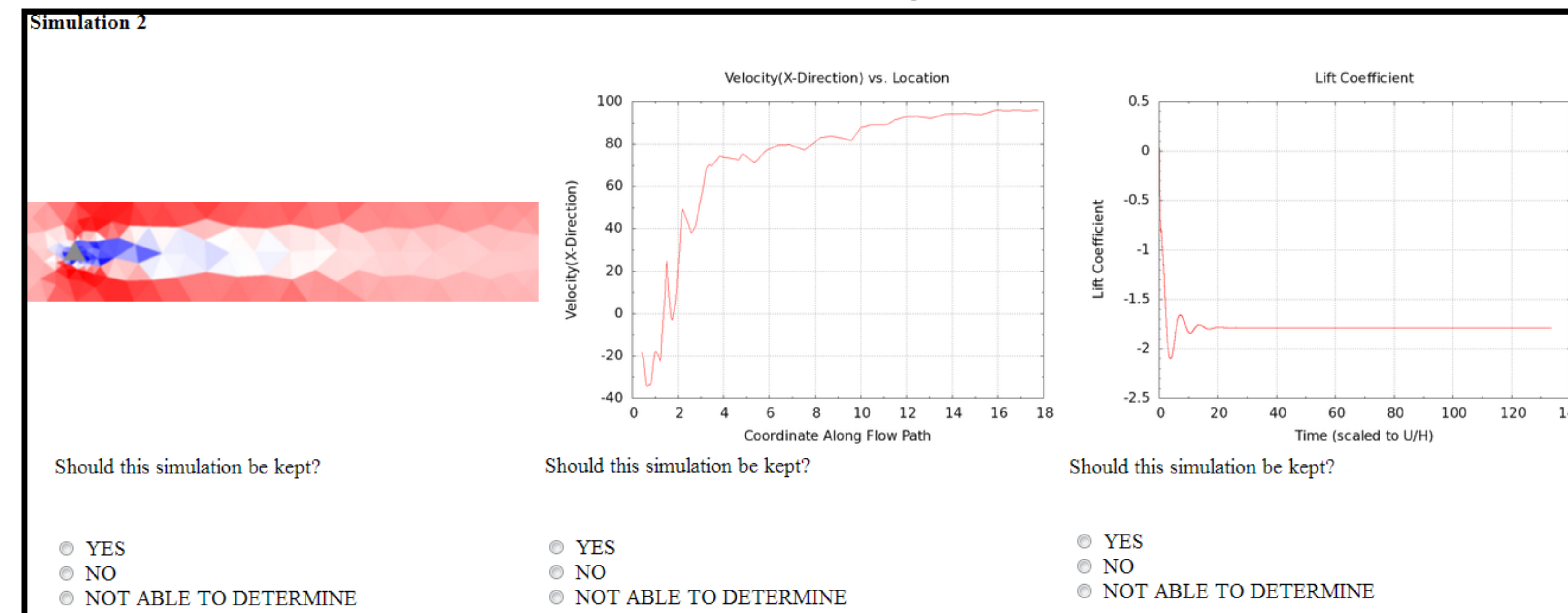
The less dense air which formed on the right side of this structure (the blue in this picture) exerted more force on the structure than the wind pushing on the building from the left.

Getting Started

Abstract

“Turkers” (online workers from Amazon Mechanical Turk) and Graduate students in Engineering from Notre Dame and China analyzed Virtual Wind Tunnel data which was generated using OpenFOAM. They were advised how to judge to quality of Wind Tunneling data graphs and were asked whether the simulations which we showed them contained quality data. Our goal was to measure the effectiveness of these “Turkers” compared with the graduate students.

Screenshot of a question



Motivation

Virtual Wind Tunnel Technology create a lot of useful data swiftly and cheaply. But if we cannot quickly analyze data, the speed of the data acquisition gets wasted by the slowness of the data analysis.

Can crowdsourcing be a cheap and quick solution to this problem? If so what else can crowdsourcing complete? Can crowdworkers be trained to effectively run simulations? Can crowd-workers design new structures? Can crowdworkers design, test, cross check new building structures? But, first we had to teach them to analyze data.

Who are the Turkers?

Gender

Male	Female
14	18

Education

Anonymous	High school	Some College	2 Year College	4 Year College	Advanced
1	2	7	3	13	6

Age

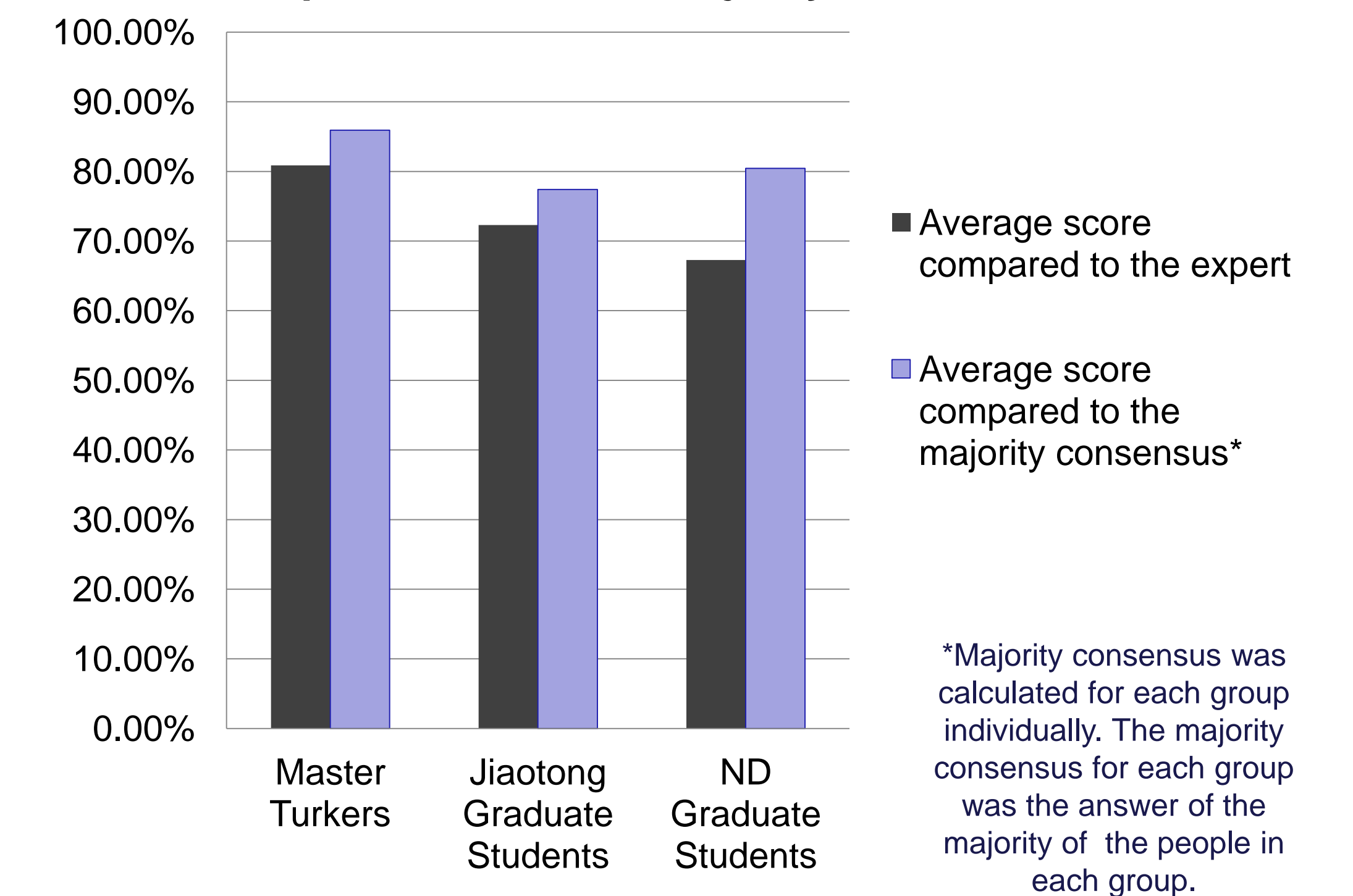
21-25	26-30	31-40	41-50	51-60	60+
1	6	16	6	2	1

Turker Motivation

Other	Money	Challenge	Good Cause	Bored
1	23	4	1	3

Results and Discussion

Percent of agreement in each group with the expert and with the majority consensus*



Why are Master Turkers better than Engineering Graduate students? One possible reason is that the overall the master turkers had a lower standard for what quality data was than the experts. Experts are used to data analysis and could have brought their own previous conceptions for what data is worth keeping, whereas the master turkers just matched the standard of the tutorial.

Conclusions

Given a good enough tutorial and compensation that averages around \$7.50 an hour, turkers can and will complete long and complex tasks with comparable efficiency to graduate students in Engineering.

Future work includes creating and running tasks which include asking turkers to run their own simulations (create their own data) and asking turkers to design their own structures. We will then use our already in place method to sort through the data and locate the quality simulations.

Acknowledgements

The research presented on this poster was supported in part by an award from the National Science Foundation, under Grant No. CBET-0941565 for a project entitled *Open Sourcing the Design of Civil Infrastructure (OSD-CI)*