Auto Feedback 3DLP

Evaluating Effects of Automatic Feedback Aligned to a Learning Progression to Promote Knowledge-In-Use

How AI supports teachers and students in science learning?

Mao-Ren Zeng (Mao-Jen Tseng)

Research Associate

tsengma1@msu.edu

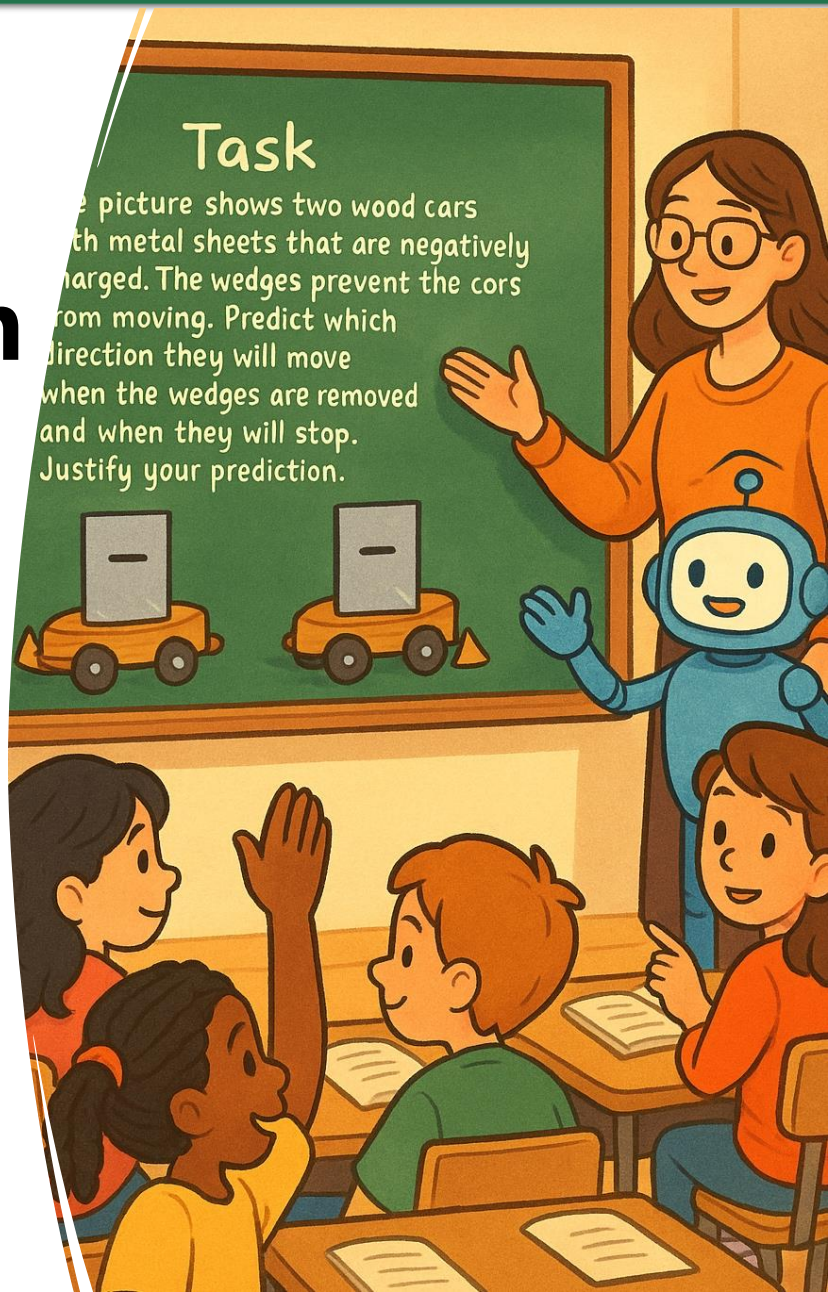
MICHIGAN STATE UNIVERSITY

CREATE for STEM Institute

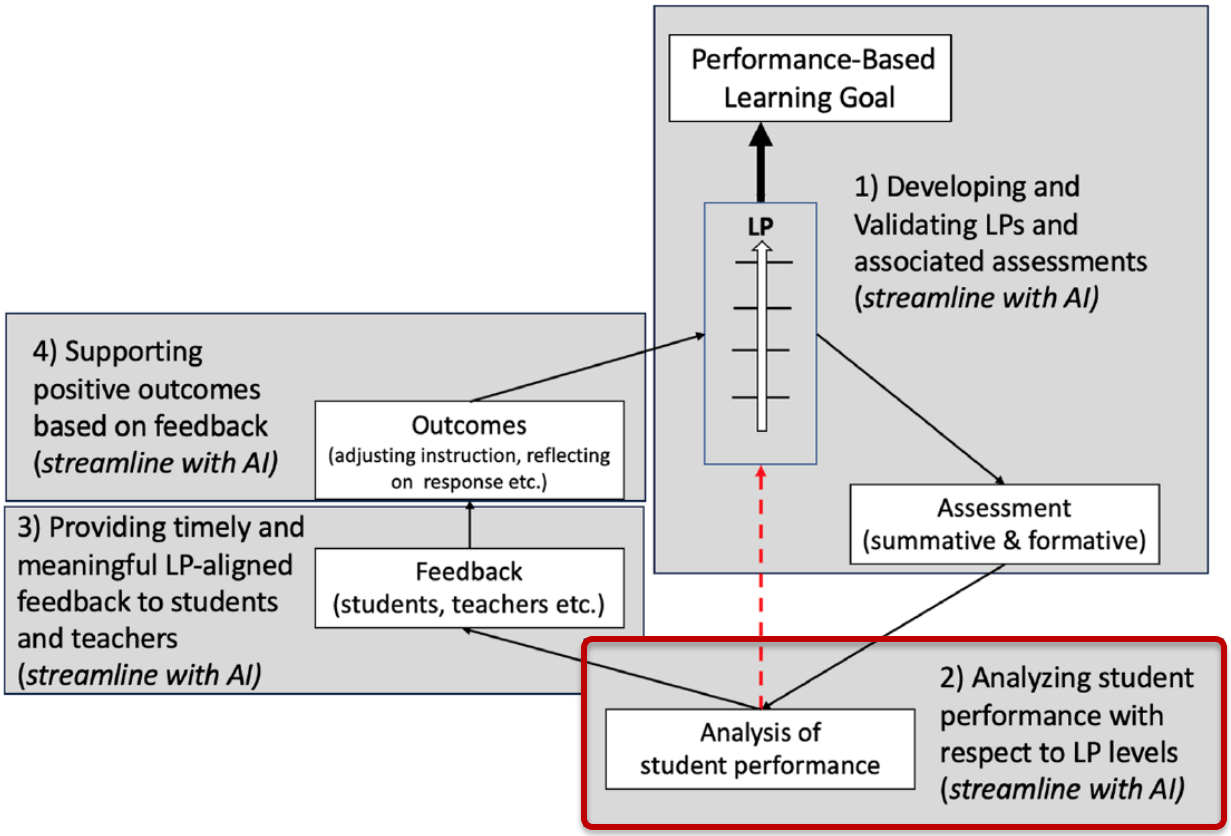
The Role of AI in Science Classroom

- **Analyzing student performance**
  - Assessing students' generated model by Machine Learning

- **Providing Feedbacks**
  - Generating meaningful feedback by Generative AI



AI Support Across Stages of LP-Based Assessment

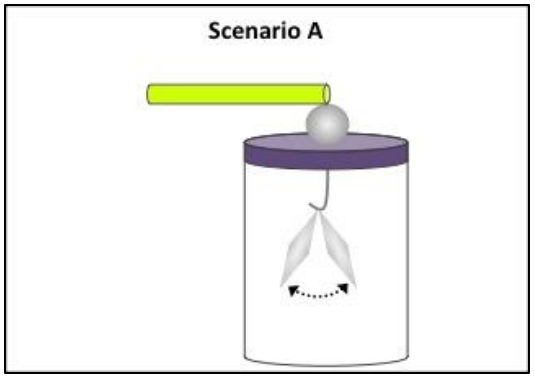


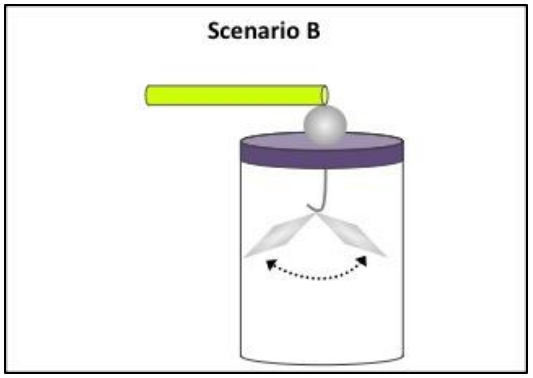
Kaldaras, L., Haudek, K., & Krajcik, J. (2024). Employing automatic analysis tools aligned to learning progressions to assess knowledge application and support learning in STEM. *International Journal of STEM Education, 11*, Article 57. https://doi.org/10.1186/s40594-024-00516-0

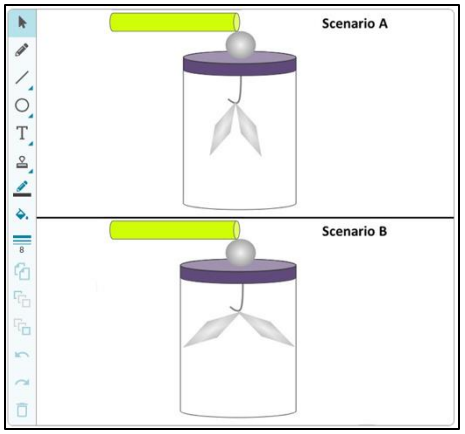
Electroscope Task

Scenario A below shows a diagram of what occurred in the video when a charged rod touched the ball.

In Scenario B, a rod touches the ball and makes the leaves move much further apart.







Task 1: Modeling

Draw a model to show what the differences are in the rod and foil leaves in the two scenarios.

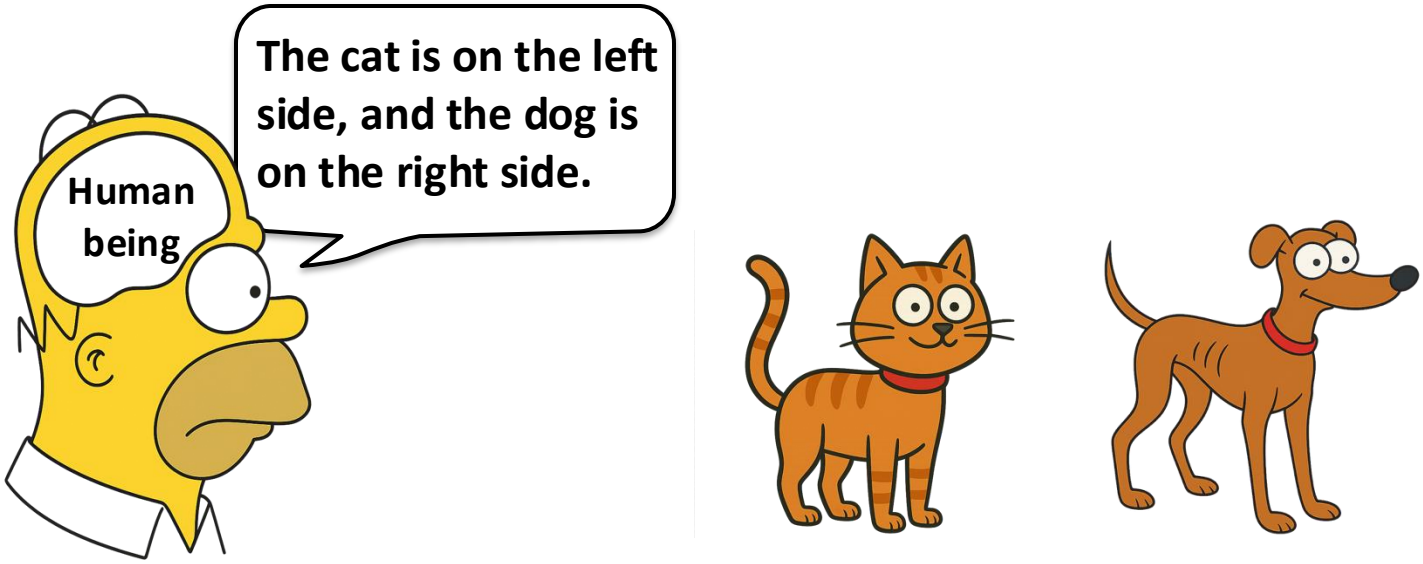
Task 2: Justification

What is different about Scenario A and Scenario B? Justify your answer.

Rubric	
Category	**Statement**
1	Point charge (either + or –) on the rod in scenario A
2	Point charge on the metal ball. The charge must be the same type as shown in the rod in scenario A.
3	Point charge on the hook of the electroscope. The charge must be the same type as shown on the rod in scenario A.
4	Point Charge on the leaves of the electroscope in scenario A. The charge must be the same type as shown in the rod in scenario A.
5	Clearly indicates repulsive Electric force causes leaves to move, by using arrows or force representations and pointing in opposite directions between the leaves in scenario A.

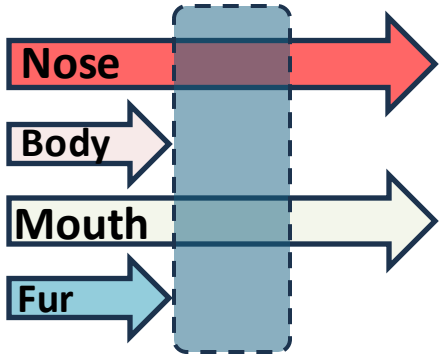
Challenge in this study… …

- How can machine “read” images?
- How can machine “understand” meaning of images?



Challenge in this study… …

- How can machine "read" images?
- How can machine "understand" meaning of images?



It is a Dog!
Not a Cat !

Select Features

Organize Information

Make Prediction

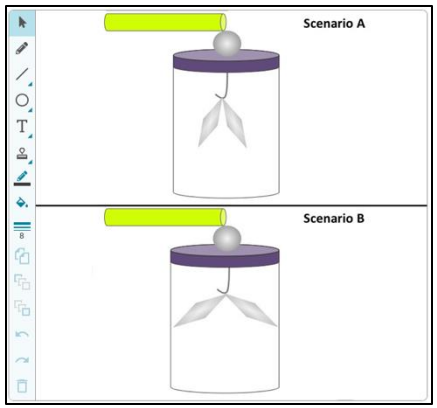
Object Detection

+

Multi-label classification

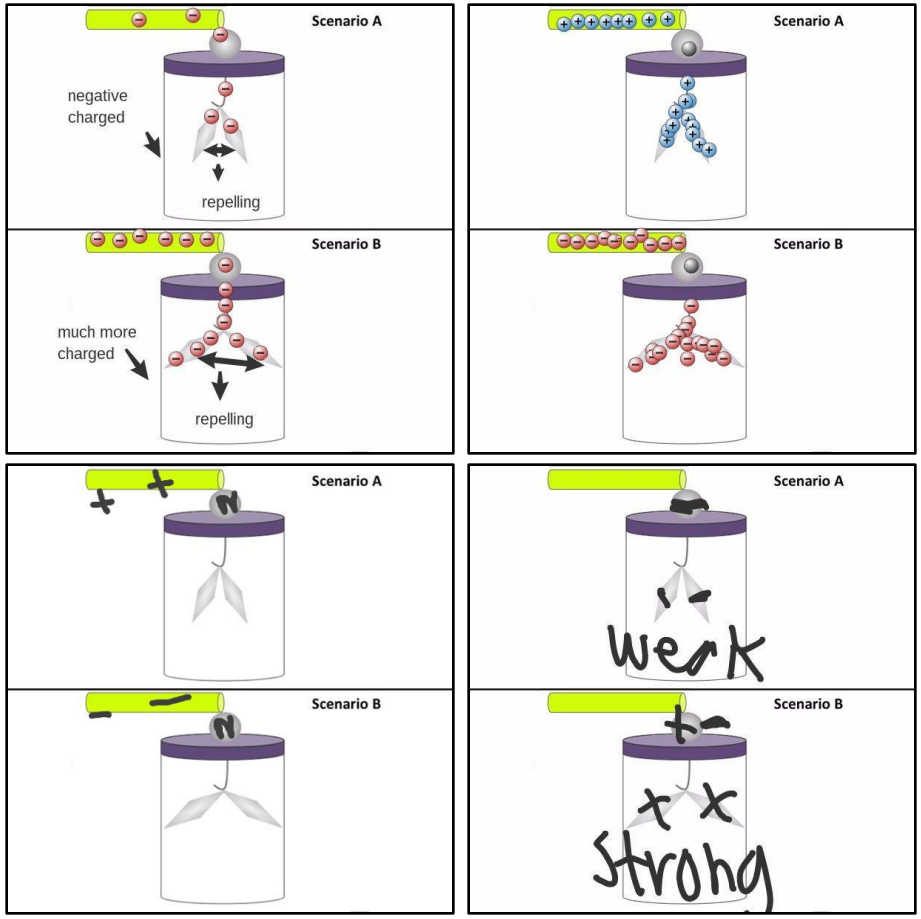
Dual-model scoring system

Students' generated model



Task 1: Modeling

Draw a model to show what the differences are in the rod and foil leaves in the two scenarios.

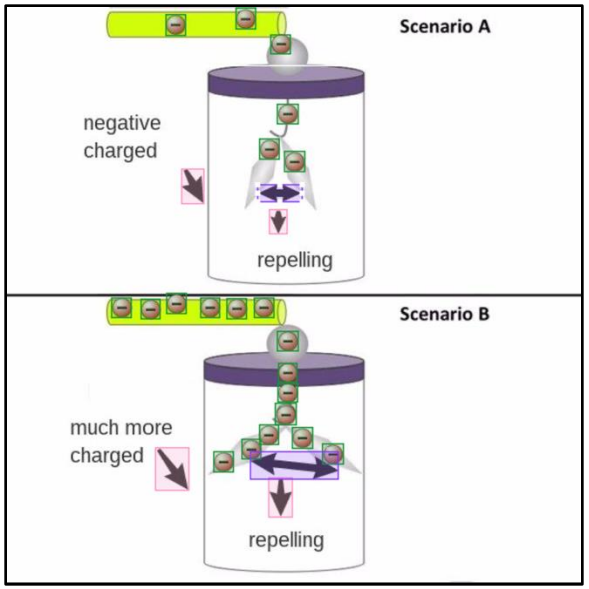


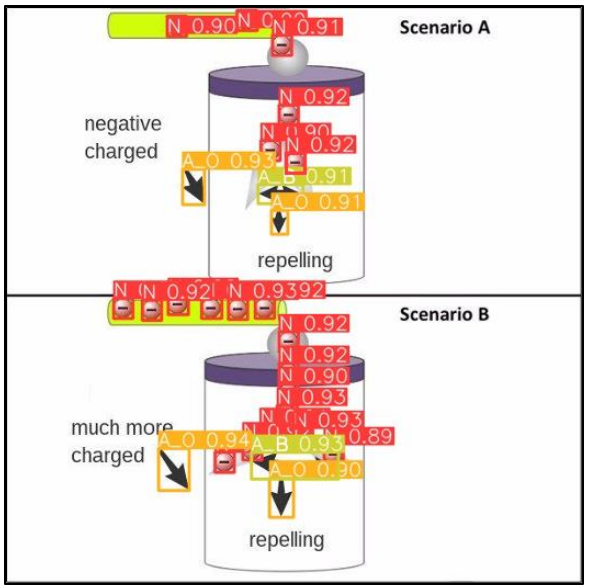
What are the features of the students' generated model?

Object detection task

- **Select features**
- **Provide the feature's Label and Location**

Rubric	
Category	**Statement**
1	Point charge (either + or –) on the rod in scenario A
2	Point charge on the metal ball. The charge must be the same type as shown in the rod in scenario A.
3	Point charge on the hook of the electroscope. The charge must be the same type as shown on the rod in scenario A.
4	Point Charge on the leaves of the electroscope in scenario A. The charge must be the same type as shown in the rod in scenario A.
5	Clearly indicates repulsive Electric force causes leaves to move, by using arrows or force representations and pointing in opposite directions between the leaves in scenario A.





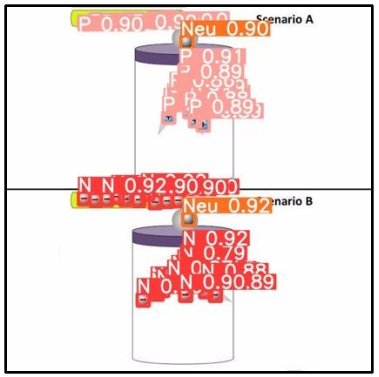
Rubric

Manual labeling

Machine detection the features

Object detection task

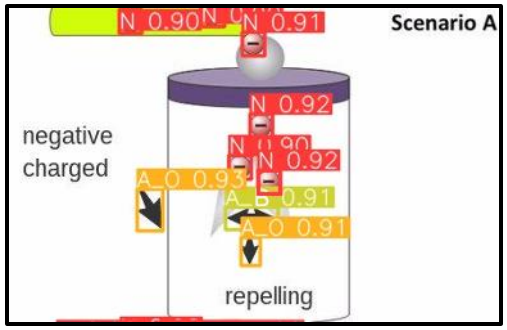
- We organized 496 images for training machine.
  (training: 396 images, validation: 50 images, and testing: 50 images)
- Object detection model performed well across individual features.
  - ✓ most achieving high mAP(mean Average Precision) values (> 0.7).



Features	Training stage		Testing stage	
	F1	$mAP_{.50}$	F1	$mAP_{.50}$
Overall performance	0.881	0.914	0.876	0.886
1. Negative charge	0.916	0.957	0.941	0.953
2. Positive charge	0.943	0.968	0.882	0.900
3. Neutral charge	0.769	0.808	1	0.995
4. One-directional arrow	0.823	0.877	0.637	0.650
5. Bidirectional arrow	0.921	0.959	0.915	0.931

Multi-label classification task

Rubric		Human scoring
Category	Statement	
1	Point charge (either + or –) on the rod in scenario A	1
2	Point charge on the metal ball. The charge must be the same type as shown in the rod in scenario A.	1
3	Point charge on the hook of the electroscope. The charge must be the same type as shown on the rod in scenario A.	1
4	Point Charge on the leaves of the electroscope in scenario A. The charge must be the same type as shown in the rod in scenario A.	1
5	Clearly indicates repulsive Electric force causes leaves to move, by using arrows or force representations and pointing in opposite directions between the leaves in scenario A.	1



- Human scoring
- Type of Features
- Location

Category 1

Point charge (either + or –) on the rod in scenario A

	Training stage			Testing stage		
Category	F1	Accuracy	Kappa	F1	Accuracy	Kappa
1	0.943	0.941	0.882	0.916	0.930	0.856
2	0.886	0.933	0.838	0.815	0.950	0.786
3	0.870	0.950	0.838	0.778	0.960	0.756
4	0.907	0.941	0.864	0.889	0.960	0.865
5	0.959	0.975	0.941	0.966	0.980	0.951

- ✓ Train machine
- ✓ Make prediction

Takeaway

Dual-model scoring system

Enhanced Interpretability

- The system measures a student's performance by identifying specific, meaningful symbols in their models. This symbol-level interpretability helps educators understand the specific conceptual components students are demonstrating.

Robust Transparency

- The system aligns object detection and classification with the rubric to assign meaning to findings. This provides the necessary transparency to build a robust validity argument for the AI-generated scores.

Providing meaningful information for teachers

- It provides evidence-based information that helps teachers reorganize their instruction and plan follow-up activities.
- Automated scoring reduces teachers' workload, and AI-based scoring increases teachers' willingness to adopt modeling practices in science classroom.

Auto Feedback 3DLP

Evaluating Effects of Automatic Feedback Aligned to a Learning Progression to Promote Knowledge-In-Use

Acknowledgements

We are grateful to the teachers and students who partnered with us.

Current & past team members:

Kevin Haudek · Joseph Krajcik · Leonora Kaldaras
Tingting Li · Selin Akgun · Clare Carlson
Wenxiu Tang · Steve Bennett · Mao-Ren Zeng

MICHIGAN STATE UNIVERSITY

ICER

This work was supported in part through computational resources and services provided by the Institute for Cyber-Enabled Research at Michigan State University.

NSF

This material is based upon work supported by the National Science Foundation (Grant No. 2200757). Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the supporting agencies

Auto Feedback 3DLP

Evaluating Effects of Automatic Feedback Aligned to a Learning Progression to Promote Knowledge-In-Use

Thank You!

Kevin Haudek
Principal Investigator
haudekke@msu.edu

Joseph Krajcik
Co-Principal Investigator
krajcik@msu.edu

Leonora Kaldaras
Co-Principal Investigator
lkaldara@central.uh.edu

Wenxiu Tang
Visiting Scholar
tangwen4@msu.edu

Steve Bennett
Research Associate
benne455@msu.edu

Mao-Ren Zeng
Research Associate
tsengma1@msu.edu

MICHIGAN STATE UNIVERSITY

Colleges of Education, Natural Science, Engineering, and Lyman Briggs

Office of the Provost