

From Sight to Skill: A Surgeon-Centered Augmented Reality System for Ureteroscopy Training



MAPLE
LAB

VISE
VANDERBILT INSTITUTE FOR
SURGERY AND ENGINEERING



VANDERBILT
UNIVERSITY

VANDERBILT UNIVERSITY
MEDICAL CENTER

Jumanh Atoum¹, Fangjie Li¹, Ayberk Acar¹, Nicholas L. Kavoussi², Jie Ying Wu¹

¹ Dept. of Computer Science, Vanderbilt University, Nashville, TN, USA

² Vanderbilt University Medical Center, Nashville, TN, USA

INTRODUCTION

- Time, patient safety constraints, and reliance on delayed verbal feedback limit surgical training
- Augmented reality (AR) gaze markers provide real-time feedback, improving efficiency in ureteroscopy
- User-centered design is critical to prevent cognitive overload and ensure AR usability in the operating room
- We previously conducted semi-structured interviews and co-design activities with surgical trainees
- to identify key factors in AR marker design [1]
- Based on findings from [1], we designed 3 AR gaze markers to provide intraoperative visual guidance
- Evaluated with all 3 AR markers and verbal guidance alone with 8 trainees and 1 expert urologist on 4 kidney phantoms of similar difficulty

[1] Atoum J. et al. (2025). Focus on the Experts: Co-designing an AR Eye-Gaze Tracking System with Surgical Trainees. arXiv:2506.21896.

Project Website

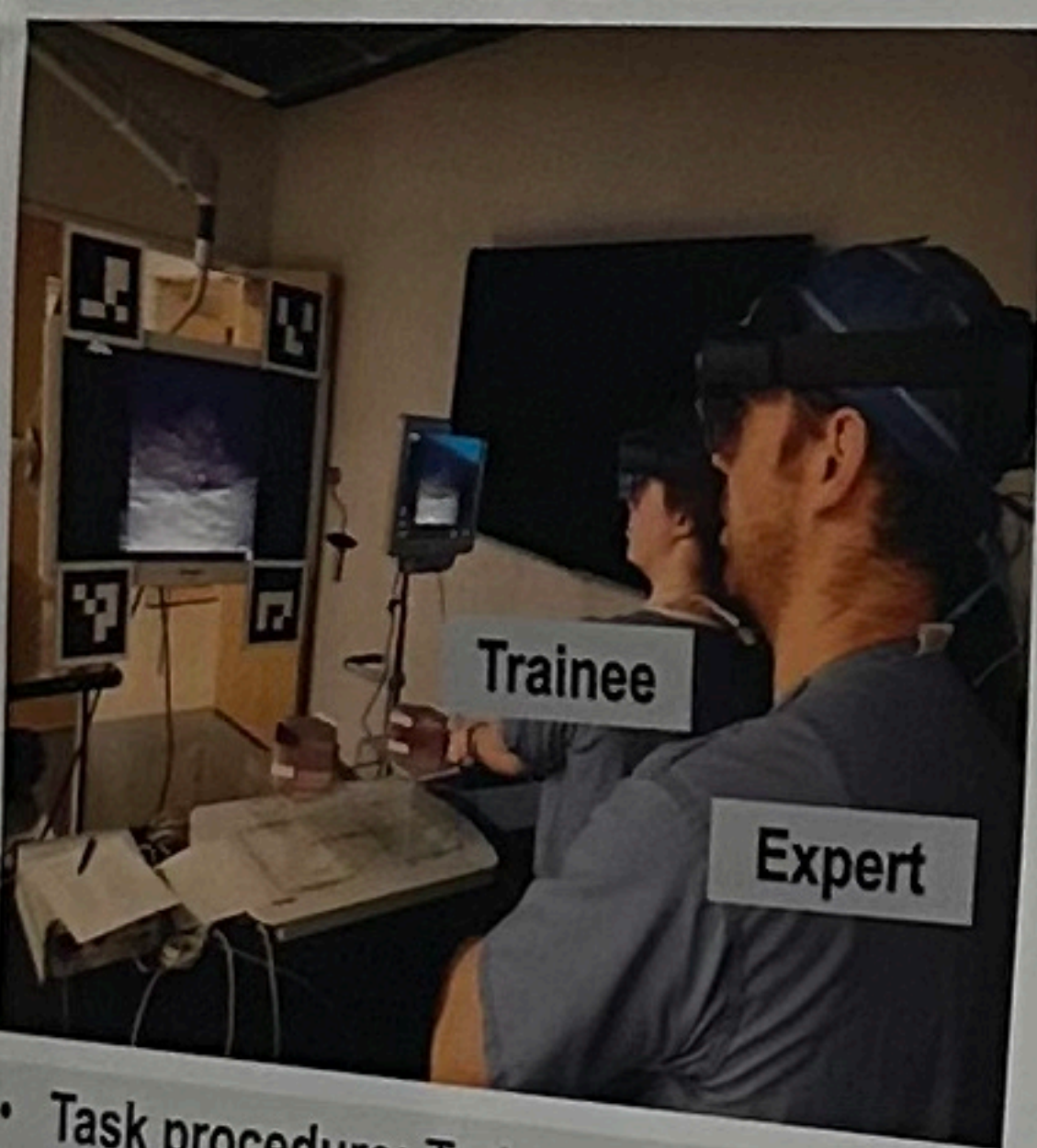


GitHub Website

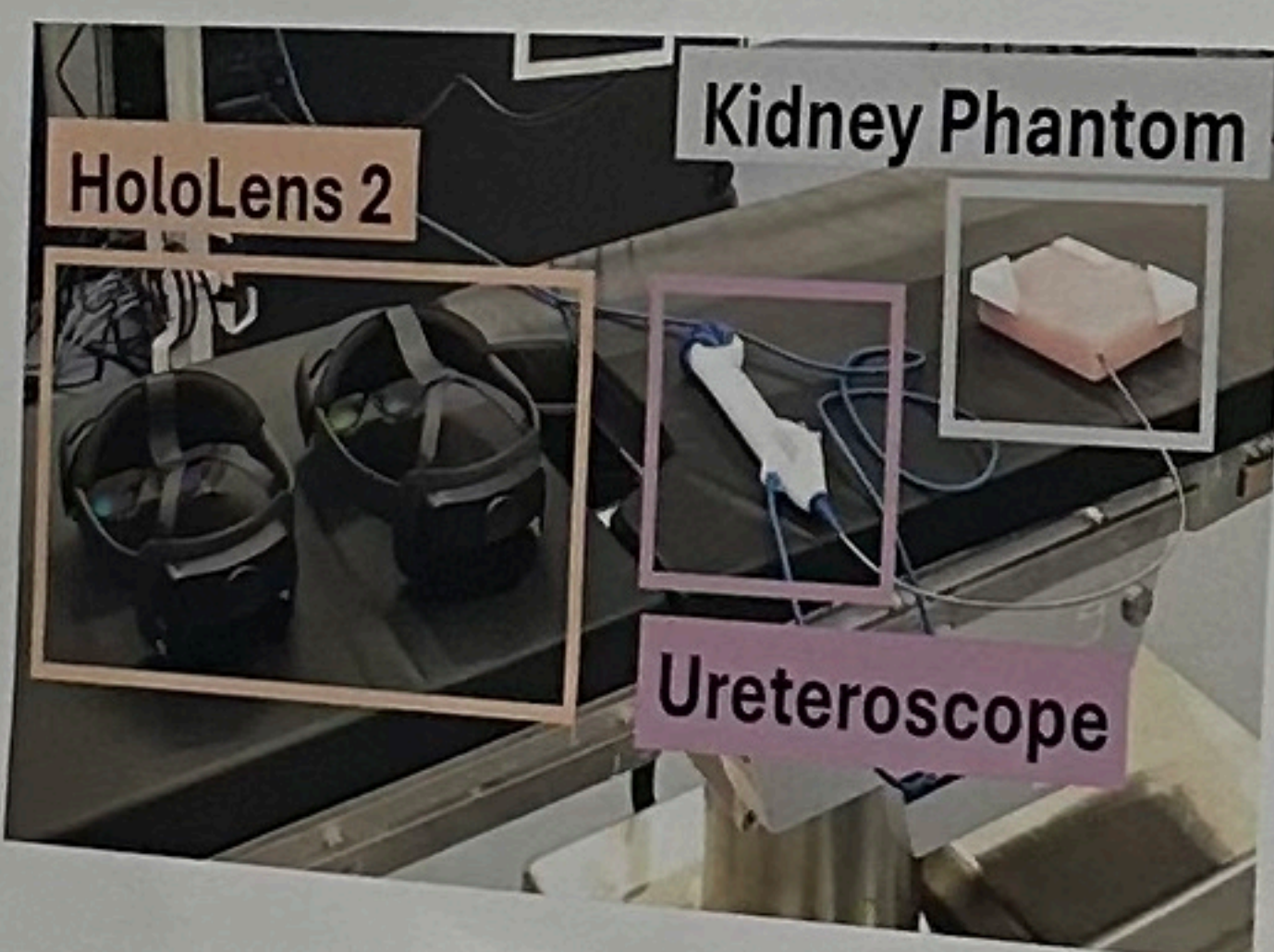


METHOD

User Study Setup



Experimental Setup



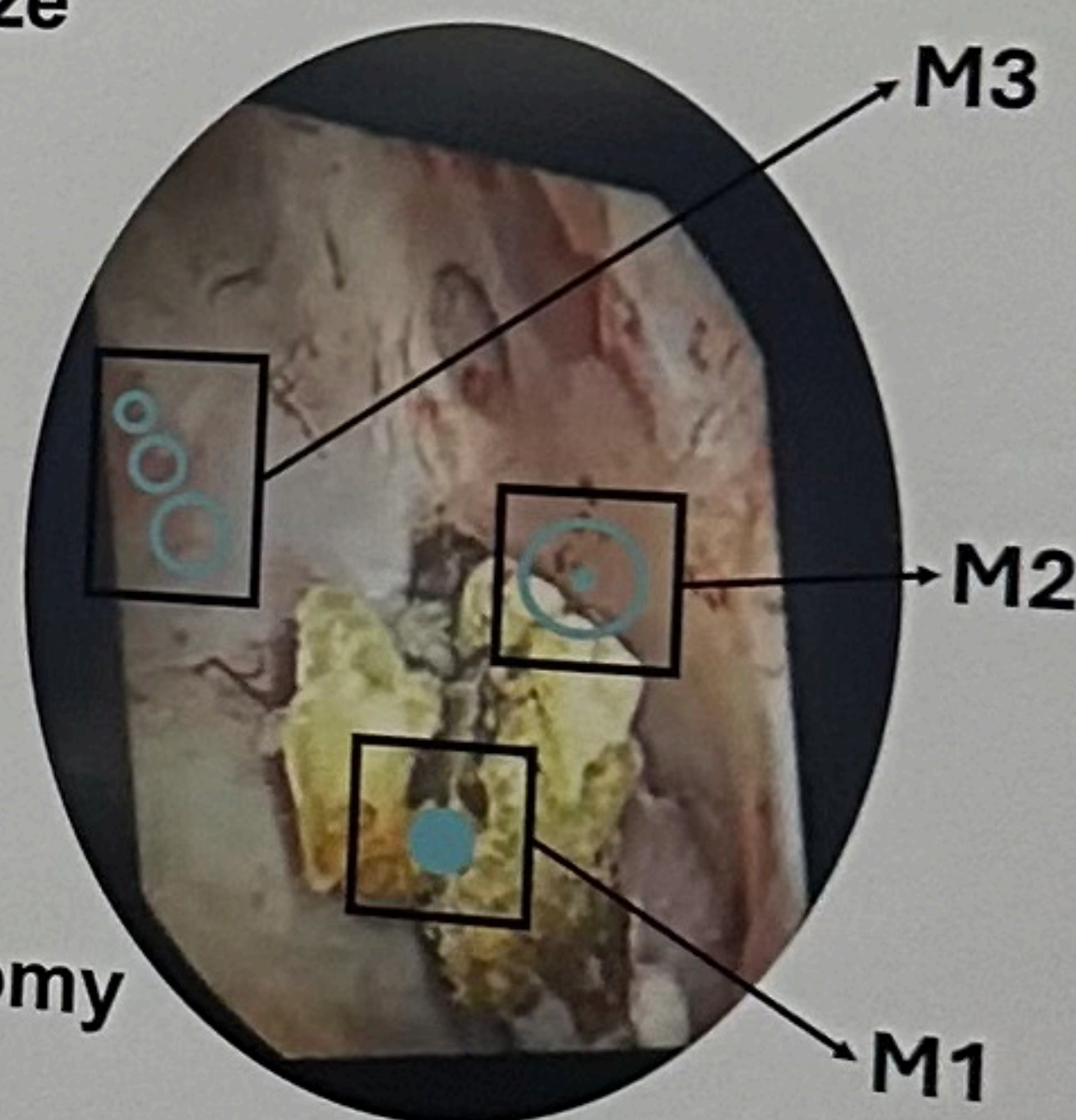
Marker Design Guidelines

Appropriate Size

Minimize
Distraction

High Color
Contrast

Preserve Anatomy
Visibility



- Task procedure: Trainee navigated ureteroscope with expert guidance, identified stones verbally, confirmed by expert and researcher
- Measures collected: 1- Task completion time, 2- Number of stones found, 3- Gaze metrics, 4- NASA-TLX
- Post-task feedback: Trainees reported marker preference.

RESULTS AND DISCUSSION

Subjective Preference



User's Marker Choice

[M2] is very subtle; I don't need to focus on it the whole time, but it still provides the needed guidance.

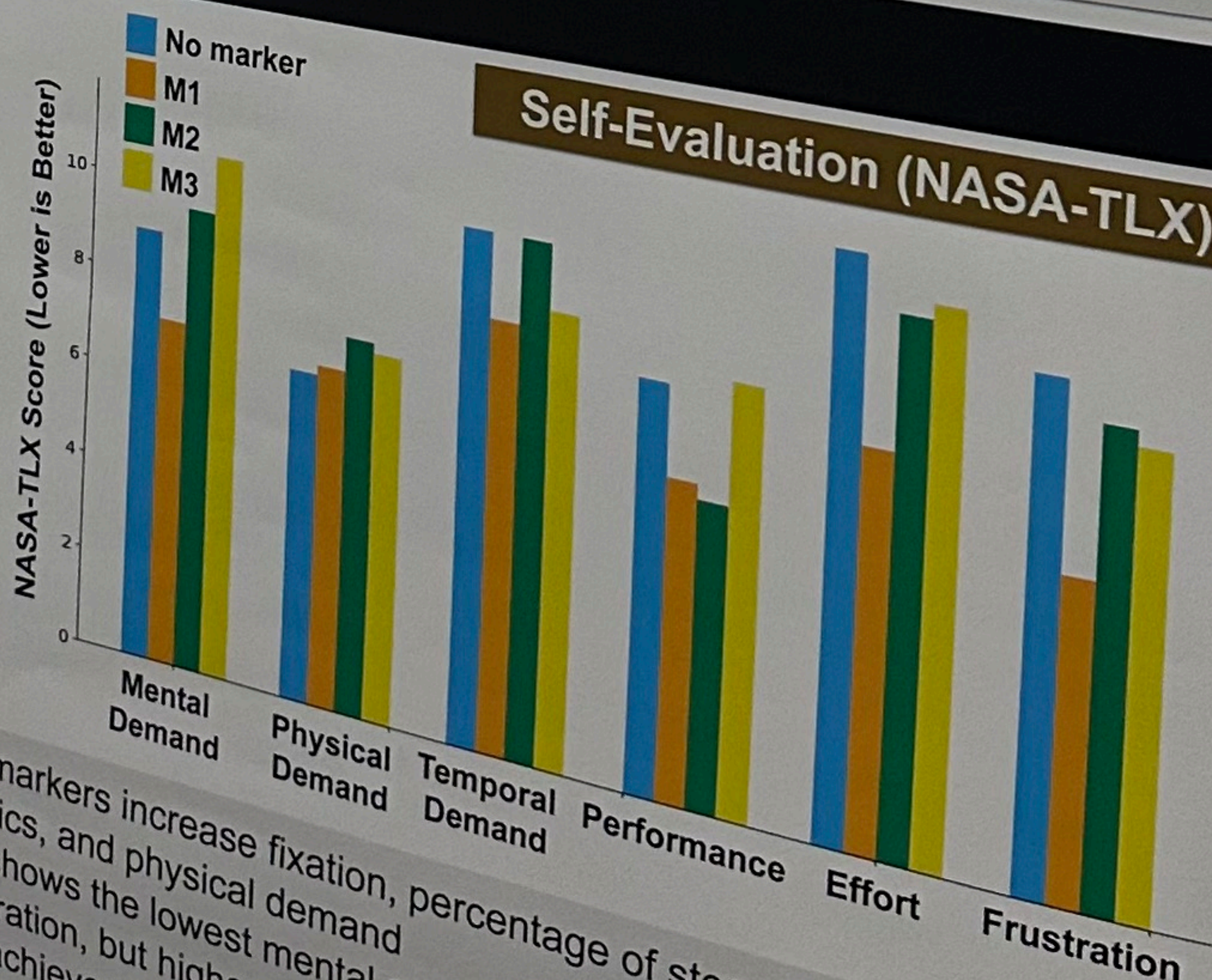
Objective Performance Metrics

Metric	No Marker	M1	M2	M3
Number of Stones	4.4 (23.9)	4.9 (20.5)	4.6 (24.7)	4.4 (23.9)
Completion Time (s)	73.7 (85.2)	72.1 (45.5)	84.7 (24.7)	88.6 (23.9)
Total Distance (cm)	523.4 (173.7)	590.1 (354.7)	502.0 (184.0)	531.3 (217.3)
Fixation/Saccades	40.7 (35.4)	51.3 (58.8)	42.1 (37.0)	52.7 (48.6)
Area in AOI (cm²)	132.2 (34.7)	148.2 (48.5)	138.9 (61.7)	122.3 (36.0)

CONCLUSION

- Well-designed markers improve task engagement – simple markers tend to work best
- Poorly designed markers can increase cognitive load and decrease performance
- Future work will evaluate with broader participant group

Self-Evaluation (NASA-TLX)



- AR markers increase fixation, percentage of stones found, gaze metrics, and physical demand
- M1 shows the lowest mental and temporal demand, but also high effort and frustration, but higher gaze distance travelled
- M2 achieves best performance in self-evaluation, showing task engagement, and reflects reduced cognitive load
- M3 design achieves the lowest percentage of stone found while incurring high mental demand and effort

ACKNOWLEDGMENT

This work was supported in part by the NIBIB of the NIH Grant 1R21EB035783 and the Wellcome Leap SAVE Program.