Measuring User Experiences of Interactions Between Digital and Physical Objects Through Augmented Reality and Internet of Things Devices

Connor Flick
Western Kentucky University
connorflick1@gmail.com

Courtney J. Harris
University of New Orleans
cjharri5@uno.edu

Nikolas Yonkers
University of South Florida
nyonkers@usf.edu

Nahal Norouzi
University of Central Florida
nahal.norouzi@knights.ucf.edu

Damla Turgut
University of Central Florida
Damla.Turgut@ucf.edu

Gerd Bruder
University of Central Florida
bruder@ucf.edu

Greg Welch
University of Central Florida
welch@ucf.edu

ABSTRACT

The use of digital interfaces to interact with physical devices has become, and is continuing to become, a common recurrence in homes, work spaces, and various industries around the world. Smart devices and Internet of Things (IoT) devices are replacing or being incorporated with traditional devices at a growing pace. Some of the current digital interfaces being used to control these IoT/smart devices include voice control via AI virtual assistants and smartphone apps. However, with augmented reality (AR) becoming more popular and accessible amongst average consumers, the seamless meshing of the digital realm with the physical realm through AR is becoming indispensable.

We explored using AR as another digital interface to interact with IoT devices to perform tasks of greater complexity. We conducted a comparative human-subjects study examining how AR interfaces are perceived against the already established methods of control, such as smartphones and voice controlled AI virtual assistants.

METHODS

We designed two novel AR methods, Gaze (**Fig 1.1**) & Gesture (**Fig 1.2**), and their networking processes **(Fig. 2)** to allow participants to control the IoT devices **(Fig. 3)**, as well as utilizing premade smartphone apps and IoT smart speakers to complete the same tasks. Development and design was informed by work completed in [1], [2], and [3].

Figure 1. User Action

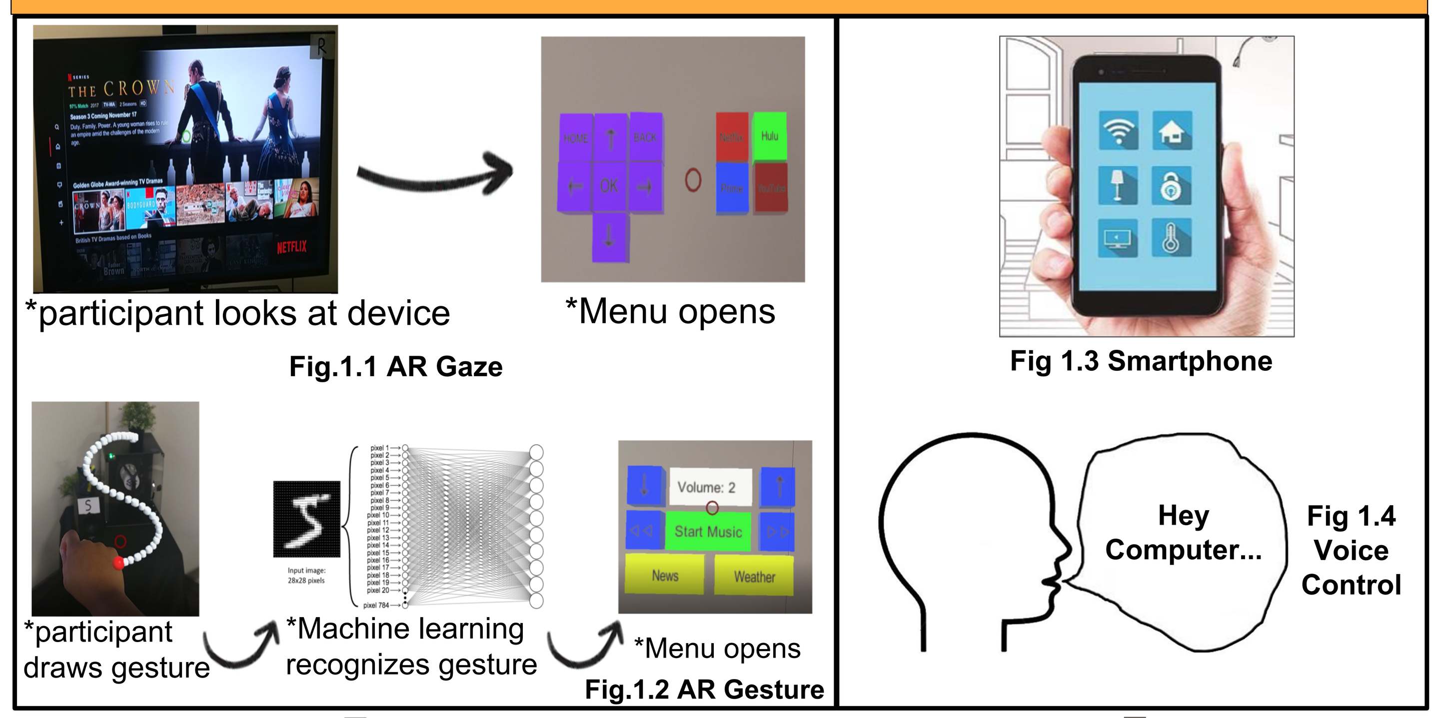


Figure 2. HoloLens Networking

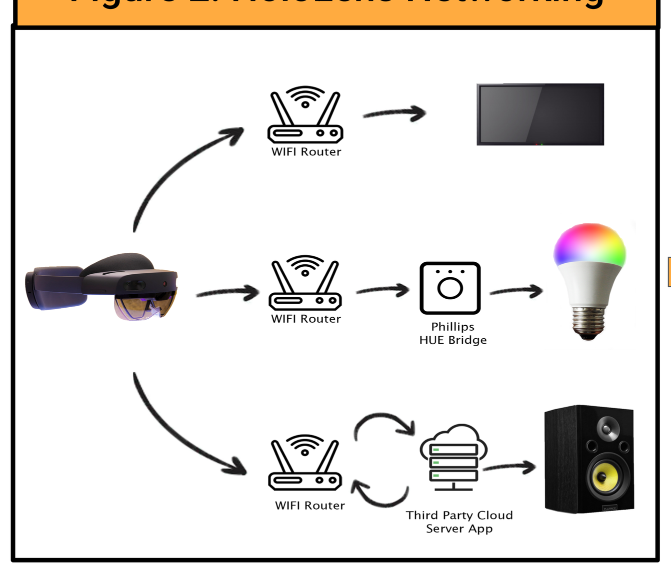
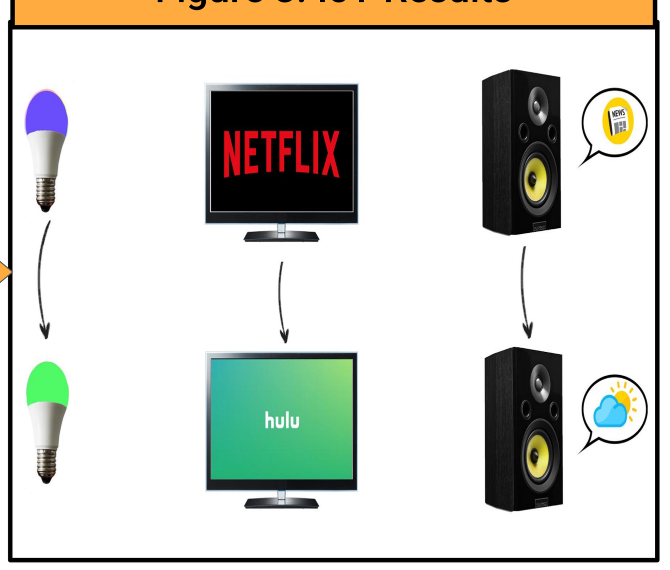


Figure 3. IoT Results



We used a full-factorial within-subjects design in this experiment. Participants were asked to complete a specific, random series of tasks with each of the methods to elicit responses from the IoT devices. Following the completion of the associated tasks, participants were asked to evaluate each method using the User Experience Questionnaire (UEQ), System Usability Scale (SUS), and the short form of the NASA Task Load Index (TLX). Once all tasks had been completed with all methods, participants ranked their preferences for each of the methods, given a statement or scenario.

RESULTS

Figure 4. A participant uses the Gaze method

We used parametric statistical tests to analyze the responses. The following figures **(Fig. 6 – 10)** demonstrate the significant findings from these tests.

*Statistical significance is shown by a starred line between categories.

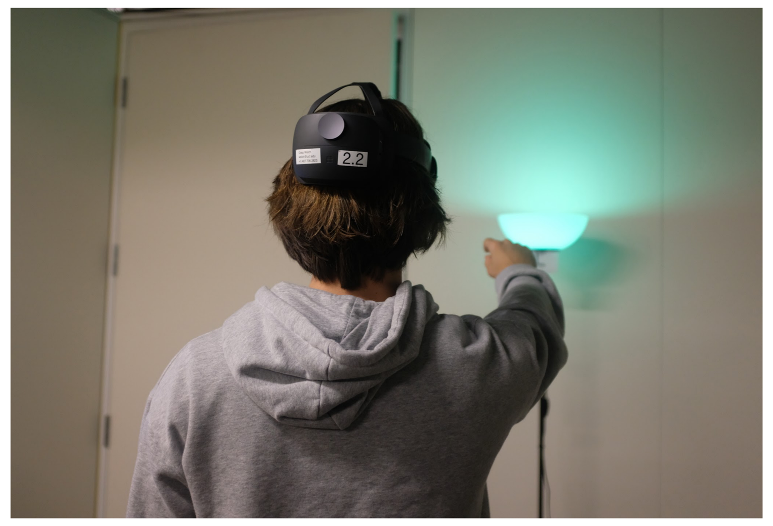


Figure 5. A participant uses the Gesture method

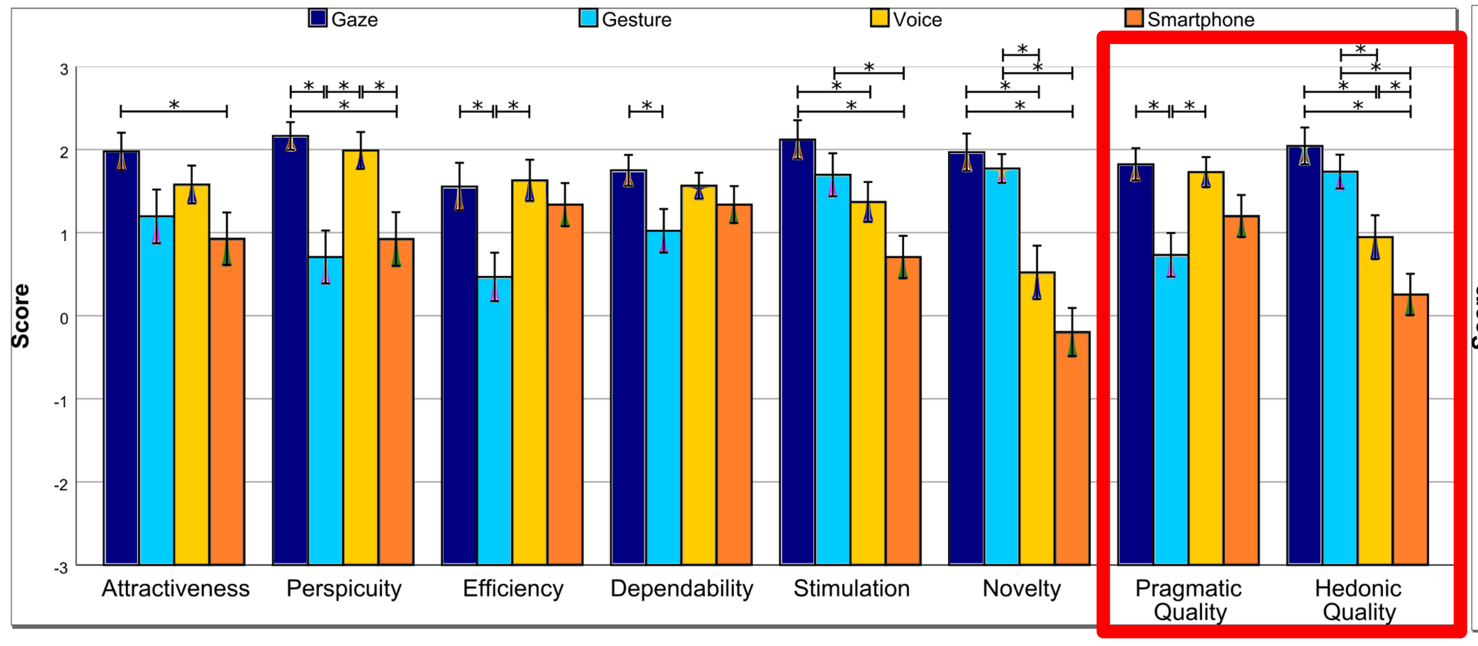


Figure 6. Subjective responses to the User Experience Questionnaire (UEQ)

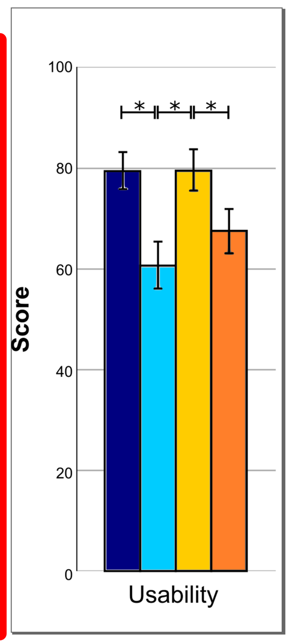


Figure 7. Subjective responses to the System Usability Scale (SUS)

Fig. 6 shows the results of the UEQ questionnaire for each of the methods.

- Gaze is among the highest scoring for all categories and significantly different from Voice in Hedonic Qualities.
- Gesture and Smartphone only have significant differences for Stimulation, Novelty, and Hedonic Quality.

Fig. 8 shows the results of the NASA TLX questionnaire.

- Gaze and Voice had the lowest demand for each of the significant categories
- Gesture and Smartphone typically had the highest demand
- Gesture is significantly more physically demanding than the other methods

Fig. 9 shows the task completion time for each of the methods.

- Gaze and Voice were significantly faster than Gesture and Smartphone.

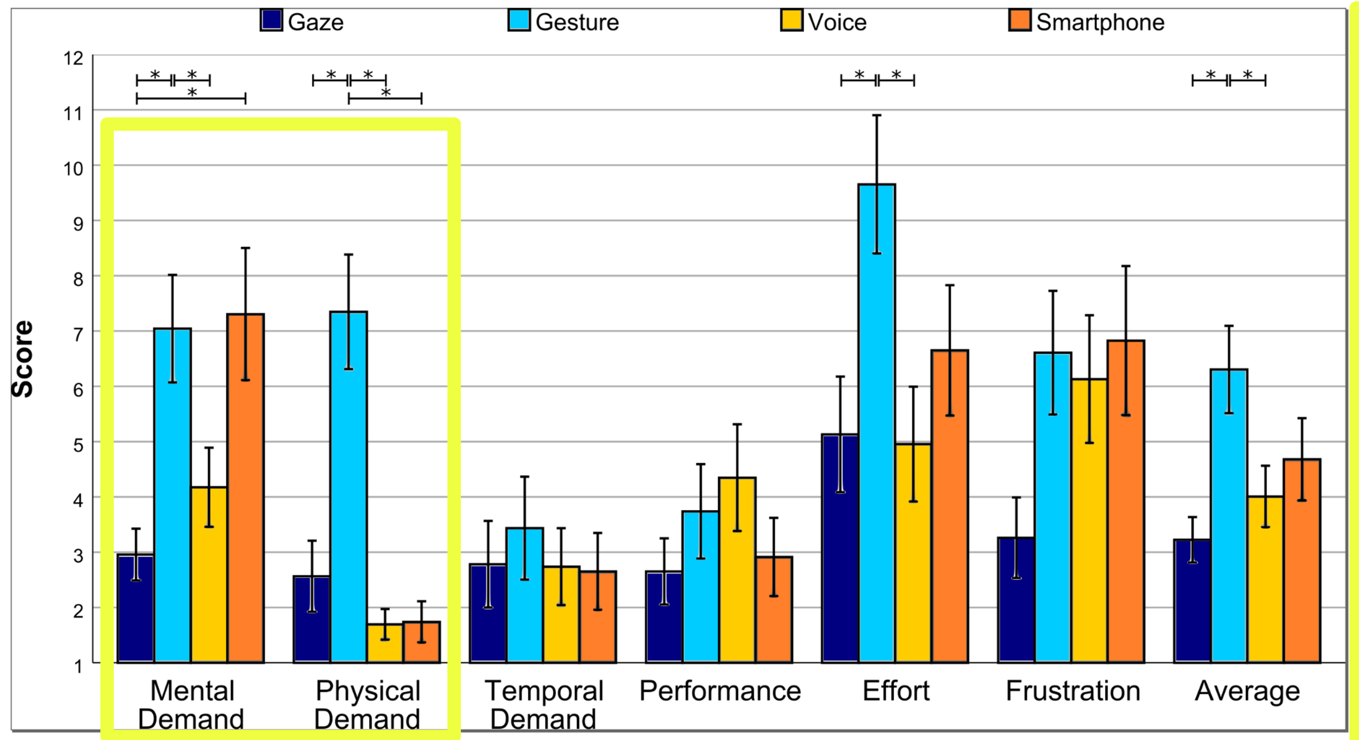


Figure 8. Subjective responses to the NASA Task Load Index (Scores out of 20, lower is better)

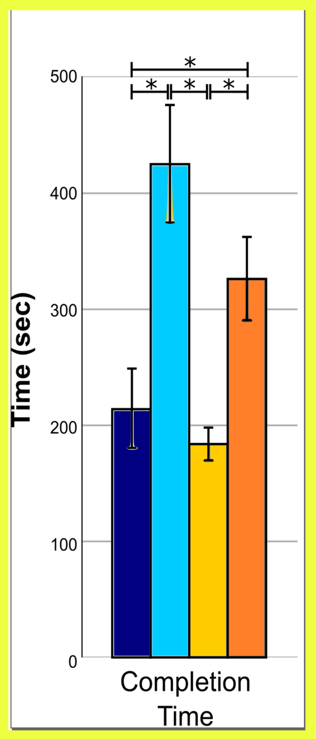


Figure 9. Mean task completion times

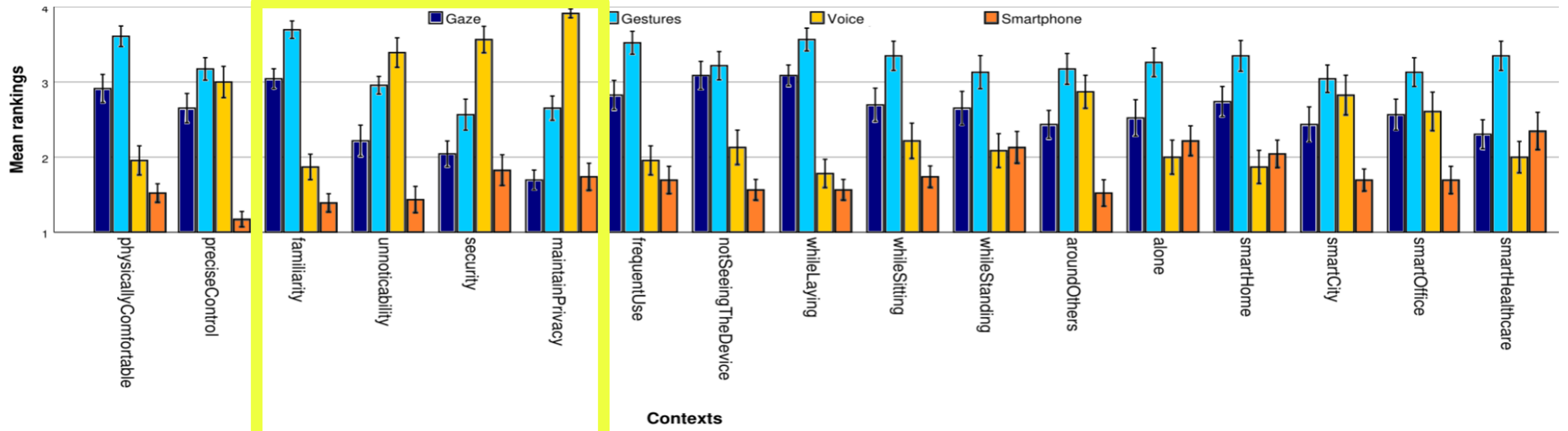


Figure 10. Mean preference rankings for the four conditions in different contexts (lower is better).

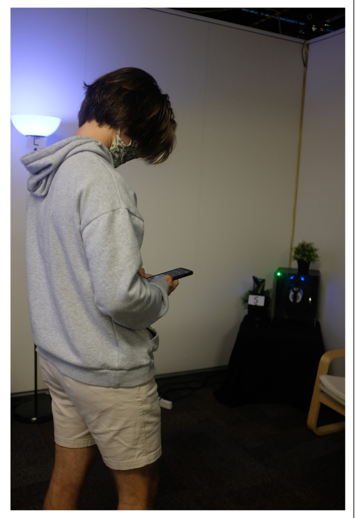


Figure 11. A participant uses the Smartphone method

Fig. 10 shows the rankings for each of the conditions, given different contexts.

- Smartphone ranks among the best in all categories.
- Voice ranks similar to Smartphone, except in contexts relating to privacy, security, or noticeability, where it ranks among the poorest.
- Gesture ranks among the poorest in all categories **except** in contexts relating to privacy, security, or noticeability.

CONCLUSION

Across almost all significant **functional and pragmatic measures**:

- *Participants typically rate, perform, and perceive Gaze and Voice highly and similarly* for overall function, despite some preference for Voice over Gaze in some specific functional contexts.
- Smartphone is typically ranked higher by participants in functional contexts despite lower performance and ratings in some functional categories.

Across **hedonic and aesthetic measures**:

- Gaze and Gesture typically outperform Voice and Smartphone in most of these categories
- This indicates that, regardless of their functional ratings, *the AR methods were viewed as more interesting and enjoyable to use by participants*, particularly when compared to the more traditional methods of Smartphone and Voice to control IoT devices.

Together, the conclusions suggest that:

- **Both AR methods were received positively** as evidenced by the high hedonic and aesthetic scores.
- When considering their functional performance, Gaze performed particularly well, but **Gesture left something to be desired** for participants.
  - Given that Gesture was the most complex and unfamiliar method of the four, the lower functional rating is not without context and further development may help to raise its rating.

Future work may include:

- A focus on natural gestures to draw to control IoT devices,
- Streamlining the design of the AR methods
- Allowing greater control over the device using AR
- Methods to overcome familiarity biases, such as controlled trainings and exploration with devices
- Introducing other AR-based methods of controlling IoT devices.

REFERENCES

[1] Kangsoo Kim; Nahal Norouzi; Dongsik Jo; Gerd Bruder; and Gregory F. Welch. The Augmented Reality Internet of Things: Opportunities of Embodied Interactions in Transreality. In: Nee, A Y C; Ong, S K (Ed.): Handbook of Augmented Reality, pp. 1-60, Springer, 2021.

[2] Nahal Norouzi; Gerd Bruder; Brandon Belna; Stefanie Mutter; Damla Turgut; Greg Welch. A Systematic Review of the Convergence of Augmented Reality, Intelligent Virtual Agents, and the Internet of Things. In: Artificial Intelligence in IoT, pp. 1-37, Springer, 2019.

[3] Kangsoo Kim; Mark Billinghurst; Gerd Bruder; Henry Been-Lirn Duh; Gregory F. Welch. Revisiting Trends in Augmented Reality Research: A Review of the 2nd Decade of ISMAR (2008–2017). In: IEEE Transactions on Visualization and Computer Graphics, 24 (11), pp. 2947-2962, 2018.

ACKNOWLEDGEMENTS

The support for this work was provided by the National Science Foundation REU program under Award No. 1852002. This material also includes work supported in part by the Office of Naval Research under Award Numbers N00014-17-1-2927 and N00014-21-1-2578 (Dr. Peter Squire, Code 34).
Any opinions, findings, and conclusions and recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or the Department of Defense.
Special thanks are offered to all staff and students at SREAL for their expertise, time, and gratuitous support.