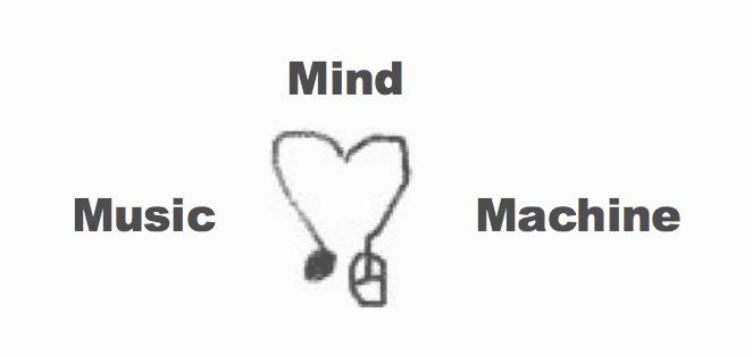
Two or Three Things You Need to Know About AI Design: From Human Factors Perspective



Myounghoon Jeon (Philart)
ISE & CS

MIND MUSIC MACHINE LAB

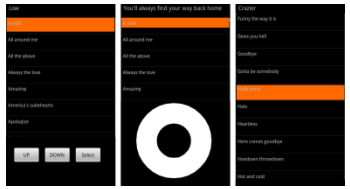
VIRGINIA TECH

My Research Areas

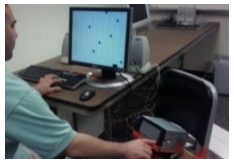
AFFECTIVE COMPUTING

ASSISTIVE TECHNOLOGY

AUDITORY DISPLAYS



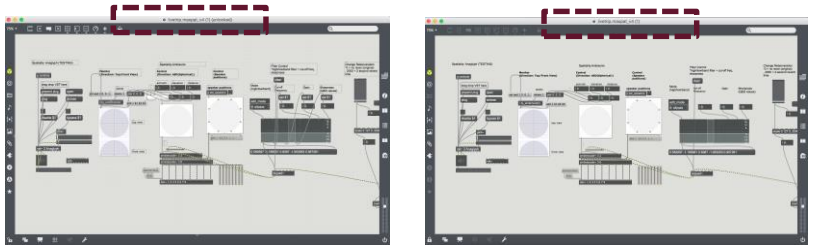
AUTOMOTIVE UI



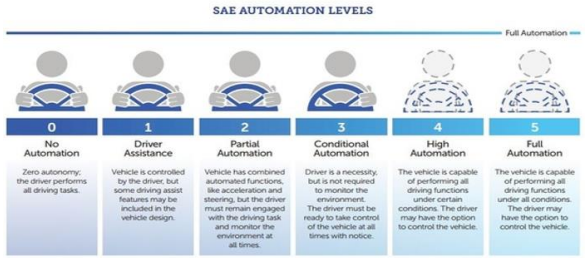
MIND MUSIC MACHINE LAB

1. AI systems that pose problems to the users (1)

- **Mode Confusion**



Unlocked vs. Locked



Automated Vehicles

Only two modes:
- When unlocked, no play: but still confusing

Consequences:
- **No serious problem**.
- Just lock and play again

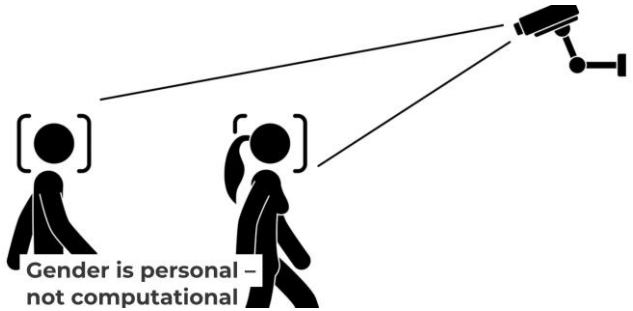
Six modes:
- In four modes, partial collaboration **(dynamic function allocation)**

Consequences:
- Fatal outcomes on road safety and lives

Research Agenda:
- **How to let drivers be aware of the level they're in**
- **How to train them what they're supposed to do in that mode**
- **Or how to reduce the number of modes itself**

1. AI systems that pose problems to the users (2)

- **Bias & Discrimination**



Automatic Gender Recognition

Only two modes:
- Male vs. Female

Consequences:
- Gender reductionism
- AGR Threatens safety of the transgender community and beyond

Research Agenda:
- **Is it necessity?**
- **AGR as a tool for oppression: It's just going to exacerbate what's already there**
- **Inform users if and how they might be gendered and let them opt out**
- **Let users define their own gender/age identity**

* Hamidi, F., Scheuerman, M. K., & Branham, S. M. (2018, April). Gender recognition or gender reductionism? The social implications of embedded gender recognition systems. In *Proceedings of the 2018 chi conference on human factors in computing systems* (pp. 1-13).

2. Design, development, and user testing methods and practices currently adopted in AI (1)

- **Optimization as product**

Facial Affect Detection

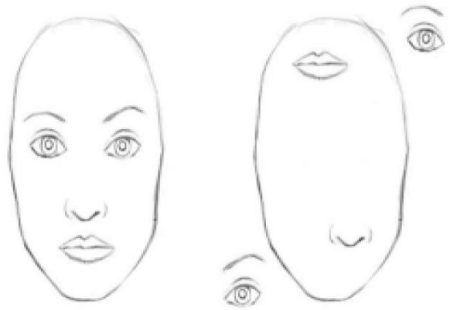


Fig. 9. To a CNN, both pictures are similar, since they both contain similar elements. [29]

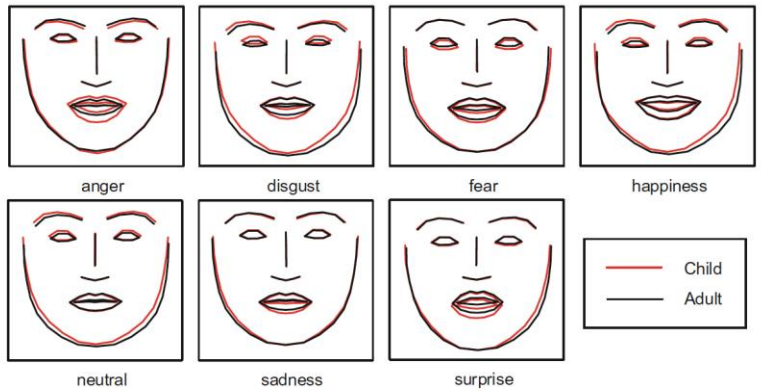


Fig. 2. Average faces of children and adults when making different facial expressions

Practices:
- Use a lot of data points (e.g., 35,000 of face images in FER2013)
- Split the datasets as training data (80%) vs. validation data (20%)
- Use different methods (algorithms) & different datasets
- Use dropouts to avoid overfitting
- Little research done with actual users

Consequences:
- High accuracy but overfitting
- Accuracy drops with validation data and drops even more with real users

Research Agenda:
- **Comparison between algorithms**
- **Use target-appropriate database**
- **Context-dependent user study**
- **Limitations (can't fully estimate emotions based on facial expressions)**
- **Ethical issues**

* Zheng, Z., Li, X., Barnes, J., Park, C. H., & Jeon, M. (2019, July). Facial Expression Recognition for Children: Can Existing Methods Tuned for Adults Be Adopted for Children?. In *International Conference on Human-Computer Interaction*(pp. 201-211). Springer, Cham.

MIND MUSIC MACHINE LAB

2. Design, development, and user testing methods and practices currently adopted in AI (2)

- **Divergence as process**

Dancer sonification

Practices:
- Use small data points (even with one dancer)
- Quickly prototype different alternatives
- Use AI as a tool for design research & creativity
- Research done always with users (or expert)

Consequences:
- Low accuracy but more opportunities
- Impossible to generalize

Research Agenda:
- **How many participants?**
- **How many data points?**
- **If accuracy is not our ultimate goal, how to evaluate our system?**

* Landry, S., & Jeon, M. (2020 March). Interactive sonification strategies for the motion and emotion of dance performances. Journal on Multimodal User Interfaces, 14:167-186.

Thank You! Questions?