introduction

- Donath – *Mediated Faces*
- Cassell & Vilhjalmsson – *Fully Embodied Conversational Avatars*
- Mutlu et al – *Nonverbal Leakage in Robots*
- Powers et al – *Comparing a Computer Agent with a Humanoid Robot*
- Yee et al – *The Unbearable Likeness of Being Digital*

donath - *mediated faces*

- The face is central to real world interactions, and is a rich source of social information
- Our faces are “never not communicating” - structural, dynamic, and decorative facial features continuously offer communicative cues
- As a communication channel, mediated faces can be realistic or “beyond” realistic

why use faces?

- We naturally and intuitively use faces as a rich source of social information (e.g. identity, emotion)
- Facial signals (e.g. gaze) aid conversational process and manage interaction (e.g. indicating attention)
- Faces are tied with politeness norms
- Faces are a major source of impression formation and stereotyping

what the face conveys

- Faces convey information through combinations of structural features, facial movements, and decorations.
- We are hard-wired to use faces to extract both personal (who you are) and social (categories you belong to) identity
- Facial expressions convey emotion, and people guess right reasonably often (e.g. Ekman)
- Gaze indicates attention and understanding; manages turn-taking; expresses intimacy and control; conveys emotion
  - Gaze “pointers” wipe out many of these cues

"bringing the face to the interface"

- Two goals: verisimilitude and "beyond being there"
- Two approaches: video and avatars
- Video
  - Pros: representative; more "subtle and natural"
  - Cons: introduces disruptions (delay, off-axis gaze, visual artifacts)
- Avatars
  - Pros: allows common virtual space; can communicate without explicitly conveying identity
  - Cons: technical issues; can be unintentionally expressive; can erase "communicative competence"

cassell & vilhjalmsson – fully embodied conversational avatars

- Most avatars don't have that communicative competence – they are “incapable of representing the kinds of knowledge that humans have about how to use the body during communication” and users had to explicitly control changes in their state
- Mechanisms for giving avatars “emotion” are a small subset of human nonverbal displays, and are unlikely to be noticed without corresponding nonverbal process cues (e.g. gaze to maintain attention/focus)
- The purpose of BodyChat was to give agents autonomy of communicative expression while maintaining user control over communicative intent.

autonomous behavior

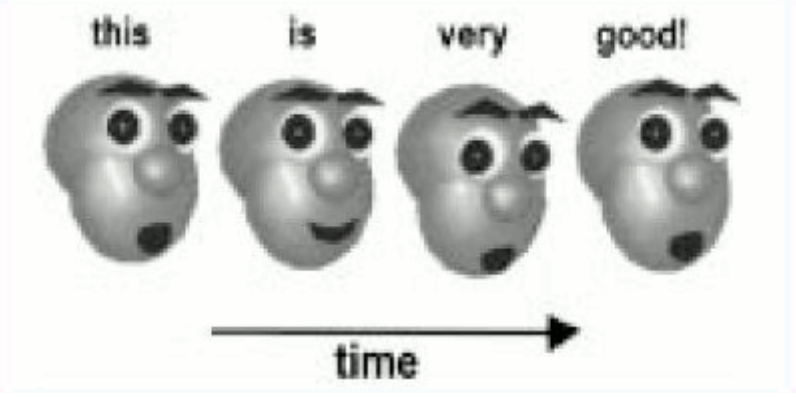
- Human nonverbal behavior can't easily be translated to avatar because they inhabit radically different spaces
- Direct manual control disrupts conversational flow
- Limited avatar autonomy solves several problems:
  - no micromanagement
  - avatar handles spontaneous reactions (e.g. backchannels)
  - avatar control over its motions makes them flow more naturally – no delay as user navigates menu/issues commands
  - maps communicative intent appropriate from real to virtual space

communication behaviors to emulate

- Avatars should “effortlessly” use hand gestures, posture/stance, orientation, facial expression, and gaze
- Non-verbal communication behavior should be multimodal
  - just one channel at a time is unnatural
- Should replicate communication phenomena:
  - Turn taking - Indicating attention
  - Marking syntactic events (e.g. questions)
  - Indicating understanding

BodyChat

- Conversational phenomena (e.g. greeting) consists of a set of communicative behaviors (e.g. smiling, orienting towards other)
- User issues high-level intention parameters (e.g. enter conversation) through text prefixes, text detection (e.g. ! = emphasis), or control panel
- Avatar reacts socially and autonomously to events in the virtual environment, both external and user-driven

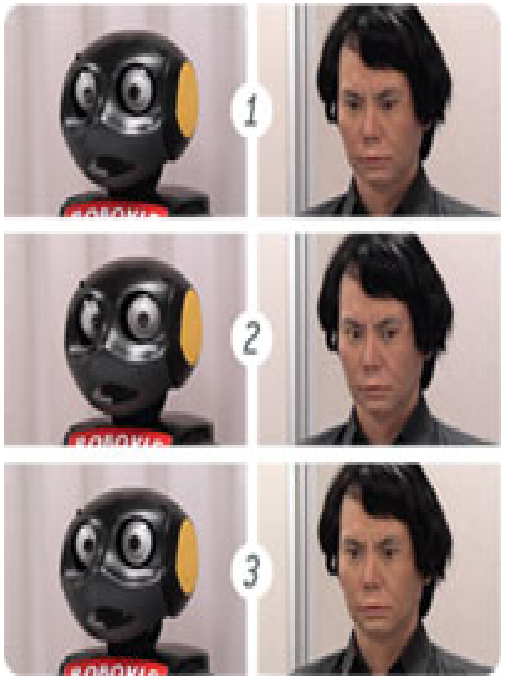


evaluation

- Authors compared people's interactions with autonomous, manual, autonomous + manual, and no non-verbal communicative behaviors
- Autonomous condition remembered more facts about interlocutors, and interactions lasted longer. Authors suggest autonomous condition led to more focus on conversation and less on process.
- Users judged the autonomous condition as more natural and more under user's control than other conditions.
- Autonomous condition was viewed as more natural but not more expressive than none condition

mutlu et al – *nonverbal leakage cues*

- Do people detect nonverbal “leakage” - unintentionally produced cues that reveal internal emotional and cognitive processes – in robots like they do in humans?
- If so, do people attribute mental states to robots?
- This experiment focuses on whether people can read gaze-based leakage cues and how people attribute intention in a stylized anthropomorphic robot (Robovie) and a highly humanlike robot (Geminoid)



design

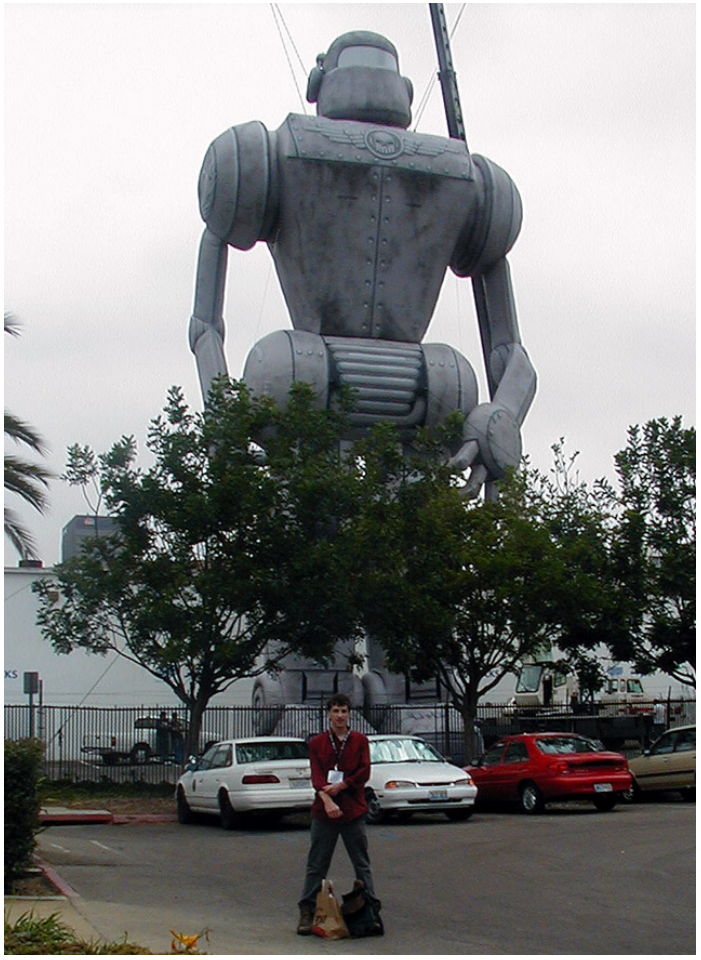
- 2x2 design
  - Gaze cue or no gaze cue
  - Geminoid vs. Robovie
- Post-test showed that people can identify robots' as well as human gaze targets at roughly similar rates, and all well above random chance.
- Experimental procedure

results

- Participants took less time and asked fewer questions with a gaze cue, but this effect was driven by Geminoid
- Participants recognized more gaze cues from Robovie, but more often attributed intention to Geminoid's gaze cues.
- Pet owners too less time and asked fewer questions (the authors imply that they are used to attributing mental states to non-humans)
- Robovie was rated as more socially desirable (mostly by women, who understandably found Geminoid creepy)
- Geminoid was viewed as more socially demanding – it is not clear whether this was an Uncanny Valley effect or because it was human enough for politeness norms to kick in

powers et al – *comparing a computer agent with a humanoid robot*

- How do people's responses to robots differ from responses to agents?
- Agents can have a “surface resemblance” to people and can be programmed with lifelike movement, but don't embody physical space
- Robots exist in physical space and operate under real-world physics
- The authors suggest that social qualities of robots and agents may vary with realism, physical proximity, sense of presence, and size



design

- Compared co-located robot, remote video-mediated robot, and an agent modeled off of the robot on both small and life-size displays
- All were representations of the NurseBot and gave a health habits survey
- They measured participants' engagement with the robots and agents, disclosure of sensitive information, social influence, and conversational memory
- They also looked at participants' mental states and subjective attitudes towards the robots

results

- Participants were more engaged (spent more time) with robots than agents
- Participants disclosed more to agents than to robots, and remembered more of the conversation
- Participants liked interacting with robots (particularly the collocated robot) more than the agents, and rated the robots more as more socially favorable (trustworthy, responsive, etc.)
- The robots were viewed as more lifelike than agents
  - This points to a possible confound. The robots had bodies but the agents were just floating heads.
- Few differences between large-display and small-display agent

yee et al – *the unbearable likeness of being digital*

- One of the study's goals was to demonstrate that behavior in virtual world is a serviceable proxy for typical behaviors in the real world, a prerequisite for harnessing virtual environments as social science testbeds
- They decided to look at how two fundamental non-verbal communication processes, proxemics and gaze, compare in real and virtual world

design

- Longitudinal study of dyadic social interaction patterns in Second Life (excluding certain areas with “activity-specific positional configurations”)
- Captured avatar gender pairings, interpersonal distance, gaze angles, and “talking” state from in-game data sampling
- Used this data to look at mutual gaze by gender pairing and interpersonal distance by gender pairing interactions

results

- Findings support argument that social interaction in virtual environments follows many of the same norms as social interaction in the physical world
  - Mixed-gender pairs stood closer than M-M or F-F pairs
  - The closer avatars were, the less likely they were to maintain eye contact
  - M-M pairs were significantly less likely to make eye contact than mixed or F-F pairs, particularly indoors when space was limited

more fun with robots!

30 Rock Explains The Uncanny Valley

KeepOn

Beck - “Hell Yes”

Japanese Fashion Robot

Robot City Workshop (Lakeview, Chicago)

Charting the Uncanny Valley