

# Abuses in Building and Using ECA Systems

Catherine Pelachaud,

LINC - University of Paris 8



## Media Equation

- ◆ People tend to treat socially intelligent computers/agents/robots as human (Reeves and Nass, 1996)
  - act with the entity as if they are human: trust them, believe them
- ◆ But people treat them also as machine:
  - turn them off (Bartneck, 2006)

## Believable Agents

- ◆ Definition by Loyall 1997: « *a character is considered to be believable if it allows the audience to suspend their disbelief* »

From Jarmo Laaksolahti & Per Persson, 2001

- ◆ believable characters:  
movie/theater/book. Subjects know
  - they are spectators of the life of another one
  - they are able to make a separation between real and fictive life

## Believable Agents

- ◆ believable agent: interactive system, autonomous entity
  - may act directly on subject's life (eg sell him an object)
  - affords trust, emotion response, comprehension etc since
    - she is responsive to user's concern
    - she interacts with him
    - she shows interest
    - ...

## Applications

- ◆ Evaluation studies have shown that ECAs can benefit the interaction, user's experience, user's outcome, etc
  - Tutor
  - Companion (long term companion)
  - Health-care
  - Language teacher
  - Language translator (Sign Language, Cued Speech)
  - Do repetitive, dangerous tasks
 Even when they are ironic about user's web choices  
 (Agneta&Frida (Höök et al, 2003))

## Model of Human-like Qualities

- ◆ Agent looks and behaves like humans:
  - appearance
  - communicative, social and emotional behaviors
  - perception
  - plannification
  - action
  - speech

## Attribution of Human-like Qualities

- ◆ Attribution of human-like qualities (Okita & Schwartz, 2006)
  - Young children (around 3 years old) attributes animalistic properties to entertainment robots
  - growing up they learn to distinguish peacemeal technology and real animal
- ◆ Social and emotional intelligence

## Non Human-like Qualities

- ◆ « They don't get hungry, they're not afraid. They don't forget their orders. They don't care if the guy next to them has just been shot. Will they do a better job than humans? Yes. » Gordon Johnson of the Joint Forces Command at the Pentagon

## ECA design

- ◆ In ECA design several entities intervene:
  - designer and developer: plan ECA's behavior and reaction to user's action; misuse of ECA
  - user: abuse of ECA
  - ECA: take advantage of user, influence user, create bond with user, abuse trustworthiness of user
  - application: provide information, e-selling, act for the user

## Abuse from developer/designer

- ◆ Different degrees of autonomy
- ◆ "The lawyers tell me there are no prohibitions against robots making life-or-death decisions," said Mr. Johnson, who leads robotics efforts at the Joint Forces Command research center in Suffolk, Va. "I have been asked what happens if the robot destroys a school bus rather than a tank parked nearby. We will not entrust a robot with that decision until we are confident they can make it. » NY times, 16 Feb 05

## Abuse from application

- ◆ influence user against his will/awareness
- ◆ take/don't take action against user's will
- ◆ manipulate/influence/control user's believes/emotions/actions

## Abuse on user from ECA

- ◆ pro-active agent: problem of privacy and security
- ◆ powerful mean of persuasion
- ◆ source of control and manipulation

## Abuse on ECA from user

- ◆ Misuse and Abuse of Interactive Technologies Workshops (2005, 2006)
- ◆ Reasons of abuse (de Angeli 2006)
  - people feel they can get away with their action
  - people may feel frustrated when interacting with device, software...

## Abuse on ECA from user

- ◆ User abuse computers through
  - verbal abuse, injury, sexual harassment (De Angeli 06, Brahnam 06)
  - physical abuse (slam the computer, break it) (Bartneck 06)
- ◆ Gender of ECAs (Brahnam, 2006) :
  - little effect on insult frequency
  - but sexual comments greatly increased with female embodiment

## Abuse on ECA from user

- ◆ Closure of a dialog system between a user and an agent (Adapt system (Bell&Gustafson, 2000))
  - non-courtesy expression: 35%
  - farewell: 19%
  - courtesy expression: 4%
  - thanks: 3%

## Abuse on ECA from user

- ◆ Replica of Milgram's study with robot (Bartneck, 2005) :
  - 40% of subjects administered deadly voltage to human in Milgram's study
  - 100% of subjects when to deadly voltage for robot. Plus subjects went to much higher voltage for robot than for human
- ◆ BUT in peer-to-peer contexts:
  - abusive behaviors from user are rare (Krenn & Gstein, 2005)

## Some answers against user's abuse

- ◆ How should the machine/ECA answers and reacts to such user's actions? (Branham 05)
  - be aggressive; counter-attack
  - disconnect user/leave environment → punishment of the user's behavior
  - be compliant (play the victim)
  - change subject of conversation
  - repeat the abuse: best solution as user gets bored.

## Media Equation revisited

- ◆ People have less concern to abuse machine than other humans (Bartneck, 05)
- ◆ User build long term relationship with ECA, create bonds with them, ...
- ◆ BUT attention: should not replace relationship and social bond with real people

## Some sort of conclusion

- ◆ Educate people and children of the risk of being manipulated, make them aware of the danger (Bickmore & Picard, 05)
- ◆ Add some forms of monitoring (Creed & Beale, 2006)
- ◆ Add some governing rules for agent design
- ◆ Ensure that at the end of an interaction with an agent, users know they are communicating with virtual entities (breaking the 'believability')

## “It Worked for me”

- ◆ Long term relation with social agent as a virtual trainer
- ◆ « She is a computer character. I don't know if she cared about me. I don't know if she feels. She's a character and has a role, but I don't know if she has feeling. But it worked for me and I'm happy ». reported in (Bickmore and Picard, 2005)