How emotional should the iCat robot be?

A children's evaluation of multimodal emotional expressions of the iCat robot

TNO | Knowledge for business

Melanie Kroes

Outline

- Research question
- Background
  - iCat
  - Emotions
- Hypotheses
- Methods
- Results
- Discussion

Research Question

To what extent do emotions in facial and vocal expressions improve the iCat's effectiveness in convincing children to show certain behavior?

Conditions:

- Neutral iCat (*neutral*)
- iCat with emotional facial expressions (*face only*)
- iCat with emotional facial and vocal expressions (*speech and face*)

iCat's roles (and children's tasks)

- Electronic pet (emotion recognition)
- Educator (play a quiz)
- Motivator (doing exercises)

The iCat robot

Features

- 13 servo motors to control:
  Eyes, eyebrows, eyelids, lips, head and body
- Lights and touch sensors in both ears and paws
- Speaker
- Webcam

Programming

- OPPR (Open Platform for Personal Robotics™)
- C++
- LUA

Emotions (1) Emotions in synthetic speech

Speech synthesis techniques

- Formant synthesis
- Diphone synthesis
- Unit selection

Emotion Modeling

- Modification of parameters →
- Emotional speech units
- Copy synthesis
- Voice conversion

	Anger	Happiness	Sadness	Fear	Disgust
Speech rate	slightly faster	faster or slower	slightly slower	much faster	very much slower
Pitch average	very much higher	much higher	slightly lower	very much higher	very much lower
Pitch range	much wider	much wider	slightly narrower	much wider	slightly wider
Intensity	higher	higher	lower	normal	lower
Voice quality	breathy, chest tone	breathy, blaring	resonant	irregular voicing	grumbled, chest tone
Pitch changes	abrupt, on stressed syllables	smooth, upward inflections	downward inflections	normal	wide, downward terminal inflections
Articulation	tense	normal	slurring	precise	normal

Emotions (2) Emotions in facial expressions

- Emotion in the human face:

	Facial action						
	Eyebrow frown	Raise eye-brows	Raise upper eyelid	Raise lower eyelid	Lip corners	Open mouth	Raise upper lip
Emotion							
Happiness				X	Raise	X	
Surprise		X	X			X	
Anger	X		X	X			
Disgust	X			X			X
Fear	X	X	X			X	
Sadness	X	X			Lower		

- How to transform these features into parameters for synthetic faces?

Hypotheses

Research question

To what extent do emotions in facial and vocal expressions improve the iCat's effectiveness in convincing children to show certain behavior?

Hypotheses

- The more emotional, the better the iCat convinces the children to show certain behavior
- The more emotional, the more children will like the robot

Methods (1) **Tasks**

- Three iCat roles:
  - **iCat as electronic pet:** a scenario is told and the iCat says something now and then. Children have to report the emotion of the iCat while saying these things.
  - **iCat as educator:** the iCat explains the rules of the sport korfball and afterwards asks questions about them. The children have to answer these questions
  - **iCat as motivator:** the iCat tries to motivate children to exercise. Two tasks have been done: searching blue marbles form a large basket and running as fast as possible for two minutes (the amount of steps is registered)

Methods (2) **Design**

- Each participant:
  - evaluated two (out of three) conditions
  - performed all tasks with both iCat robots (the order of the tasks was balanced across participants)

- There are 3 possible combinations of conditions and 6 (3!) possible orders to present the tasks. So, 3 x 6 = 18 participants were needed.

- As each participant evaluated two iCats,12 data points were collected in each condition for each task.

Methods (3) **Participants**

- 20 children of eight or nine years old from a primary school in Soesterberg (2 removed, because of missing data)
- The experiment was done at TNO, in the so called ‘living room’
- Parents were also welcome and could see the child on a video screen during the experiment

- Children received a cheque of €10,- and a picture of themselves with the iCat robots.

Methods (4) **Data**

Emotions

- The emotions anger, fear, happiness and sadness have been used, as these are basic emotions and can be modified both in the speech and facial expressions.

Speech

- Neutral speech synthesized with Fluency (voice Diana)
- Emotional parameters modified with EmoFilt

Facial Expressions

- Standard settings from Philips for sadness and fear
- Modified emotions for happiness and anger (because of better recognition in the pilot study)

Methods (5) **Materials**

User:

- Two iCat robots
- 2 pc's and one laptop (no monitors)
- Webcam
- 2 video camera's

Experimenter:

- One pc with two monitors
- Video screen

Methods (6) **Questionnaires**

Questionnaire A: general questions at the beginning

Questionnaire B (after each dialog with the iCat):

1. Fun with robot
2. Intelligibility
3. Fun with task
4. Difficulty with task
5. Want to do the task again?

Questionnaire C (after each task):

1. Difference
2. Preference

Questionnaire D (at the end):

1. Fun
2. Empathy
3. Trust
4. Overall preference

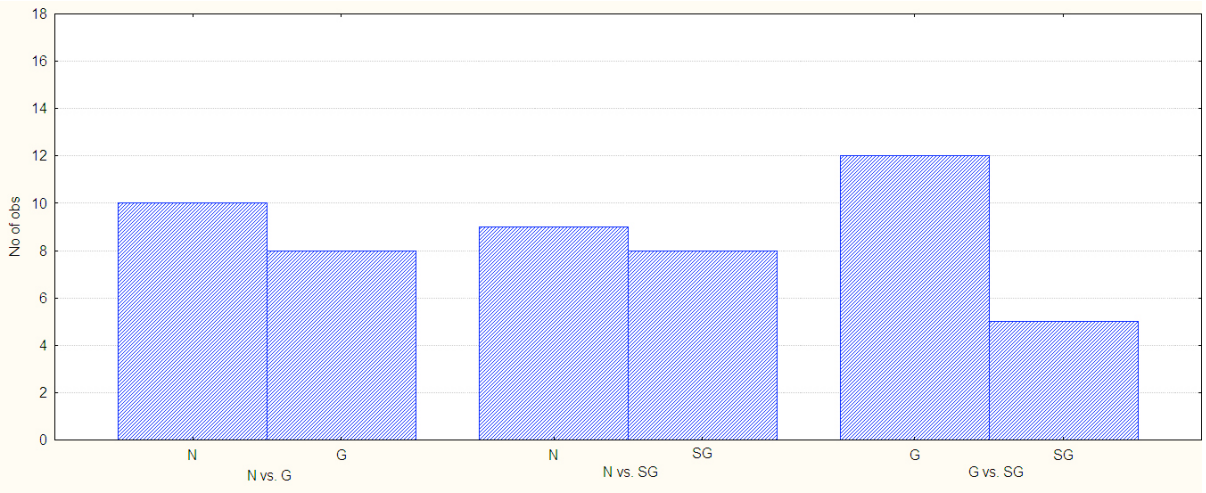
Results (1) **Questionnaire B**

- In general, emotional speech was less intelligible than neutral speech.
  - This effect was the greatest with the educator and the smallest with the electronic pet
- The quiz was judged the most difficult of the three tasks

Results (2) **Questionnaire C:**

- On the question ‘did you notice a difference between the two iCats with this task?’ the answers ‘yes’ and ‘no’ were given equally often.

- Only in the *face only* versus *speech and face* conditions, there was a trend that *face only* was chosen more often.



Results (3) **Questionnaire D**

- On the questions about ‘fun’, ‘empathy’ and ‘trust’ no significant differences were found as all iCat conditions scored very high on these questions.

	fun		empahty		trust	
	mean	std.dev.	mean	std.dev.	mean	std.dev.
neutral	4,67	0,471	5,00	0,000	1,92	0,862
face only	4,50	0,764	4,75	0,433	2,00	0,816
speech and face	4,75	0,433	4,75	0,595	2,00	0,707

  - Children like to work with the iCat as it is new for them so even the neutral iCat is judged very well.

Results (4) **iCat as electronic pet**

- Emotions were very well recognized, even in the neutral condition:
  - *neutral* 78,1 % correct
  - *face only* 88,4 % correct
  - *speech and face* 84,4 % correct

		Perceived				
		Happiness	Fear	Anger	Sadness	Neutral
Intended	Happiness	**79.2**	0	0	0	20.8
	Fear	0	**81.9**	5.6	5.6	6.9
	Anger	0	2.8	**84.7**	4.2	8.3
	Sadness	0	0	5.6	**88.7**	5.6

- Reaction times do not differ across conditions, but are longer for incorrect answers than for correct answers.

Results (5) **iCat as educator**

- 69,4 % of the questions were answered correctly.
  - *neutral* 75,0 % correct
  - *face only* 66,7 % correct
  - *speech and face* 66,7 % correct

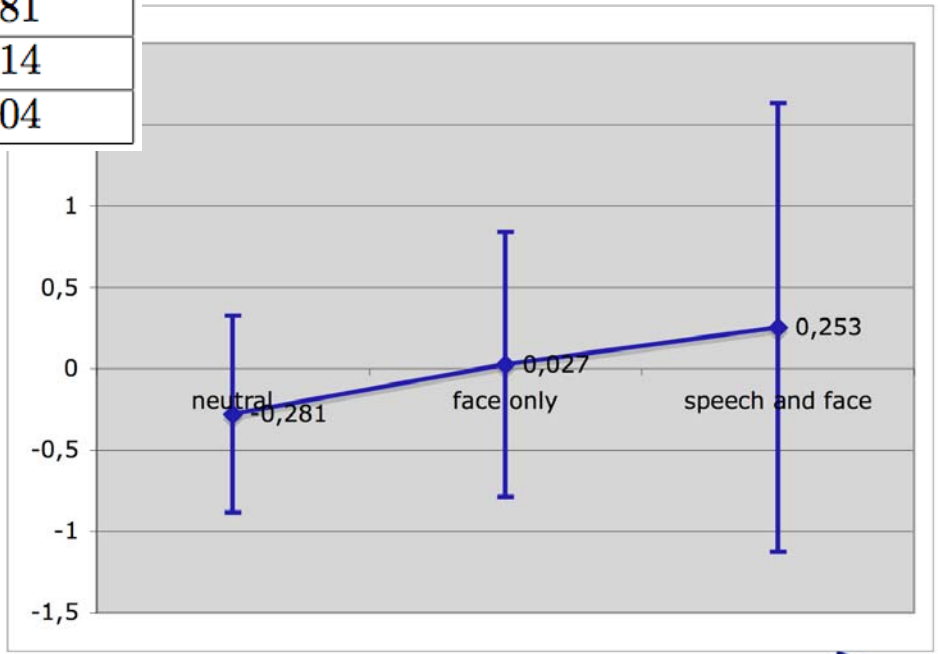
- Reaction times were shorter with the *speech and face* condition than in the other two conditions.
  - *neutral* 5.94 sec (std.dev. = 1.87 sec)
  - *face only* 5.88 sec (std.dev. = 1.67 sec)
  - *speech and face* 4.69 sec (std.dev. = 1.54 sec)

Results (6) **iCat as motivator**

- Mean amount of marbles collected: 71.8 (std.dev. = 12.24)
- Mean amount of steps taken: 458.3 (std.dev = 131.8)
- z-scored in order to analyze these results together

	Valid N	Mean	Std.Dev.
all	36	0	1
speech and face	12	0.253	1.381
face only	12	0.027	0.814
neutral	12	-0.281	0.604

- Results lie in the expected direction, i.e. *neutral* < *face only* < *speech and face*, but were not significant



Discussion

- Intelligibility of the emotional speech should be improved, because emotions are useful:
  - With the quiz, children answered faster with the emotional iCat
  - During the exercises, more marbles have been found and steps have been taken with the emotional iCat

- Suggestions for further research:
  - Test more children (more than one session)
  - Other tasks (maybe more difficult or less fun)
  - More differences between the conditions

Questions?