On the recording and measurements of social robotics experiments in education

M. Graña

Computational Intelligence Group

University of the Basque Country

Index

- Short overview of CybSPEED project
- Classes of observation measures:
  - Remote sensing
  - Contact sensing
  - Indirect observations
- Some examples from the literature
- Conclusions and discussion

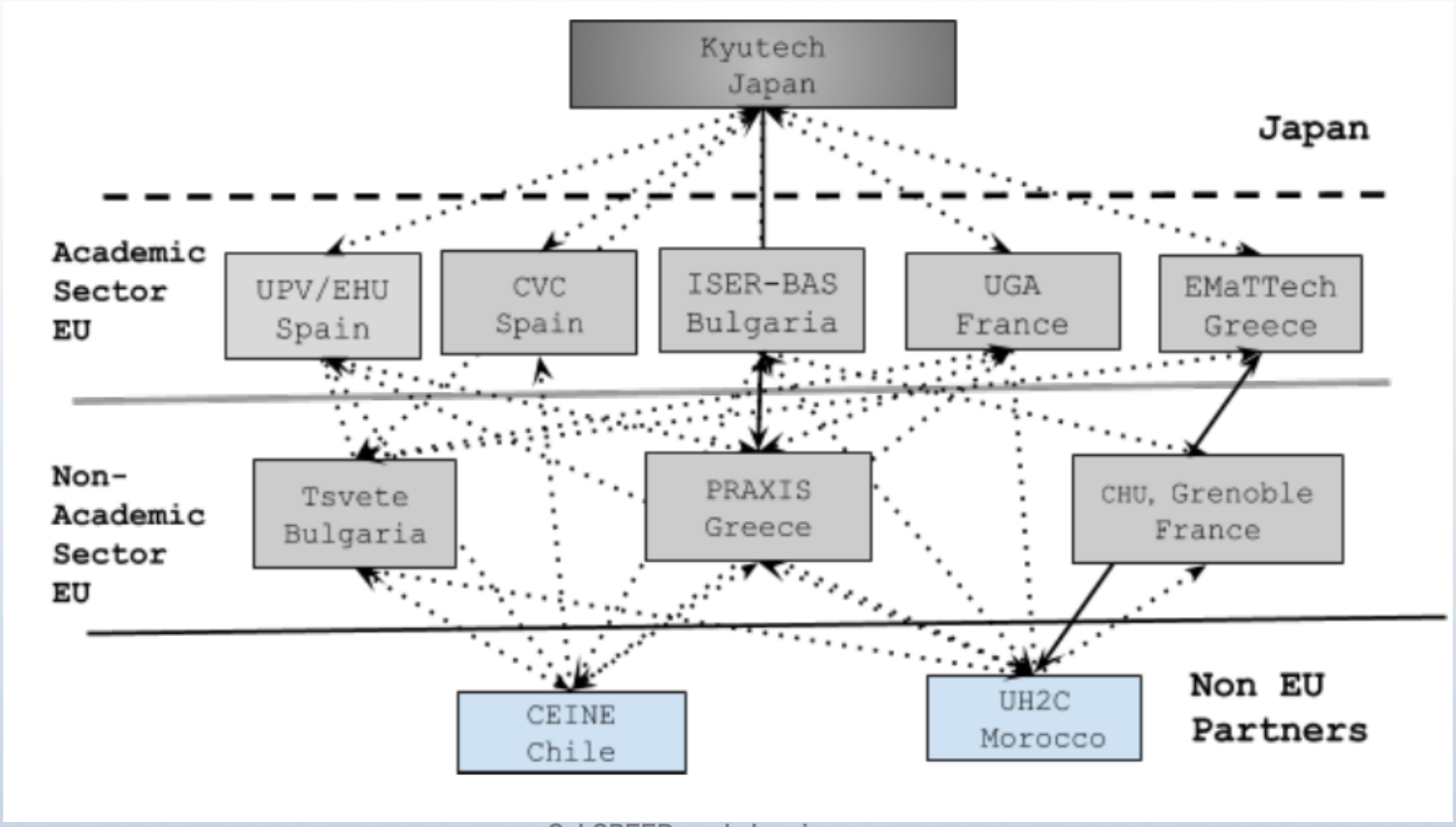
CybSPEED

Project proposal

CybSPEED proposal

- Research on three levels
  - analysis of cognitive biometrics signals,
  -
  - modeling of the learner-robot interaction and
  - development of novel instruments
    - towards an optimal design of Cyber-Physical Systems
    - for improved pedagogical rehabilitation in education

Consortium


Kyutech
Japan
Japan
Academic
Sector
EU
UPV/EHU
Spain
CVC
Spain
ISER-BAS
Bulgaria
UGA
France
EMaTTech
Greece
Non-
Academic
Sector
EU
Tsvete
Bulgaria
PRAXIS
Greece
CHU, Grenoble
France
CEINE
Chile
UH2C
Morocco
Non EU
Partners

Topology of the project

Artificial perception

User awareness

CVC

UCH

UPV/EHU

Robot soft-adaptation

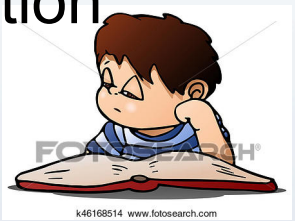
IR-BAS

Advanced human-robot interfaces

Games and storytelling

Tsevete

Student progression measurement



Kyutech

PRAXIS

EMaTTech

UH2C

Behavior measurement

Neural activity correlation with behavior

UGA

CHU

Legal framework

Ethics requirements

- There is an increased concern about ethics in research related to humans
- Screening is increasingly stringent requiring
  - Clear hypothesis formulation
  - Trade-off between intervention and benefits
  - Protection of the subject well-being
  - For children with special needs safety and absence of negative effects
  - Strict experimental control: when to stop it.

Ethics requirements

- Ethics refer mostly to experiment definition quality regarding
  - intervention carried out
  - Experimental control
    - When it will be stopped
  - Information given to the experimental subjects
    - Informed consent
    - Withdrawal from the experiment

Data protection

- The EU has made effective a new regulation that increases the control of people over its personal data.
- For the experimental research point of view it requires stronger controls on the data processing quality
- Data must be anonymized for processing
- Subjects have complete control on when and how the data will be deleted

Data protection

- Institutions require now a data protection officer that ensures correct data management policies and corrections.

Scientific requirements

- From the scientific point of view we need
  - Exhaustive information
  - Repeatable experiments
  - Instrument calibration experiments
  - Precise and complete experimental designs
    - Identification of causes (interventions) and expected effects (measures)
    - Good definition of observable measures
      - With little error risk
      - As complete as possible
      - Good statistical properties

Trade off

- There is an intrinsic conflict between scientific and ethical/data protection interests
- The definition of good observation measures is at the core of the problem
  - They must effectively quantify the expected effect
    - Minimally invasive and harmless
    - They must not interfere with the measured process
  - They must minimize data protection and ethical issues
  - They must be reliably obtained

Kinds of observation measures

Remote sensing

- Sensors that do not touch the subject
  - Cameras
    - Visible spectrum
    - Infrared spectrum
  - depth sensors (i.e. kinect),
  - position sensors (i.e. wireless localization)
- Analysis can be carried out by human observer or automatically

Remote sensors

Pros	Cons
•Great quantity of information •Allow repeatability of analysis •New measures can be applied much later than the actual experiment •They provide quite intuitive demonstration of effects of intervention •They provide quantitative information	•Usually are data protection critical •People can be easily identified •Images are very sensitive to evolving ethical standards

Wearable sensors

- Motion sensors
- Physiological sensors
  - EEG
  - MEG
  - Body physiology (heart ratio, sweat and conductivity, others)

Wearable sensors

Pros	Cons
•Provide quantitative information •They do not raise data protection issues	•They are difficult to interpret in terms of behavioral units •Their signals are very noisy •There can be strong inter-subject variabilities •Their statistical analysis is difficult

Indirect measures

- Questionaires
  - The subject
  - The experiment controller/observer
- Observations made during the experiment

Indirect

Pros	Cons
•They do not pose big data protection issues •They are the last resort to gather information in very unstructured conditions	•They are very subjective •Not easily quantifiable (Likert scales)

Some examples

Adaptive Emotional Expression in Robot-Child Interaction

Myrthe Tielman
TNO / Utrecht University
m.l.tielman@gmail.com

Mark Neerincx
TNO / Delft University of Technology
mark.neerincx@tno.nl

John-Jules Meyer
Utrecht University
J.J.C.Meyer@uu.nl

Rosemarijn Looije
TNO Human Factors
rosemarijn.looije@tno.nl

2014

HRI'14, March 3–6, 2014, Bielefeld, Germany.

ACM 978-1-4503-2658-2/14/03 ...$15.00.
http://dx.doi.org/10.1145/2559636.2559663.

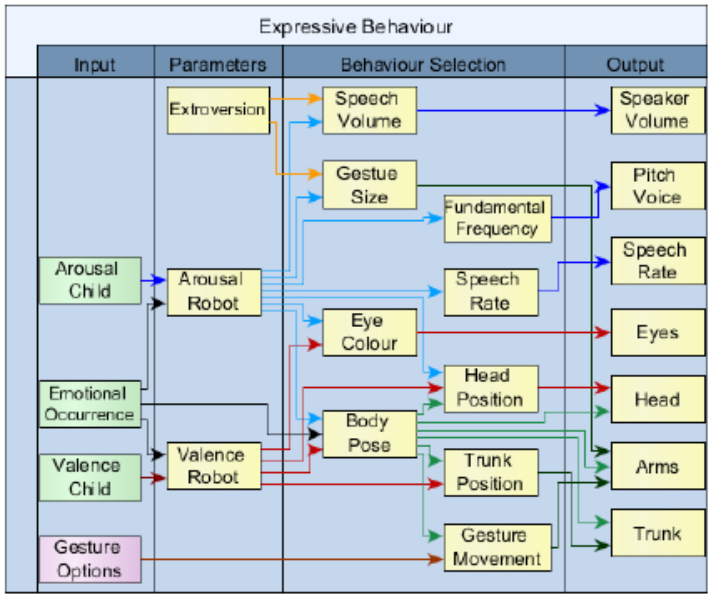


Figure 1: Model for expressive behaviour of the Nao robot

Figure 2: A child playing the quiz with the robot using the tablet and seesaw

Task: quiz answering in alternating roles using a tablet

Table 1: Expressions and their definitions

Expression	Properties
Smiles	All instances where the mouth of the child angles upwards. As we only count instances and not duration, this was only counted when there was a change. So only when the mouth angles rose upwards.
Laughter	All cases in which the child laughed. Laughter is here classified as those smiles which are accompanied by sound or movement of the chest related to the happy feelings.
Excited bouncing	All cases in which the child either bounced up and down out of obvious excitement, or in which the child made a large excited gesture. An example of the latter is raising both arms, and other such gestures of success.
Positive vocalization	Every positive exclamation not directly related to the dialogue. Common words are *yay* or *yes*.
Frowns	All facial expressions obviously related to thinking, concentrating or misunderstanding. Also all facial expressions where the eyebrows are lowered.
Shrugging & Sighing	Raising the shoulders and dropping them again, or audibly letting out air. These two expressions are seen as signs of boredom
Startle	All signs of involuntary fright from the child, such as it being startled by sudden movement.
Negative vocalization	All negative exclamations not directly related to the dialogue, such as *nou zeg* or *jammer*.

Goal: to assess wether emotional reactions by the robot increase social bonding

Measurement of child reaction: via video labeling by a human using a scoring system

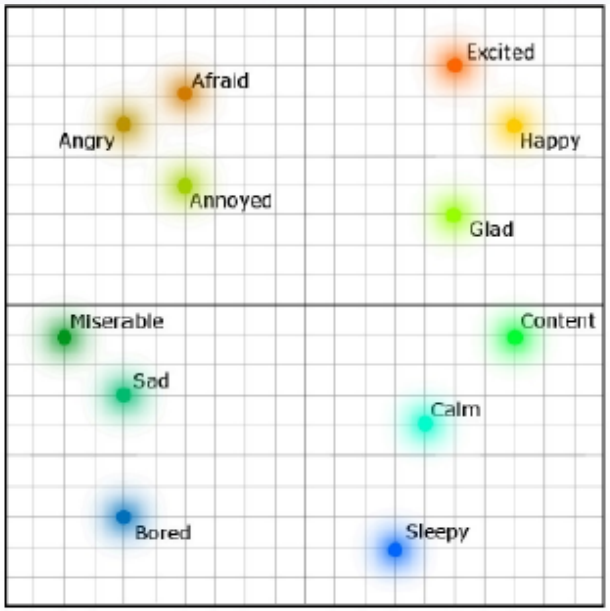


Figure 3: The interface via which the experimenter provided information about the arousal and valence of the child. The horizontal axis represents the valence of the child, the vertical axis the arousal. The coloured dots reference specific emotions as context.

Emotional feedback to the robot about the child emotional state
Carried out by a human

Table 2: Topics of questions in questionnaires

Subject	Nr. of questions individual robot	Nr. of questions forced choice
Fun	9	1
Acceptance	3	1
Empathy	3	1
Trust	3	1
Emotions	3	1
Preference	0	1

Questionary to the child about the robot preferences (with vs. without emotion) x

2017 26th IEEE International Symposium on
Robot and Human Interactive Communication (RO-MAN)
Lisbon, Portugal, Aug 28 - Sept 1, 2017.

Functional Imitation Task in the Context of Robot-assisted Autism Spectrum Disorder Diagnostics: Preliminary Investigations

Frano Petric[1] and Damjan Miklić[1] and Maja Cepanec[2] and Petra Cvitanović[2] and Zdenko Kovačić[1]

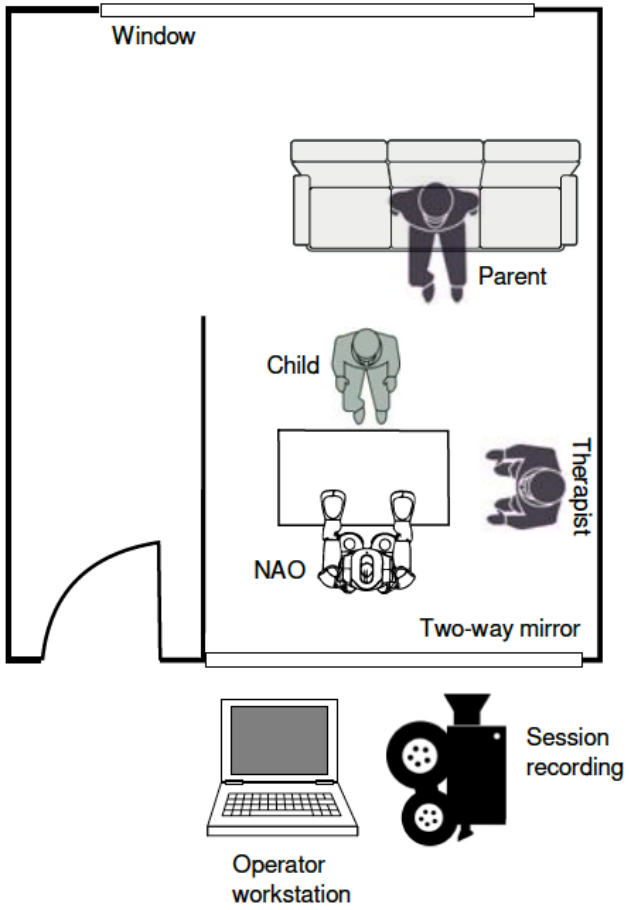


Fig. 1: The layout of the room where the imitation experiments were taking place.

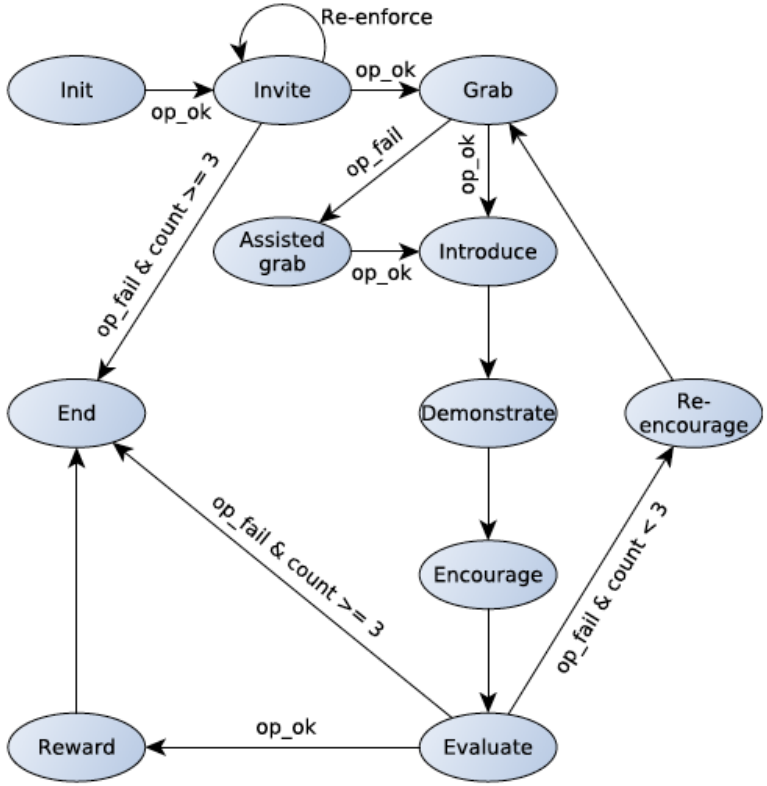


Fig. 2: A state machine model of the imitation task. Transition conditions labelled with `op_*` indicate that robot operator confirmation is required for completing the transition. Unlabelled transitions imply successful completion of the task associated with the source state.

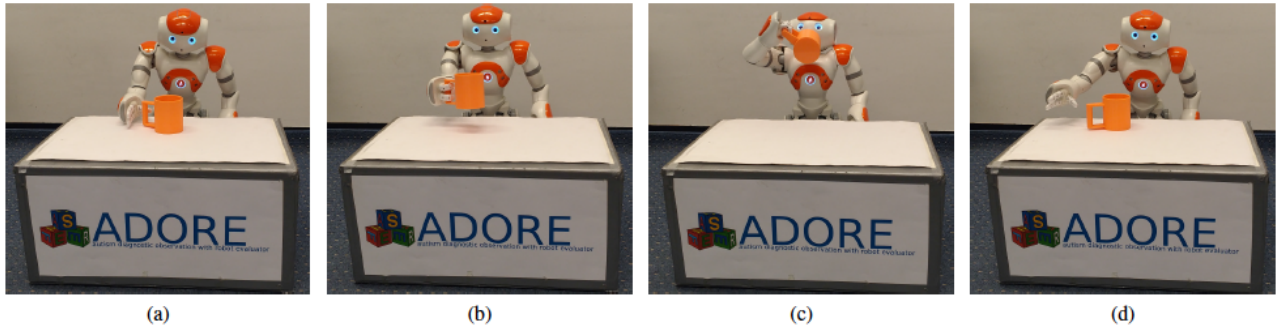


Fig. 5: The drinking and frog gesture demonstration sequences.

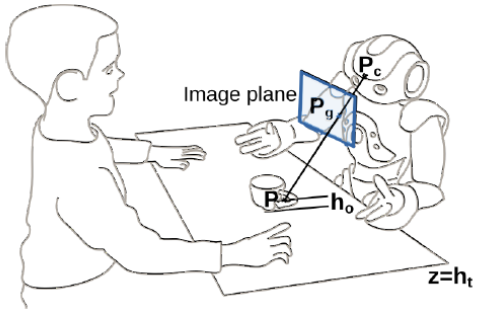


Fig. 4: Grab point calculation using calibrated camera and known height of the grab point with respect to NAO's base coordinate frame.

Measuring the success of the imitation game

Automated child gesture recogntion
External camera
HMM based gesture recognizer

TABLE III: Laboratory evaluation of gesture recogntion algorithm. Values indicate number of true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN). The classifier chooses the gesture with higher likelihood when more than one gesture satisfies the threshold.

TP	TN	FP	FN	accuracy	recall	precision
98	123	27	3	0.88	0.97	0.78

However results during actual experiment with children were very bad

Four Tasks of a Robot-assisted Autism Spectrum Disorder Diagnostic Protocol: First Clinical Tests

Frano Petric, Kruno Hrvatinić, Anja Babić, Maja Cepanec, Jasmina Stošić and Sanja Šimleša

The robot carrying out the evaluation of child behavior by detection of specific events while performing some tasks

TABLE V. SESSION RESULTS FOR CHILD ASD001.

Task	Successful		Vocalization		Speech		Code	
	H	R	H	R	H	R	H	R
Response to name	✓	✓	✓	✓	✓	×	0	1
Joint attention	×	×	×	✓	✓	✓	3	3
Play request	✓	✓	✓	×	✓	✓	0	1
Imitation	✓	✓	×	✓	✓	×	1	1

Results show a strong divergence between human rater on video recording and robot

2013 IEEE RO-MAN: The 22nd IEEE International Symposium on
Robot and Human Interactive Communication
Gyeongju, Korea, August 26-29, 2013

Comparing two LEGO Robotics-Based Interventions for Social Skills Training with Children with ASD

Jordi Albo-Canals, Marcel Heerink, Marta Diaz, Vanesa Padillo, Marta Maristany, Alex Barco, Cecilio Angulo, Ariana Riccio, Lauren Brodsky, Simone Dufresne, Samuel Heilbron, Elissa Milto, Roula Choueiri, Dan Hannon, and Chris Rogers

TABLE II. CODE SCHEME FOR BEHAVIOUR ANALYSES BARCELONA

Group	*Behavior*	*Description*
Social Interaction	Ask for Help	How many times the children ask for help from therapist or technician
	Ask for permission	How many times the children ask for permission from therapist or technician
	Group proxemics	When groupmates stand within 120 cm, or what is describes as the limit of ''personal distance'' in conversational interaction, of each other by [3]
	Shared gaze	When groupmates look at the same object or at each other [3]
	Pointing Behaviour	Indicating the robots, computers or activity material (i.e.: cards, board, etc.) to either the experimenter or groupmates (i.e.: during a conversation/explanation even if they don't saying nothing) through pointing at them [3]
	Shared Positive affect	How many times the children would laugh or smile with groupmates [3]
	Joint attention	Initiation and response
States of play	No playing	The play it hasn't started or user it isn't doing nothing related with the play
	Disengagement	Participant is no focusing to the task or other individuals within the group or the other group (not really interested) [15]
	Co-operative activity	Subject works with another person by turn-taking, or discussing play outcomes but where tasks are distributed Individual works together with somebody e.g. hands on something at same time or discussing outcome together [15]
	Onlooker	Participant is watching what the other individuals within the own group are doing but does not actively take part or is watching the experimenter [15]
	Onlooker of the other group	Participant is watching what the other group are doing and isn't playing or are speaking with the other group
	Playing alone	Subject is playing (with activity material, pc or computer) or focused to the task alone (the other user can be onlooker)
Childre n	Robot manipulatio	Direct interaction manipulation with the robot (e.g. holding, connecting,

Human observer labeling of events and behaviors on video recordings

TABLE III. CODING SCHEME FOR BEHAVIOUR ANALYSES BOSTO

Group	*Behavior*	
Social Skills	Non Verbal Communication	Joint Attention (initiation and response)
		Gestures/pointing
		Showing
	Conversation with Partner	Initiation of conversation
		Response to conversation
		Conversation turns
		Commenting
		Interrupts partner
		Asks for help
		Arguing
		Resolved by themselves or adult intervention
	Conversation with Adult	Asks for help
		Teacher interferes to resolve arguing
		Teacher prompts an interaction between partners
States of play	Behaviour	Echolalia/Scripting
		Self-stemming behaviors
		Hyper/Hypo active (yes/no, duration)
		Frustration
		Sharing positive affect
		Difficulty turn taking/grabbing from partner/other children
		Description of other behaviors (ex. personal space, transitions)

Self assessment questionaire

Figure 1. Session questionaire in Boston

	Questions:	 No	 Neutral	 Yes	Teacher
1	I listened to my partner's ideas				
2	I worked well together with my partner				
3	I had fun in robotics today				
4	My partner had fun in robotics today				
5	I controlled my anger or frustration well today				
6	I followed the classroom rules				
7	How was today's activity?	Hard	Ok	Easy	

Figure 2. Session questionaire in Barcelona

¿?	 NO	 UN POCO	 BASTANTE	 SÍ
Me ha gustado la actividad de hoy				
He cumplido las reglas del grupo				
He trabajado en equipo				
La actividad de hoy me ha parecido fácil				
Me he divertido trabajando con mis compañeros				
He sido un buen compañero				
He escuchado las ideas de mis compañeros				
Estoy esperando la próxima sesión con el robot				
Me he divertido trabajando con el robot				
He controlado mi enfado				

2015 IEEE 14th International Conference on Machine Learning and Applications

A Machine Learning Based WSN System for Autism Activity Recognition

Sami S. Alwakeel[⇑±], Bassem Alhalabi [±], Hadi Aggoune[±→], Mohammad Alwakeel[±]

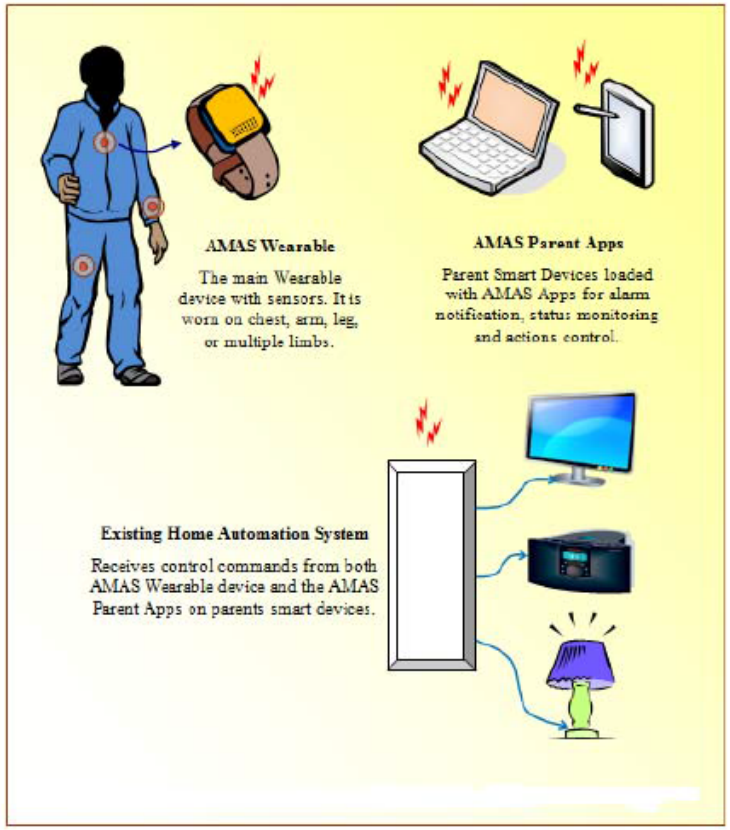


Figure 1: ACSA wearable Sensor &ACSA Parent Apps

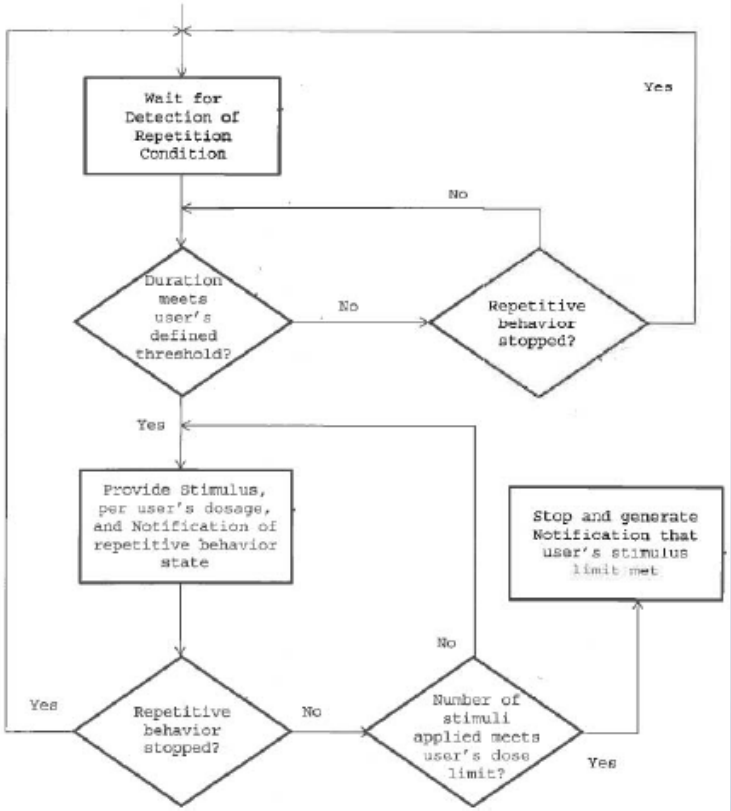


Figure 3; A flowchart representation of the wearable ACSA's logical operations

Conclusions

Conclusions

- Stringent ethics and data protection limit the measurement devices and records
- Most aseptic measure is the self evaluation, but it suffers from subjectivity and imprecision
- Most accurate measures are video recordings, but they are very sensitive to data protection issues.
- Wearable sensors provide anonymity but interpretation and analysis is difficult