ASC-Inclusion: Interactive Emotion Games for Social Inclusion of Children with Autism Spectrum Conditions

Björn Schuller,
Erik Marchi
Machine Intelligence and Signal Processing group
Technische Universität München
80333 München, Germany
schuller@tum.de
erik.marchi@tum.de

Simon Baron-Cohen,
Helen O'Reilly
Autism Research Centre
University of Cambridge
Cambridge, UK
sb205@cam.ac.uk
heo24@cam.ac.uk

Peter Robinson,
Ian Davies
Computer Laboratory
University of Cambridge
Cambridge, UK
peter.robinson@cl.cam.ac.uk
Ian.Davies@cl.cam.ac.uk

Ofer Golan,
Shimrit Friedenson,
Shahar Tal
Department of Psychology
Bar-Ilan University
52900 Ramat-Gan, Israel
ofer.golan@biu.ac.il
shimfri@gmail.com
shahar0190@gmail.com

Shai Newman,
Noga Meir,
Roi Shillo
Compedia Ltd
Ramat-Gan, Israel
newmans@compedia.net
nogam@compedia.co.il
roishillo@compedia.net

Antonio Camurri,
Stefano Piana
InfoMus Lab, DIST
University of Genova
16145 Genova, Italy
antonio.camurri@unige.it
steto84@infomus.org

ABSTRACT

Autism Spectrum Conditions (ASC) are neurodevelopmental conditions, characterized by social communication difficulties and restricted and repetitive behaviour patterns. The ASC-Inclusion project aims to create and evaluate the effectiveness of an internet-based game platform, directed for children with ASC and those interested in their inclusion. The platform combines several state-of-the art technologies in one comprehensive virtual world, including analysis of users' gestures, facial and vocal expressions using standard microphone and webcam, training through games, text communication with peers and smart agents, animation, video and audio clips. We present the progress in realising such a serious game platform and provide results for the different modalities.

Keywords

Auism Spectrum Conditions, inclusion, virtual worlds, computerised environement, emotion recognition

1. INTRODUCTION

Autism Spectrum Conditions (ASC) are neurodevelopmental conditions, characterized by social communication difficulties and restricted and repetitive behaviour. Individuals with ASC have marked difficulties using verbal and non-verbal communication for social interaction. They lack the sense of social reciprocity and fail to develop and maintain age appropriate peer relationships [8], [45]. Current studies suggest 1% of the population might fit an ASC diagnosis [3]. This finding has been replicated across cultures and age bands, stressing the importance of accessible, cross-cultural means to support this growing group. The social communication deficits, characteristic of ASC, have a pervasive effect on the ability of these individuals to meet age appropriate developmental tasks, from everyday negotiation with the teacher or the shopkeeper to the formation of significant relationships with peers. Due to these difficulties, children with ASC are at high risk of social exclusion. They often suffer from rejection, bullying and isolation [19]. In the long run, these social communication difficulties hamper the independent functioning of individuals with ASC, including their attainment of occupational and residential goals. They affect as their ability to find friends and intimate partners, and increase the likelihood of psychiatric disorders [31]. It is therefore important to attend to the social communication difficulties individuals with ASC experience as early as possible. Indeed, studies of intervention into ASC have shown that the earlier the intervention is provided, the more effective it is [32].

The ability to attend to socio-emotional cues, interpret them correctly and respond to them with an appropriate expression plays a major role in social communication. Three decades of research have shown that children and adults with ASC experience significant difficulties recognizing and expressing emotions and mental states [2, 28]. These difficulties are apparent when individuals with ASC attempt to recognize emotions from facial expressions [11, 15, 23, 29], from vocal intonation [9, 24], from gestures and body language [26, 47], and from the integration of multimodal emotional information in context [22, 52, 57]. Limited emotional expressiveness in non-verbal communication is also characteristic in ASC, and studies have demonstrated individuals with ASC have difficulties directing appropriate facial expressions to others [35, 36, 34], modulating their vocal intonation appropriately when expressing emotion [39, 42, 46] and using appropriate gestures and body language [1]. Integration of these non-verbal communicative cues with speech is also hampered [14].

Attempts to teach emotion and mental state recognition, either on an individual basis [10, 18] or as part of social skills group training [6, 27, 53], have shown mixed results. Besides improvement on taught material, most studies reported limited generalization from taught curriculum to situations not included in the training program. In addition, such training requires facilitation by trained specialists, is not always available, and may be quite expensive. When conducted in groups, it may also deter the more anxious or withdrawn children [25]. Adding to these limitations is the significant growth in the prevalence of ASC, mentioned above. This increase in prevalence requires greater intervention and rehabilitation resources, to be provided by well-trained clinicians and educators. At present, many countries across Europe – and globally – lack an adequate policy response to the needs of this growing population of young people, resulting in their exclusion from society and societal resources. There is therefore an urgent need to seek for new innovative methods for approaching children with ASC and for fostering their integration into society. A solution to the shortage of trained service providers for individuals with ASC may be found in ICT, which enables users everywhere to enjoy state-of-the-art professional support online. The computerized environment is especially appealing for individuals with ASC, due to its predictable, controllable and structured nature [43], which enables them to use their strong systemizing skills. Systemizing is the drive to analyze or build rule-based systems, allowing one to predict the behaviour of the system and control it. In contrast to their socio-emotional difficulties, individuals with ASC have intact and sometimes superior systemizing skills [5]. Their attraction to systems is apparent in the circumscribed interests they possess [4], such as spinning objects, mechanics, and computers. These are attractive because they are systematic, repetitive, and therefore predictable. These special interests could therefore be harnessed when teaching children with ASC, in order to keep them intrinsically motivated [50]. Motivation is indeed a major challenge when aiming to teach socio-emotional communication skills to children with ASC. Whereas as adolescents, the motivation for social communication and interaction increases [55, 56], implicit social motivation is usually lower in children with ASC [13, 58] and their interest in such training needs to be initiated and retained externally [37]. As noted above, harnessing the child's circumscribed interests when teaching socio-emotional communication may assist in raising interest and motivation [20].

The ASC-Inclusion project suggests advanced ICT-enabled solutions and serious games for the empowerment of children with ASC who are at high risk of social exclusion. ASC-Inclusion aims to create an internet-based platform that will assist children with ASC, and those interested in their inclusion, to improve their socio-emotional communication skills, attending to the recognition and expression of socio-emotional cues and to the understanding and practice of conversational skills. ASC-Inclusion combines several state-of-the-art technologies in one comprehensive game environment, including analysis of users' gestures, facial and vocal expressions, training through games, text chatting, animation, video and audio clips. Despite the innovative technologies involved, the ASC-Inclusion is aimed for home use. Though designed to assist children with ASC, ASC-Inclusion could serve other population groups characterized by deficient emotional understanding and social skills, such as children with learning difficulties [7, 30], attention deficit and hyperactivity disorder (ADHD) [12], behavioural and conduct problems [54], or socio-emotional [49].

The remainder of this paper is structured as follows: first, a detailed description of the user requirements and specification is given (Section 2); then we describe the three modalities namely face, voice and body gesture (Sections 3, 4, 5) and the platform (Section 6). We next comment on content creation (Section 7) before concluding the paper in Section 8.

2. USER REQUIREMENTS

Previous literature reports social communication difficulties in individuals with ASC as well as enhanced abilities in other non-social areas such as systemizing [5]. This provided us with a greater insight into the social, cognitive and behavioural abilities of these potential end users. Review of the design, content and effectiveness of the Mind Reading DVD as an intervention [21] and The Transporters DVD [20] which have shown to enhance the ability of children with ASC to recognize emotions, permitted identification of a method through which learning material can be presented in an engaging manner which capitalizes on the enhanced systemizing skills seen in individuals with ASC.

7 focus groups were observed to help identify the preferences and interaction style of children with ASC when using virtual environments (VEs). The feedback from these groups helped to specify the user requirements of the virtual environment and to design accordingly the computerized therapeutic intervention. The results from Hebrew user focus groups found that the children enjoyed using the virtual environments and the majority of children could navigate independently in the environment, activate different functions of the VE and change the settings. It was noted that having an overall mission to guide the children's interaction in the VE was a motivating tool. The user focus groups also provided feedback on the game genres that the children enjoyed most. Specifically it was found that the children enjoyed being able to design their own avatar and home, although they would like the avatar to have a more interactive role. The children also enjoyed strategy and thought games, drawing games, and a provided communication tool. As one would assume, the children also liked to gain rewards from playing games. Another aim of these focus groups was to obtain feedback from the children on the theme of the proposed ASC-Inclusion VE – a research camp, and also their comprehension of what research is and the types of tools a researcher might use. The majority of children understood and liked the theme of the VE and also the concept of a researcher. Both feedback from the children and parents on the games has been very beneficial in refining and adjusting the games to enhance the children's learning and also to engage and maintain the interests of the children. The final aim of

the focus groups was to (1) confirm if the type of content (social scenes, short stories) we intended to create/film for the platform was understandable to children, (2) to make sure they could understand and could give judgements of people's feelings and thoughts just from verbal descriptions, to see if this is a another possible mode of games design, (3) to get ideas for social scenes that are part of the children's world, and to test if they matched those created by the ASC-Inclusion team's scenes and (4) to examine the children's ability to sustain attention to emotion tasks and their ability to complete tasks these tasks successfully. It was found that the children produced similar stories to those created by the ASC-Inclusion team, the children could understand verbal descriptions of emotions and could sustain attention to these tasks. The results from a second set of Swedish user focus groups mainly investigated different game genres with children with ASC. The results from these focus groups found that the games that were easy to understand and had a good game concept were rated highest by the children. The researchers noted that a number of the games presented to the children and which the children enjoyed playing could be easily adapted to become emotion recognition games, but that better reward systems and the ability to change difficulty level was necessary to keep the children engaged and motivated. The most valuable feedback from these focus groups was in relation to the game Tinkatolli[1]. In Tinkatolli each child is represented in the VE through an avatar character, the other avatars in the VE are their classroom friends, and the children can send messages to one another. It took some time for the children to realise that the other avatars were their classroom friends but once this realization occurred, the children were very excited and instantly began communicating with one another both virtually within the game and across the class room. Tinkatolli promoted and prompted social interaction and has a therapeutic benefit as the children initiated play and social contact with the other children in the virtual world. A similar tool will thus be developed for the ASC-Inclusion VE.

Two specialist focus groups were held, these focus groups consisted of experts in the field of ASC. The purpose of these groups was to provide feedback regarding the VE concept and implementation, as well as to highlight difficulties common to ASC children which may affect their ability to interact with the VE. A number of potential issues and recommendations were made by the group. These recommendations have been used to guide the development of the virtual environment.

The information gathered from review of literature and technologies developed for children with ASC, as well as the results from the focus groups have been fundamental in designing both the visual and functional aspects of the VE to match the needs of the intended user – children with ASC 5–10 years of age.

3. FACE ANALYSIS

A stand-alone prototype of the Facial Affect Computation Engine (FACE) has been constructed. This combines earlier work [33] on inference of complex mental states (agreeing, concentrating, disagreeing, interested, thinking, unsure, and further) from continuous video with a more recent facial tracker (Figure 1) based on head and facial actions such as head yaw, pitch, and roll, eyebrow raise, lip pull and pucker, lips parting, jaw drop, mouth open and teeth present. The resulting system is robust and runs in real time. Our approach integrates machine vision and probabilistic machine learning to represent a computational model of mind-reading. The model is a multi-level Dynamic Bayesian Network that abstracts raw video input into three levels. Each level conveys face-based information at different extents of spatial and temporal abstraction. The automated mind-reading system implements this model by combining top-down predictions of mental state models with bottom-up vision-based processing of the face to recognize complex mental states in real-time video.

[1] https://www.tinkatolli.me/

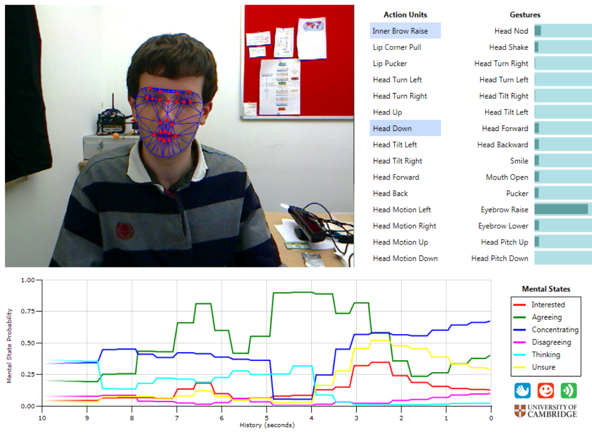

Figure 1: Face analyser.

4. VOICE ANALYSIS

For the training of a voice analyser, a data set of prototypical emotional utterances containing sentences spoken in Hebrew by a total number of 20 children has been collected with typically developing children and children with ASC. The focus group consists of nine children (8 male and 1 female) at the age of 6 to 12 years. All children were diagnosed with ASC by trained clinicians. 11 typically developed children (5 female and 6 male) at the age of 6 to 9 years served as control group. In order to limit the effort of the children, the experimental task was designed to focus on the Ekman 'big six' basic emotions except disgust: happy, sad, angry, surprised, afraid plus other three mental states: ashamed, clam, proud, and neutral. A complete description of the database is given in [41]. Considering the outcome of the evaluations described in [41] and [40], we defined potential features and descriptors that are relevant for emotion perception. Applying the openSMILE audio feature extractor [17], a large set of prosodic, spectral and voice quality low-level features has been extracted from the recorded utterances. A deeper look was taken at individual features in order to make quantitative statements on the differences between the classes. We have chosen features modelling pitch, energy, and duration basing on the assumption that such features can be conveyed to children rather easily by prompting them to speak louder or more quietly, to raise or lower pitch, or to speak faster or slower, etc., resulting in a list of prosodic descriptors such as statistical functionals of: energy by the sum of auditory spectrum at different frequency bands (from 20 Hz to 8 kHz) and root-mean-square signal frame energy, pitch by fundamental frequency contour, and duration by modelling temporal aspects of F0 values, such as the F0 onset segment length.

A new networking layer was added to openSMILE. This includes a basic platform independent socket API available both for Linux and Windows, a network message and data sender, and a remote control interface. With this interface, openSMILE can be run on live audio input on a backend system; output data can be streamed to a client over a standard TCP/IP network. This is required for easy integration with other components of the game. The network components enable the processing in the backend to be started, paused and resumed, and stopped remotely from a client process. The acoustic parameters are computed on several temporal levels. The lowest level is the frame level, i. e., acoustic low-level descriptors are computed from short overlapping frames of audio data (20–60 ms), which are sampled at a rate of 10 ms. The next level is constructed by applying statistical functionals to the low-level descriptor contours over fixed length windows. The length of these windows can either be constant, typically 1–5 seconds, or dynamically correspond to linguistic meaningful units such as words, phrases, or utterances. We refer to these levels as the constant length or dynamic length supra-segmental levels. The data rate required for the transmission of acoustic parameters decreases from the frame level to the supra-segmental levels. The dynamic length supra-segmental level has no fixed rate at which data is sent, while the other two levels have constant rates. The Graphical User Interface (GIU) for the current ASC-Inclusion Voice Analyser (Figure 2) includes two parts: the upper part in which the system gives a visual feedback on the recognised emotion and the bottom part in which the parameters for emotion perception are tracked in real time. The upper part of the GUI concerns emotion recognition. It shows arousal and valence in a 2D plot by colouring the 4 quadrants depending on the recognition results, according to Plutchik's wheel of emotions [48]. In addition to arousal and valence, it shows also the emotion name while for the others it will just colour up accordingly to arousal and valence colours mapping. The bottom part of the GUI is dedicated to real-time parameters tracking. It shows Energy and Pitch over time. Moreover speed and pitch standard deviation are showed as moving bars.

5. GESTURE ANALYSIS

Full-body expressive features to be analysed in the game are inspired by several sources: biomechanics, computer vision, experimental psychology, and humanistic theories, including Laban's Effort Theory [38]. The list of candidate features is the following: kinetic energy of body parts (hands, head, upper body), symmetry of hands movements and posture with respect to vertical axis and forward/backward, forward/backward leaning of upper part of the body, directness (direct/flexible in terms of Laban's Effort), impulsivity, fluidity, light/heavy gesture quality, the relative position between head and shoulder, openness of the posture (contraction/expansion), rigidity and swinging of the whole body or of body parts. Several solutions have been explored for the recordings of full body movements of subjects. Extraction and analysis of non-verbal expressive features from full body movement, indeed, is strongly dependent on the quality of the recorded data. In particular, the presence of noise might negatively affect the measurement of features; for comparing

Figure 2: Voice analyser.

the data from several acquisition systems, it is necessary to guarantee the synchronization of the recorded data. The EyesWeb[2] software platform – which provides the basis of the analysis in the game – has been used to perform the tests. EyesWeb provides many modules and fully supports the synchronization of data from several sources. New EyesWeb XMI software modules to extract expressive features were developed that can compute features from both 2D and 3D user tracking data coming from a 3D depth sensor (Kinect, ASUS Xtion, Softkinetic). The newly developed algorithms and software (Figure 3) include modules for the computation of Kinematic features such as positions, trajectories, velocities, accelerations, and jerks directly computed from mocap data or images. Joints positions and trajectories are extracted from mocap data. User tracking can be performed using traditional web-cams, in this case only 2D coordinates will be available, or using cheap depth camera based motion capture systems like Kinect, ASUS Xtion Pro or other RGB-D sensors, that can extract both 2D and 3D MoCap data simultaneously. Velocities, accelerations, and jerk related to each single point identified during the tracking, are computed directly starting from the coordinates. 2D and 3D velocities and higher order derivative coordinates computations are performed by two EyesWeb XMI blocks developed specifically to carry on these operations. Further EyesWeb software modules extract in real-time movement expressive features. Main software modules for the Asc Inclusion EyesWeb library are described in the following. Kinetic Energy (movement activity, or motion index) is a relevant feature in recognizing emotion from the full-body movement. An EyesWeb software module computes Kinetic Energy index from 3D or 2D coordinates. Contraction Index (CI) is a measure (range from 0 to 1) of how the user's body uses the space surrounding it: the 3D version of CI is computed by comparing the volume of the Bounding Volume (BV) that is the minimum parallelepiped surrounding the user's body and an approximation of the volume of the density (DI) of the 3D coordinates. The 2D CI uses a technique related to the bounding box, i. e., the minimum rectangle surrounding the user's body. Smoothness: an index derived from the

[2] http://www.infomus.org/eyesweb_ita.php

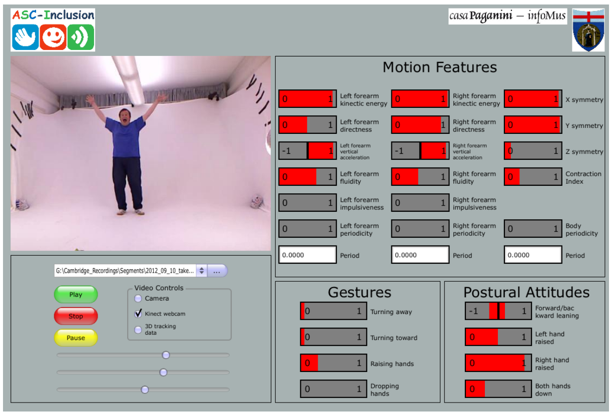

Figure 3: Body gesture analyser.

measure of the curvature of a speed trajectory; both 2D and 3D curvature can be measured. Symmetry: symmetry indexes starting from 2D and 3D MoCap data. Directness: given a target in space, a direct movement is characterized by an almost straight line trajectory, in contrast with a flexible movement which changes direction at each instant. The directness index is computed from a trajectory tends to assume values near to 1 if a movement is direct and low values (near 0) when flexible. 2D and 3D Directness of a movement is computed by a single EyesWeb XMI block that, given a trajectory of a joint, calculates its directness. Periodicity: and index based on the periodicity transform [51]. This module analyzes the input and produces a sequence of periods and their contribution to the input sequence. Rigidity: a measure of the relative movement of different parts of the body. Impulsiveness: an impulsive gesture ([16], [44]) is performed without a significant preparation phase. We have tested the above list of implemented indexes on a movement archive and with real-time measures using Kinect. Currently, we are refining and tuning the algorithms by new recordings of children using the Sualysis motion capture system and hd videocamera. This will result in the second version of the EyesWeb ASC library for body gesture analysis.

6. PLATFORM

An overview of the game platform architecture is given in Figure 4. The platform communicates with the Emotion Agent that obtains the current emotion which the player is having from Apache ActiveMQ (message queue). The partners' agents analyse input from the sensors (webcam, microphone, etc.) and send their results to the ActiveMQ. The data is encoded by the W3C EmotionML. Each message includes time stamps as per the time interval from which the emotion estimation was extracted. The sub systems' (gestures, face, and voice analysis) agents always run in the background and add results into the queue. The higher level API between the Emotion agent and the platform will be similar to the above and will define for each emotion detected at the low level API (as well as for arousal and valance dimensions) its confidence by each modality.

In addition, the platform includes the following basic procedures: User Registration and Login – including setting the profile, Carers' registration and login: Child login: Once registered, the child has a unique username and avatar (see below). User profile: For each user, a profile is created by the system, containing details such as age, preferences, and on-going play patterns. The personalisation infrastructure of the system uses this information to ease personalise the child's experience, e. g., present the games in an initial difficulty level that fits the child. Saving user data: Anything a user achieves, purchases, and creates in the VE is saved in the system data base, e. g., progress in the training program, paintings and designs created by the child in a creative activity, and collectible items. The data is sent using web services. Initial parents' site, Interaction manager: The interaction manager was defined and can be activated and enable the various user selections and activities, navigation between activities and locations. Avatar Choosing and customizing: In the avatar application designed for the VE, the child can create an avatar, customise its facial features and expressions, and choose cloths and accessories for it. Being one of the strongest motivational elements of VE's in general, and as found for ASC-Children too , the avatar application is strongly connected to the rewards system of the VE. Upon entering the VE the child has a basic set of items for the avatar, and he/she can purchase more using virtual money they earn by advancing in learning tasks. Some items can only be purchased after reaching a certain amount of experience points and rank in the VE storyline. Content templates: ASC children are known to be fond of collecting, so these applications are also to be harnessed to motivate the VE users to be active in the VE and progress in learning, and are therefore related to the rewards system.

The world's theme and narrative were created to face the challenge of making the children go through a learning program which aims to help them in coping with one of the issues that are maybe the hardest for them: recognising, understanding, and expressing emotions. The learning plan is the heart of the VE and was designed to be both didactic and engaging. The rewards system and motivational elements were designed to support it [58]. The look and feel of the

#	Emotion	Valence	Category	# Levels
1	Happy	positive	basic	2
2	Sad	negative	basic	2
3	Afraid	negative	basic	2
4	Angry	negative	basic	2
5	Disgusted	negative	basic	2
6	Surprised	neutral	basic	2
7	Excited	positive	individual	1
8	Interested	positive	individual	1
9	Bored	negative	individual	1
10	Worried	negative	individual	1
11	Disappointed	negative	individual	1
12	Frustrated	negative	individual	1
13	Hurt	negative	individual	1
14	Kind	positive	social	1
15	Jealous	negative	social	1
16	Unfriendly	negative	social	1
17	Joking	positive	social	1
18	Sneaky	negative	social	1
19	Ashamed	negative	self-conscious	1
20	Proud	positive	self-conscious	1

Table 1: 20 emotions specified for the game. Levels refers to the number of different intensity levels.

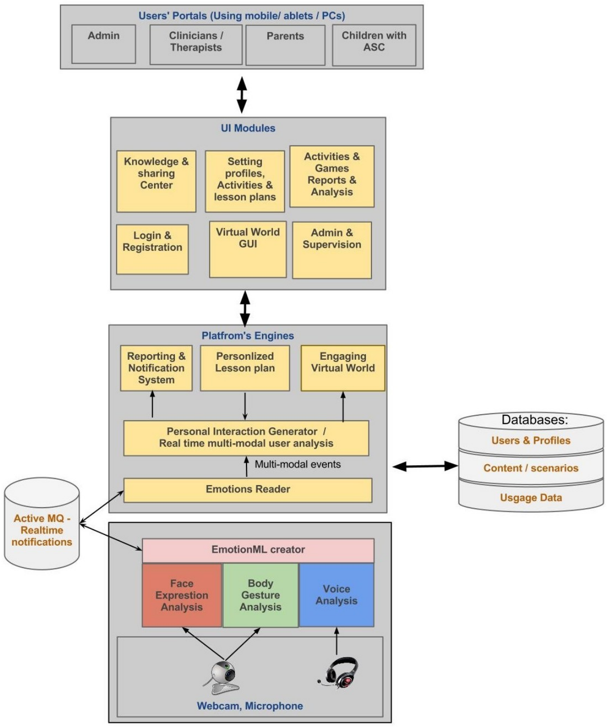

Figure 4: Platform architecture.

virtual world environment were designed according to the theme and narrative , and the main locations and Non Player Characters were developed.

Current motivational elements in the VE are: a designable virtual home, including a 'tickets machine' application to take a canoe and go to your own home or visit a friend's home, the "Linko" communication device, and a collector's album application including buying and replacing cards.

7. CONTENT CREATION

In order to determine the most important emotions for social interaction in children with ASC, an emotions survey was developed and completed by parents, typically developing adults and parents of children with ASC. 20 emotions were identified by the survey as being the most important for social interaction and selected to be taught through the platform (cf. Table 1). To teach children with ASC about the recognition and expression of emotions through facial expressions, vocal expressions and body gestures; audio and video material was recorded of actors portraying these emotional displays. To demonstrate these emotions within a social context, 120 social scenes were also filmed, 6 per emotion. Scripts were written to help the actors portray the emotions through each of the filmed modalities. In the game, the child can earn coins for carrying out their investigation work, the emotion tutorials. The children can then use these coins to buy things or earn play time in their favourite games. Positive feedback will always be provided each time the child answers a question in a game correctly, or makes the wanted action. The amount of praise provided will match the degree of success. If the child answers incorrectly, the game will provide a hint and will then encourage him/her to try again.

Figure 5: Research centre environement.

8. CONCLUSIONS

We introduced the current state of the gaming platform ASC-Inclusion targeted to children aged 5 to 10 years with ASC. User requirements have been analysed, identifying the needs of target audience for the creation of user scenarios for the platform subsystems. A Facial Affect Computation Engine (FACE) that takes video from a standard webcam, infers probabilities for discrete mental state conditions, and displays the results. The on-line acoustic feature extractor, which calculates the speech features in real-time, is further integrated and extended according to the relevant descriptors that need to be tracked for ASC vocal expression analysis. A GUI gives visual feedback of these parameters. A collection of EyesWeb XMI software modules implementing computational models and algorithms extracts in real-time a set of low-level expressive gesture features. Further, a basic platform that provides procedures and infrastructure forms the core piece of the game. Video and audio content forms the main material of the learning tutorials and serious game and has accordingly been created. An evaluation design which will test the psychological and clinical effectiveness of the sub-systems as well as their integration has been developed.

The upcoming improvements will enable integration of the different modalities. Further, we will focus on evaluation with children in the loop.

Acknowledgment

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 289021 (ASC-Inclusion).

9. ADDITIONAL AUTHORS

Sven Bölte, Daniel Lundqvist, Steve Berggren (Center of Neurodevelopmental Disorders, Karolinska Institute, Stockholm, Sweden, email: `(sven.bolte |daniel.lundqvist | steve.berggren) @ki.se`), Aurelie Baranger and Nikki Sullings (Autism-Europe aisbl, 1000 Brussels, Belgium, email: `(Aurelie.Baranger | nikki.sullings) @autismeurope.org`).

10. REFERENCES

[1] T. Attwood. *Asperger's syndrome: A guide for parents and professionals*. Jessica Kingsley Pub, 1998.

[2] S. Baron-Cohen. *Mindblindness: An essay on autism and theory of mind*. MIT press, 1997.

[3] S. Baron-Cohen, F. J. Scott, C. Allison, J. Williams, P. Bolton, F. E. Matthews, and C. Brayne. Prevalence of autism-spectrum conditions: Uk school-based population study. *The British Journal of Psychiatry*, 194(6):500–509, 2009.

[4] S. Baron-Cohen and S. Wheelwright. 'obsessions' in children with autism or asperger syndrome. content analysis in terms of core domains of cognition. *The British Journal of Psychiatry*, 175(5):484–490, 1999.

[5] S. Baron-Cohen, S. Wheelwright, J. Lawson, R. Griffin, C. Ashwin, J. Billington, and B. Chakrabarti. Empathizing and systemizing in autism spectrum conditions. *Handbook of autism and pervasive developmental disorders*, 1:628–639, 2005.

[6] N. Bauminger. The facilitation of social-emotional understanding and social interaction in high-functioning children with autism: Intervention outcomes. *Journal of autism and developmental disorders*, 32(4):283–298, 2002.

[7] N. Bauminger, H. S. Edelsztein, and J. Morash. Social information processing and emotional understanding in children with ld. *Journal of learning disabilities*, 38(1):45–61, 2005.

[8] C. C. Bell. Dsm-iv: Diagnostic and statistical manual of mental disorders. *JAMA: The Journal of the American Medical Association*, 272(10):828–829, 1994.

[9] J. Boucher, V. Lewis, and G. M. Collis. Voice processing abilities in children with autism, children with specific language impairments, and young typically developing children. *Journal of Child Psychology and Psychiatry*, 41(7):847–857, 2003.

[10] D. M. Bowler and E. Strom. Elicitation of first-order'theory of mind in children with autism. *Autism*, 2(1):33–44, 1998.

[11] G. Celani, M. W. Battacchi, and L. Arcidiacono. The understanding of the emotional meaning of facial expressions in people with autism. *Journal of autism and developmental disorders*, 29(1):57–66, 1999.

[12] D. Da Fonseca, V. Seguier, A. Santos, F. Poinso, and C. Deruelle. Emotion understanding in children with adhd. *Child Psychiatry & Human Development*, 40(1):111–121, 2009.

[13] G. Dawson et al. Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. *Development and Psychopathology*, 20(3):775, 2008.

[14] A. de Marchena and I.-M. Eigsti. Conversational gestures in autism spectrum disorders: Asynchrony but not decreased frequency. *Autism Research*, 3(6):311–322, 2010.

[15] C. Deruelle, C. Rondan, B. Gepner, and C. Tardif. Spatial frequency and face processing in children with autism and asperger syndrome. *Journal of autism and developmental disorders*, 34(2):199–210, 2004.

[16] J. L. Evenden. Varieties of impulsivity. *Psychopharmacology*, 146(4):348–361, 1999.

[17] F. Eyben, M. Wöllmer, and B. Schuller. openSMILE – The Munich Versatile and Fast Open-Source Audio Feature Extractor. In *Proceedings of the 18th ACM International Conference on Multimedia, MM 2010*, pages 1459–1462, Florence, Italy, October 2010. ACM, ACM.

[18] N. Fisher and F. Happé. A training study of theory of mind and executive function in children with autistic spectrum disorders. *Journal of Autism and Developmental Disorders*, 35(6):757–771, 2005.

[19] U. Frith. Autism: Explaining the enigma. *Cambridge, En*, 1989.

[20] O. Golan, E. Ashwin, Y. Granader, S. McClintock, K. Day, V. Leggett, and S. Baron-Cohen. Enhancing emotion recognition in children with autism spectrum conditions: an intervention using animated vehicles with real emotional faces. *Journal of autism and developmental disorders*, 40(3):269–279, 2010.

[21] O. Golan and S. Baron-Cohen. Systemizing empathy: Teaching adults with asperger syndrome or high-functioning autism to recognize complex emotions using interactive multimedia. *Development and Psychopathology*, 18(02):591–617, 2006.

[22] O. Golan, S. Baron-Cohen, and Y. Golan. The 'reading the mind in films' task [child version]: Complex emotion and mental state recognition in children with and without autism spectrum conditions. *Journal of Autism and Developmental Disorders*, 38(8):1534–1541, 2008.

[23] O. Golan, S. Baron-Cohen, and J. Hill. The cambridge mindreading (cam) face-voice battery: Testing complex emotion recognition in adults with and without asperger syndrome. *Journal of autism and developmental disorders*, 36(2):169–183, 2006.

[24] O. Golan, S. Baron-Cohen, J. J. Hill, and M. Rutherford. The 'reading the mind in the voice' test - revised: a study of complex emotion recognition in adults with and without autism spectrum conditions. *Journal of autism and developmental disorders*, 37(6):1096–1106, 2007.

[25] O. Golan, P. G. LaCava, and S. Baron-Cohen. Assistive technology as an aid in reducing social impairments in autism. *Growing Up with Autism: Working with School-Age Children and Adolescents*, pages 124–142, 2007.

[26] J. Grèzes, B. Wicker, S. Berthoz, and B. De Gelder. A failure to grasp the affective meaning of actions in autism spectrum disorder subjects. *Neuropsychologia*, 47(8):1816–1825, 2009.

[27] J. Hadwin, S. Baron-Cohen, P. Howlin, and K. Hill. Can we teach children with autism to understand emotions, belief, or pretence? *Development and Psychopathology*, 8(02):345–365, 1996.

[28] R. Hobson. *Autism and the development of mind*. Psychology Press, 1995.

[29] R. P. Hobson. The autistic child's appraisal of expressions of emotion. *Journal of Child Psychology and Psychiatry*, 27(3):321–342, 1986.

[30] H. B. Holder and S. W. Kirkpatrick. Interpretation of emotion from facial expressions in children with and without learning disabilities. *Journal of Learning Disabilities*, 24(3):170–177, 1991.

[31] P. Howlin. *Autism and Asperger syndrome: Preparing*

for adulthood. Routledge, 2004.

[32] P. Howlin, M. Rutter, M. Berger, R. Hemsley, L. Hersov, and W. Yule. *Treatment of autistic children*. Wiley Chichester, 1987.

[33] R. Kaliouby and P. Robinson. Real-time inference of complex mental states from facial expressions and head gestures. *Real-time vision for human-computer interaction*, pages 181–200, 2005.

[34] C. Kasari, B. Chamberlain, and N. Bauminger. Social emotions and social relationships: Can children with autism compensate? In J. Burack, T. Charman, N. Yirmiya, and P. Zelazo, editors, *Development and autism: Perspectives from theory and research*, pages 309–323. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, 2001.

[35] C. Kasari, M. Sigman, P. Mundy, and N. Yirmiya. Affective sharing in the context of joint attention interactions of normal, autistic, and mentally retarded children. *Journal of autism and developmental disorders*, 20(1):87–100, 1990.

[36] C. Kasari, M. Sigman, N. Yirmiya, and P. Mundy. *Affective development and communication in young children with autism*. Paul H. Brookes Publishing, 1993.

[37] R. L. Koegel, T. W. Vernon, and L. K. Koegel. Improving social initiations in young children with autism using reinforcers with embedded social interactions. *Journal of autism and developmental disorders*, 39(9):1240–1251, 2009.

[38] R. Laban and L. Ullmann. *The mastery of movement*. ERIC, 1971.

[39] H. Macdonald, M. Rutter, P. Howlin, P. Rios, A. L. Conteur, C. Evered, and S. Folstein. Recognition and expression of emotional cues by autistic and normal adults. *Journal of Child Psychology and Psychiatry*, 30(6):865–877, 2006.

[40] E. Marchi, A. Batliner, B. Schuller, S. Fridenzon, S. Tal, and O. Golan. Speech, Emotion, Age, Language, Task, and Typicality: Trying to Disentangle Performance and Feature Relevance. In *Proceedings First International Workshop on Wide Spectrum Social Signal Processing (WS3P 2012), ASE/IEEE SocialCom 2012*, Amsterdam, The Netherlands, September 2012. ASE/IEEE, IEEE.

[41] E. Marchi, B. Schuller, A. Batliner, S. Fridenzon, S. Tal, and O. Golan. Emotion in the Speech of Children with Autism Spectrum Conditions: Prosody and Everything Else. In *Proceedings 3rd Workshop on Child, Computer and Interaction (WOCCI 2012), Satellite Event of INTERSPEECH 2012*, Portland, OR, September 2012. ISCA, ISCA.

[42] J. McCann and S. Peppe. Prosody in autism spectrum disorders: a critical review. *International Journal of Language & Communication Disorders*, 38(4):325–350, 2003.

[43] D. Moore, P. McGrath, and J. Thorpe. Computer-aided learning for people with autism–a framework for research and development. *Innovations in Education and Teaching International*, 37(3):218–228, 2000.

[44] C. T. Nagoshi, J. R. Wilson, and L. A. Rodriguez. Impulsivity, sensation seeking, and behavioral and emotional responses to alcohol. *Alcoholism: Clinical and Experimental Research*, 15(4):661–667, 2006.

[45] W. H. Organization. *The ICD-10 classification of mental and behavioural disorders: diagnostic criteria for research*. World Health Organization, 1993.

[46] R. Paul, A. Augustyn, A. Klin, and F. R. Volkmar. Perception and production of prosody by speakers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 35(2):205–220, 2005.

[47] R. Philip, H. Whalley, A. Stanfield, R. Sprengelmeyer, I. Santos, A. Young, A. Atkinson, A. Calder, E. Johnstone, S. Lawrie, et al. Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. *Psychological medicine*, 40(11):1919–1929, 2010.

[48] R. Plutchik. *Emotion, a psychoevolutionary synthesis*. Harper & Row New York, 1980.

[49] S. D. Pollak, D. Cicchetti, K. Hornung, and A. Reed. Recognizing emotion in faces: developmental effects of child abuse and neglect. *Developmental psychology*, 36(5):679, 2000.

[50] C. Sainsbury. *Martian in the playground: understanding the schoolchild with Asperger's syndrome*. Sage Publications Limited, 2009.

[51] W. A. Sethares and T. W. Staley. Periodicity transforms. *Signal Processing, IEEE Transactions on*, 47(11):2953–2964, 1999.

[52] L. B. Silverman, L. Bennetto, E. Campana, and M. K. Tanenhaus. Speech-and-gesture integration in high functioning autism. *Cognition*, 115(3):380–393, 2010.

[53] M. Solomon, B. L. Goodlin-Jones, and T. F. Anders. A social adjustment enhancement intervention for high functioning autism, asperger's syndrome, and pervasive developmental disorder nos. *Journal of Autism and Developmental Disorders*, 34(6):649–668, 2004.

[54] D. Stevens, T. Charman, and R. J. Blair. Recognition of emotion in facial expressions and vocal tones in children with psychopathic tendencies. *The Journal of genetic psychology*, 162(2):201–211, 2001.

[55] D. Tantam. Psychological disorder in adolescents and adults with asperger syndrome. *Autism*, 4(1):47–62, 2000.

[56] D. Tantam. The challenge of adolescents and adults with asperger syndromes. *Child and Adolescent Psychiatric Clinics of North America; Child and Adolescent Psychiatric Clinics of North America*, 2003.

[57] N. Yirmiya, M. D. Sigman, C. Kasari, and P. Mundy. Empathy and cognition in high-functioning children with autism. *Child development*, 63(1):150–160, 1992.

[58] S.-V. Zeeland, A. Ashley, M. Dapretto, D. G. Ghahremani, R. A. Poldrack, and S. Y. Bookheimer. Reward processing in autism. *Autism Research*, 3(2):53–67, 2010.