

# **Teaching Facial Expression Production in Autism: The Serious Game JEMImE**

Charline Grossard<sup>1,2\*</sup>, Stéphanie Hun<sup>3</sup>, Arnaud Dapogny<sup>1</sup>, Estelle Juillet<sup>2</sup>, Fanny Hamel<sup>2</sup>, Heidy Jean-Marie<sup>4</sup>, Jérémy Bourgeois<sup>3</sup>, Hugues Pellerin<sup>2</sup>, Pierre Foulon<sup>4</sup>, Sylvie Serret<sup>3</sup>, Ouriel Grynszpan<sup>1</sup>, Kevin Bailly<sup>1</sup>, David Cohen<sup>1,3</sup>

<sup>1</sup>Institut des Systèmes Intelligents et Robotiques (ISIR), Sorbonne Université, Paris, France <sup>2</sup>Service de Psychiatrie de l'Enfant et de l'Adolescent, GHU Pitie-Salpetriere Charles Foix, APHP, Paris, France

<sup>3</sup>Université de Nice Sophia Antipolis, Nice, France

<sup>4</sup>Genious Healthcare, Montpellier, France

Email: \*charline.grossard@aphp.fr

How to cite this paper: Grossard, C., Hun, S., Dapogny, A., Juillet, E., Hamel, F., Jean-Marie, H., Bourgeois, J., Pellerin, H., Foulon, P., Serret, S., Grynszpan, O., Bailly, K., & Cohen, D. (2019). Teaching Facial Expression Production in Autism: The Serious Game JEMIME. *Creative Education*, 10, 2347-2366.

https://doi.org/10.4236/ce.2019.1011167

Received: September 30, 2019 Accepted: November 4, 2019 Published: November 7, 2019

Copyright © 2019 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

http://creativecommons.org/licenses/by/4.0/

#### Abstract

Background: Children with autism spectrum disorder (ASD) show impairment in producing facial expressions adapted to social contexts. Several serious games have been computed to help them dealing with facial expression recognition but very few focused on facial expression production adapted to a given social context. Method: JEMImE is a new serious game which aims to help the player to learn how to produce happiness, anger and sadness in a 3D virtual environment with social situations that should be resolved by producing the correct facial expression. He is guided in the game thanks to facial expression feedback and gauges that help him evaluating the quality of his/her production in real time. The feedbacks on the children productions are timely given by a facial expression recognition algorithm integrated in JEMImE architecture. Specific attention was paid to the visual and motivational aspects of the game. Using a brief feasibility study with children with ASD (N = 23), we evaluated the impression of the players on the game aspect and the possibility to insert algorithmic feedbacks in real time inside JEMImE. Results: During the training phase, children with ASD showed a significant progression during training for facial expression production after algorithmic autonomous feedbacks. This means children understand the challenge and that the algorithmic feedbacks are transparent enough to allow gaming. They expressed an overall good subjective experience with JEMImE in terms of ergonomics, playability, visual aspect and motivation. Conclusion: We conclude that the beta-version of JEMImE shows promising potential and that research should proceed on computing more games and scenarios to offer a longer game exposure to children to allow adequate clinical validation.

#### **Keywords**

Autism Spectrum Disorder, Information Communication Technologies, Facial Expressions Production, Serious Games, Facial Recognition Algorithm

# **1. Introduction**

#### **1.1. Facial Expression in Autism**

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by difficulties in social adaptation and communication and repetitive behavior. Social skills are impaired including facial expression (FE) that is crucial to communicate internal state to others. In typical development (TD), FE can be observed very early in infancy, even if they do not perfectly correspond to their adult's counterparts (Oster, 2005; Sullivan & Lewis, 2003). The course of adult-like FE seems to be progressive: 12-month infants show more specific expressions to a situation than 4-month infants (Bennett, Bendersky, & Lewis, 2005). The learning of adult-like FE increases even in late childhood and at 13 years; children do not yet produce perfectly all FE (Ekman, Roper, & Hager, 1980; Grossard et al., 2018). In addition, FE is modulated by several factors: 1) emotion subtype, with positive emotions being easier to produce than negative emotions (Brun, 2001; Grossard et al., 2018); 2) gender that interacts with emotion subtype (Chaplin & Aldao, 2013; Grossard et al., 2018); 3) cultural context (Grossard et al., 2018); and 4) feedback on FE production, as it seems to increase their quality (Brun, 2001; Langdell, 1981).

In contrast with facial emotion recognition in autism that has received much attention in the literature (e.g. Jones et al., 2011; Vannetzel, Chaby, Cautru, Cohen, & Plaza, 2011), the study of facial emotion production in children with ASD is not well documented (Trevisan, Hoskyn, & Birmingham, 2018). Compared to TD children or to children with intellectual disability, it seems that no difference with ASD children exists in FE based on blind naïve judging during a task of storytelling, neither in terms of quantity or quality (Grossman, Edelson, & Tager-Flusberg, 2013). Similarly, the number of positive FE in children with ASD versus children with ID and TD children is similar between groups during an interaction with an experimenter (Yirmiya, Kasari, Sigman, & Mundy, 1989). However, children with ASD show significantly more negative FE than the 2 other groups (Yirmiya, Kasari, Sigman, & Mundy, 1989). Also the quality of FE in ASD appears more awkward (Grossman et al., 2013), more ambiguous (Yirmiya et al., 1989) and bizarre/mechanical than that of TD children or children with intellectual disability (Loveland et al., 1994). In terms of emotion subtypes, children with ASD produce sadness particularly worse than TD children (Volker, Lopata, Smith, & Thomeer, 2009).

Interestingly, the type of tasks and the context of production (spontaneous production vs production under request) could have an impact on FE (Trevisan,

Hoskyn, & Birmingham, 2018). FE seems to be better produced when children with ASD are instructed to produce them than in spontaneous condition (for a review, see Senju, 2013); e.g. evoked FE is described as more natural than posed facial expression (Faso, Sasson, & Pinkham, 2015).

#### **1.2. Teaching Facial Emotion in Autism**

The way people with ASD produce FE can impact their integration in society (Grossman et al., 2013). That's why it is essential to teach them how to produce correct FE adapted to the social context and interaction. Recently, a lot of researchers have been using Information Communication Technologies (ICTs) for this purpose. ICT tools offer many advantages in working with children with ASD on social skills: 1) they usually enjoy playing video games and are interested in ICTs (Boucenna et al., 2014); 2) their predictability make them reassuring (Mitchell, Parsons, & Leonard, 2007); 3) ICT tools allow creating environments close to real life but keeping the child in a protected area (Josman, Ben-Chaim, Friedrich, & Weiss, 2008). To date, many different supports exist as tabletouch, CAVE, computers, robots, tabletop format...

From all ICT tools, serious games (SG) are the most promising to teach children with ASD (Grynszpan, Weiss, Perez-Diaz, & Gal, 2014). SG could be described as "*digital games and equipment with an agenda of educational design and beyond entertainment*" (Park, Abirached, & Zhang, 2012). In a recent literature review, we found 31 SG aiming to train social skills in subject with autism (Grossard et al., 2017). In this review, we explored two aspects of this SG: methodology and playability. We observed that the playability was often forgotten and that most games do not involve the characteristics of video games as described by Yusoff (2010). The implication of the player in the game is often limited, as the possibility to personalize some aspects of it. This is unfortunate as playability allows the player to be more motivated during game sessions and so more available to learn new skills. Also, visual aspects of the games are not well described although it is an important part of a video game allowing players' comprehension and progression in the game (Yusoff, 2010).

Among the 31 games we found, 16 games targeted facial emotion but only 4 were addressing facial expression production: CopyMe (Tan, Harrold, & Rosser, 2013), LifeIsGame (Fernandes, Alves, Miranda, Queirós, & Orvalho, 2011), SmileMaze (Cockburn et al., 2008) and the framework of game described in Park et al. (2012). The first 3 games have been computed including a real time facial expression analysis. However, none offers a qualitative judgment on player's FE nor proposes helping the player to adjust his/her FE when needed. CopyMe offers a feedback of player's face but the face is covered by markers of facial motion capture system. LifeIsGame proposes a feedback thanks to an avatar that reproduced player's FE.

To face this lack of smooth and transparent adaptive feedbacks on FE during gaming, we developed the JEMImE software that is a SG targeting FE in context.

The specificity of JEMImE SG is that it is able to give a feedback on the FE quality of the child-player in real time thanks to a system of gauges.

This study pursues two aims: 1) to describe the game in terms of integration, playability and ergonomics through a study of usability on a sample of children with ASD aged between 6 and 12 years. We expected a satisfaction rated good/very good for players and 2) to evaluate the feasibility of integrating a visual feedback in real time to support facial production in children with ASD. We expected that children will adapt their productions according to feedbacks provided by the algorithm.

# 2. Materials and Methods

#### 2.1. Participants

Twenty-five children were recruited in the Autism Resource Center PACA in Nice (N = 15) and in the Child and Adolescent Psychiatry Department of the Pitié-Salpêtrière Hospital in Paris (N = 10) thanks to purposeful sampling. Two participants were excluded because of an error during data recording. The final sample included 17 boys and 6 girls, aged 6 to 12 years (mean = 10.5 years, SD = 1.5). All participants received a diagnosis of ASD (DSM-IV-R criteria for ASD) after clinical assessment and the use of one of the following instrument: the Autism Diagnostic Interview-Revised (ADI-R) or the Autism Diagnostic Observation Schedule (ADOS). The cognitive level was assessed using the Wechsler Intelligence Scale for Children-4<sup>th</sup> version (WISC IV). The mean IQ was 88.57 (SD = 16.61). All were included in schools with specific adaptation. One participant was in 1<sup>st</sup> grade, 4 were in 3<sup>rd</sup> grade, 4 in 4<sup>th</sup> grade, 5 in 5<sup>th</sup> grade, 6 in 6<sup>th</sup> grade and 3 in 7<sup>th</sup> grade. Informed consent was obtained from all participants and their parents prior to participation. All procedures were approved by the Ethical Committee of the principal investigator (Comité de Protection des Personnes Sud Méditerranée V: reference number 15.071).

# **2.2. JEMImE Description**

JEMIME is a video game environment designed to teach children with ASD how to produce adapted facial expression (joy, anger or sadness) according to a social context. The proposal of adding a feedback regarding FE follows a clinical study conducted with a sister game JeSTIMULE based on the same gaming principles but without feedback (Serret et al., 2014). The game is divided into two phases: the training and playing phases. Before starting to play the game, each player has to create a profile, where he/she has to fill several information such as his/her age, his/her name, his/her grade and his/her ability to read or not. These informations can also be completed by a therapist or a parent. As recommended by Yusoff (2010) we offered the player different possibilities of personalization: the player has to build his/her avatar that can be either a preconceived character or the player can build his/her own character by choosing clothes and accessories as glasses or bag (Figure 1(a)). To finish, the player has to choose the companion who will give him/her the different instructions during the game (Figure 1(b)).

During the training phase, players are trained to produce adapted FE through four games: two imitation games (**Figure 2**) and two emotional production games on request (**Figure 3**). Each type of game (imitation or production on request) has two versions: the first less confusing without background images causing emotion; the second with a social context. Different feedbacks are given to the child player during the training phase about the adequacy of his production:



Figure 1. (a) Example of a screen during the creation of the player's avatar; (b) The player can choose between a robot or a monster to be his/her companion.



Figure 2. Example of screens during the imitation task without social context (a) and with social context (b).





1) a first visual feedback in real time, the child visualizing himself producing emotion; 2) a second visual feedback from gauges of colors (a color for each of the three emotions) that go up and down depending on the quality of the facial expression; 3) a reinforcing feedback consisting in a rate of success that is associated with coin gains when a specific emotion is correctly produced based on the algorithm assessment. In addition, another gauge indicates the time left to produce the facial expression to be adapted with the child's performance considering that high speed gaming is more difficult than low speed gaming. The adult (the therapist or the parent) can modify the conditions of the game as the number of presentations of each emotion, the number of successes needed to complete each emotion training, the percentage of recognizing score needed for the algorithm to validate the production of each emotion and the time given to the player to produce the right facial expression. During the training phase, the 4 games were always presented in the same order: imitation without social context, imitation with social context, production on request without social context and production on request with social context.

During the playing phase, players control their avatar in a virtual environment with three social contexts: a birthday party, an exchange with a soccer ball, and a game in which the child is asked to inflate a balloon (**Figure 4**). To involve the player in the game, the move of the avatar is totally free and the social contexts can be played in any order.

In the current beta-version of the game, we computed 2 different scenarios for each social context, one positive and one negative, in order to avoid that children think that one social context bring always the same emotion. The children must spontaneously produce the facial expressions adapted to each context, thus making it possible to put into practice the learning of the training phase. As in the previous phase, performance feedback is offered to the child as well as visual feedbacks (video and gauges) and reward (a coin and for some version of the scenario, another objet as a cake in the birthday party) (**Figure 5**). Similarly, a



**Figure 4.** Example of screen showing the virtual environment of a birthday party during the playing phase.



Figure 5. Reward after correctly completing the positive version of the birthday party scenario.

gauge for time is also included. The "playful" feature of serious games has been materialized by a collection of candies hidden in the city to encourage exploration, animations of objects in the environment and an interactive map of the city to facilitate travel.

#### 2.3. Computing the FE Classifier within the JEMImE Architecture

In terms of architecture, we pursued two aims 1) to make JEMImE compatible with an online platform in order to increase its diffusion and 2) to incorporate the FE classifier in a way that allows a feedback in real time without bringing freezes or latencies during the game. The current prototype of JEMImE allows the child to play alone in the therapist's office or with the parent at home. The diffusion of this Serious Game on the website Curapy.com would allow a simplified accessibility and large scale diffusion. Curapy.com is a platform that has been opened to health professionals since 2017. It has already proven its efficiency in terms of data integrity, privacy, restricted data access and seamless communication. It will record the data of the sessions performed by the child while playing JEMImE, as well as its scores (reaction time, success rate, duration of play per session and in total, frequency of games selected). In this way, therapists will be able to regularly monitor the use and the progress in children's games, identify their difficulties, adapt their therapeutic intervention and guide parents via Curapy.com. The serious game could then be used by both therapists at the hospital and by families at home. Natural environment at home and dialogue between parents and therapists are key factors of attractiveness. This new modality offers a transition from a traditional work mode based in a medical environment to a home-based child focused approach, which is non-invasive and innovative (Bono et al., 2016).

We used our method, namely the Pairwise Conditional Random Forests (PCRF), in order to assess the facial expressions (Dapogny, Bailly, & Dubuisson, 2015). It consists in extracting heterogeneous derivative features (e.g. facial land-

mark movements or texture variations) upon pairs of images. These features are used as input of a random forest that is trained to recognize emotion transitions from previous frames to the current frame. The robustness is ensured by conditioning forests on the expression label of the first frame of each pair and by averaging pairwise decision upon time. PCRF was trained on the state-of-the-art BU-4DFE database of adults producing facial emotions (Yin, Sun, Worm, & Reale, 2008). The JEMIME's environment is developed in C# with the 3D game engine Unity for computer and playable with a gamepad or a keyboard. The FE classifier provides face analysis systems developed in C++ that operate in real-time. The communication between the FE classifier and the serious game environment is easily established using the Publisher/Subscriber pattern from the middleware ZeroMQ with zero latency and fully transparent for the players who have an immediate feedback on their FE productions.

#### 2.4. Assessments

One of the objectives of the current study was to evaluate the course and the adaptation of children with ASD playing JEMImE and to assess whether children and professionals appreciate it. We used three measures. First, we computed the course of each child's performance during the training phase thanks to the software records. We expected children to increase their knowledge (meaning their facial expression accuracy) during the training phase. Second, we created an ergonomic and satisfaction questionnaire to be completed by the child after the game session (see Annex). Third, we also created a child behavior observation grid to be completed by two professionals during the game sessions. The ergonomics and satisfaction questionnaire contains 26 affirmative sentences (e.g. "It was easy to choose my companion") concerning the different parts of the software (creation of the avatar, choice of the companion, learning phase, training phase) and also on the software in a global way (see Annex). Among the 26 affirmative sentences, 8 sentences were used to evaluate the appreciation of the aspect of the software (menus, buttons, avatars) (e.g. "I liked the visual aspect of my avatar which I created"). Seven sentences were used to assess the efficiency and the ergonomics (the ease of creating the avatar, understanding the role of the color gauges or moving in the virtual environment) (e.g. "It was easy to create my avatar"). Finally, 11 sentences were used to assess the child's satisfaction (using the software, creating an avatar, choosing a companion, producing emotional expressions and collecting coins or sweets) (e.g. "I liked the part where I had to imitate the characters"). The questionnaire was completed after the session of using the software. After each statement, the child had to circle his/her level of agreement with the proposal on a 4-point Likert scale ranging from "strongly disagree" to "strongly agree", illustrated by smileys corresponding to each level. The child behavior observation grid was developed to assess the players' behavior during the gaming sessions during the 3 different stages: the beginning phase with the choice of the avatar and the companion, the training and the playing phase. A part of the grid is dedicated to reporting response statistics (Number of trials/errors) recorded in the game. In the part concerning the choice of the avatar and the companion, the grid makes it possible to note the time put by the subject to choose his avatar and his companion, the type and their characteristics. Regarding the training phase, different informations are recorded about the duration of this phase, the child's reaction to gains and failures (type of emotion, type of behavior, intensity of the reaction on an analogic scale from 0 to 10) and the active or non-active use of the color gauges. Regarding the playing phase, the following information are collected: the total duration of this phase, the number of times each scenario was seen, the number of candies collected, the active attitude of the child in this phase, the reactions during wins and failures and the active use of gauges as in the previous phase.

The second objective was to evaluate the possibility and interests of integrating an algorithm that will give feedbacks in real time to the player on their productions. For that, we recorded the score of the children during each mini game. We expected their results to show a significant improvement between consecutive sessions meaning that players learned to do better FE thanks to the automated feedbacks and that they adapted their productions thanks to this feedback.

#### 2.5. Procedure

The evaluation was conducted during a single session lasting approximately 75 minutes and was conducted as follows. The child was invited to sit in front of a computer, connected to a joystick and a camera placed above the computer screen allowing the child to see him/herself on video. After this phase of installation, the child began to manipulate the software, accompanied by a therapist. He realized the different games of the training phase then visualized all the scenarios and social contexts of the gaming phase, guided by the therapist. During this phase of use (60 minutes), video recorded by an outdoor camera, a second therapist present in the room observed the behavior of the child and filled the observation grid. After the training phase, the child was asked to complete the first part of the game. Once the protocol was completed and the child has left the room, the therapist completed the observation grid behavior of the child with his/her colleague using if necessary the film recorded from the meeting.

#### 3. Results

During the training phase, we expected children with ASD to increase the quality of their FE thanks to the FE algorithm classifier that was integrated in JE-MIME. Since we know that several factors (age, gender, game timing (early games as opposed to last games) or emotion subtypes) can influence facial emotion production, we performed a multivariate model. As our variables did not follow a normal distribution, we used a logistic mixed effect model in R software version 3.4.0 (lme4 package) with the formula: issue—Age + Gender + Game timing + Emotion + (1|id\_patient). The results are presented in **Table 1**. We found no effect of age (p = 0.17) or gender (p = 0.16), but we did found a significant effect of the targeted emotion (p < 0.001) and game timing (p < 0.001).

Then, we compared 2 by 2, each of our modalities within the variable that showed a significant effect (**Table 2**). We used a multivariate analysis with the software R with the formula: Success—Age + Gender + Game timing + Emotion + (1|id\_patient). The multivariate analysis shows a significant effect of the targeted emotion on the quality of the children's FE. Happiness is better produced than sadness (p < 0.01) and anger is better produced than sadness (p < 0.001). However, anger and happiness are equally produced (p = 0.46). We also found a significant effect of the game subtype on the quality of children's FE. The game 4 is better succeeded than the game 1 and the game 2 (p < 0.001); the game 3 is better succeeded than the game 1 (p < 0.001) and the game 2 (p = 0.017); and the game 2 is better to succeed than the game 1 (p < 0.01).

Regarding the ergonomics and satisfaction questionnaire, **Figure 6** presents the histogram of children's responses by general themes. Overall, most of the children agreed or very much agreed with the statements. Concerning the visual aspect, 52.60% of the children completely appreciated it. Concerning the ergonomics, 64.28% of the children totally agreed our choices. In a more global way,

**Table 1.** Effect of age, gender, game subtypes and targeted emotion on the quality of the players FE recognized by the algorithm during the training phase (N = 23).

Factors	Degree of liberty	AIC	LRT	р
Age	1	1170.8	1.891	0.169
Gender	1	1170.9	1.972	0.1603
Game timing	3	1235.7	70.734	< 0.001
Emotion	2	1363.1	196.194	< 0.001

**Table 2.** Comparison 2 by 2of each of our modalities within the variables that showed a significant effect during the training phase in children with ASD (N = 23).

Variable	Coefficient	Standard error	р
Anger vs Sadness	2.006	0.195	<0.001
Happiness vs Sadness	2.181	0.204	<0.001
Happiness vs Anger	0.174	0.237	0.463
Game 1 vs Game 2	0.812	0.194	<0.001
Game 1 vs Game 3	1.348	0.213	<0.001
Game 1 vs Game 4	1.627	0.231	<0.001
Game 2 vs Game 3	0.535	0.224	0.017
Game 2 vs Game 4	0.814	0.241	<0.001
Game 3 vs Game 4	0.279	0.254	0.272



Figure 6. Histogram of children's responses to the satisfaction questionnaire (N = 23).

nearly 60% of the children were completely satisfied by JEMImE. All the children (100%) found easy to use the software overall.

The observational grid gathered pretty concordant positive results in term of playability. During the personalization, 91% of the children wished to personalize their avatar and did it quickly (less than two minutes) and easily (100% of ease for the creation of the avatar). They also selected quickly (7 seconds) and easily their companion (82% ease of companion selection). During the training phase, it was observed that 100% of the participants actively used the software returns (color gauges and video feedback) in real time. The role of these returns was easily understood by most children (83% for the functioning of the gauges and 78% for the video support). During the playing phase, all the children easily circulated in the virtual environment. The observations showed that the map was also used by 92% of the children to find their way or find hidden candies.

The majority of children (94%) appeared satisfied with the use and overall look of the software. During the personalization, 93% of children enjoyed being able to choose and customize their avatar and companion in the game (88%). The majority opted for the choice of the robot as a companion. The appearance of the avatars that the children had to imitate or those present in the virtual environment pleased 77% of the children. During the training phase, the games and the contextual images used for the production of emotion, were generally appreciated (84% of satisfied children). During the playing phase, the visual aspect of this phase, navigation in the virtual environment and production in context satisfied the majority of children (94%). Finally, the possibility of obtaining gains in case of success was appreciated by 96% of children.

# 4. Discussion

The aims of our study were: 1) to evaluate the interest of a visual feedback on children's FE production in JEMImE by measuring their performance during the 4 steps/games of the training phase and 2) to assess children's satisfaction concerning JEMImE thanks to an observation grid and a questionnaire.

The real time FE analysis provided by the FE algorithm gives an immediate feedback on children's FE quality without latency. Concerning children's progression during training, we found 2 significant effects. The first was the effect of the target emotion on the quality of their production: sadness was significantly worse produced than happiness and anger. These results are congruous with the literature, as sadness seems to be particularly difficult to be correctly produced (Volker et al., 2009). The algorithm integrated in the game reproduces the impression of human judges, and appears to correctly interpret their facial expressions. The second significant effect was the effect of game timing. Indeed, the first game presented was less successful than all the other games and the best successful game was the last presented. It appears that children increased their productions through the 4 games of the training phase and learned how to adapt their production to fit with the expectation of the algorithm. Even if we can't conclude that children learn how to produce better facial expressions regarding quality, we can observe that they adapt their FE to respond more adequately to the feedbacks given inside the serious game. So, the facial feedback in real time but also the gauges that give them an appreciation on the quality of their production really helped the children to correct themselves. However, we also have to temper our results, as the effect of the task could affect the results of children and favorize the on request task (game 3 and 4) compare to the imitation task (game 1 and 2). Yet, we noted a significant effect between the game 1 and 2 with results that increased in game 2 compared to game 1. This suggests that children can adapt their facial expressions by training with automatic feedbacks whatever the task (imitation as opposed to on request).

We found no significant effect of gender but also no effect of age on children's production evaluated by the algorithm. Concerning the effect of age, we do not have strong data in literature either, however it seems that facial expressions production quality should increase with age (Trevisan, Hoskyn, & Birmingham, 2018). In our study, the lack of effect could be due to the small number of children we recruited and the lack of statistical power. Moreover, we did not take into account IQ, who could modulate the effect of age (Trevisan, Hoskyn, & Birmingham, 2018). However, for the lack of effect of age on the production, we could also suppose that children do not learn spontaneously how to correctly produce facial expressions and so do not increase the quality of their productions with age. This observation supports the necessity to explicit facial expressions and work on it in order to allow children with ASD to adapt their facial expressions in a social context.

The second aim of our study was to assess the playability and the satisfaction of children playing JEMIME. Some of studies do not develop theses aspect, however it is an important part to bring motivation to the player and facilitate the immersion in the virtual environment (Grossard et al., 2017; Yusoff, 2010). That's why we gave a specific attention to the items usually find in commercial video games as for example reward, fun animations, offering help during the game, or fluidity of the game. All these characteristics are essential to allow the immersion and maintain attention.

All the children who came to try the game were highly satisfied with the game. They showed a special interest in building their own character and were happy to play with it in the playing phase. Concerning the visual aspect, they globally appreciated the choice of design and colors. All the children were able to play using the joystick or the keyboard and found easy to move and explore the environment. However, the participants played only one session of the game. They would be naturally excited to use a new game. Therefore their responses would tend to be more positive. We should propose more sessions to see if the level of positivity remains after that initial session. This was not possible as JEMIME is only a prototype and that a lot more scenarios must be developed. The players generally complain about the small environment with the impossibility to go inside the buildings. They also suggested adding more people, cars and objects in the town, as now there are only 3 groups of people placed where the social scenes can be played.

# **5. Limitations**

However, our study presents some limitations. 1) We recruited a small sample of children with ASD, which could explain why we did not find an effect of age on the quality of children's productions. 2) We did not assess the effect of IQ on their productions. 3) The game allows them to play just one hour, which is not enough to ensure that their motivation and satisfaction of the game would last in time.

#### 6. Conclusion

JEMIME prototype was used with 23 children with ASD in order to evaluate the usability of the game with this population but also to collect their experience regarding the ergonomics, the playability and the visual aspect of the serious game. It appears that JEMIME is adapted to children with ASD. The integration of the algorithm for facial expression recognition is perfectly done and does not bring latencies during gaming. The feedback in real time allows the children to modify their facial expression during a play session to fit with the expectations of the game. They correctly used the gauges to evaluate their productions. In terms of satisfaction, all the children were able to play and were very satisfied of their experience in the game. This paper contributes to support that information and communication technologies are adapted to children with ASD. Specifically, the current study shows that the integration of a facial expressions algorithm allows to work on facial emotion production.

Future works include: 1) To improve the algorithm by using a dataset of children FE to train the automated classifier (Grossard et al., 2018) as we used a FE classifier trained on adult FE databases for this first study. We believe that taking into account the morphological difference of children during the emotion

recognition process is important. 2) To develop new scenarios in the game to have enough social situations to work on a longer exposure to JEMIME. We could propose a more clinically grounded pilot study in order to evaluate the course of the children concerning their facial expression production in context after playing JEMIME for several weeks. This study could be favored by the integration of JEMIME on an online platform (Curapy.com) that will make it easily accessible to therapists and families.

# Acknowledgements

The study was supported by the *Agence Nationale pour la Recherche* (ANR) within the program CONTINT (JEMIME, n° ANR-13-CORD-0004).

# **Conflicts of Interest**

Pierre Foulon is general director of e-Genious, a private company that develops serious games for health purposes. Other authors have no conflict of interest to declare.

## References

- Bennett, D. S., Bendersky, M., & Lewis, M. (2005). Does the Organization of Emotional Expression Change over Time? Facial Expressivity from 4 to 12 Months. *Infancy, 8*, 167-187. https://doi.org/10.1207/s15327078in0802\_4
- Bono, V., Narzisi, A., Jouen, A. L., Tilmont, E., Hommel, S., Jamal, W., Chetouani, M. et al. (2016). GOLIAH: A Gaming Platform for Home-Based Intervention in Autism-Principles and Design. *Frontiers in Psychiatry*, 7, 70. https://doi.org/10.3389/fpsyt.2016.00070
- Boucenna, S., Narzisi, A., Tilmont, E., Muratori, F., Pioggia, G., Cohen, D., & Chetouani, M. (2014). Interactive Technologies for Autistic Children: A Review. *Cognitive Computation, 6*, 722-740. <u>https://doi.org/10.1007/s12559-014-9276-x</u>
- Brun, P. (2001). Psychopathologie de l'émotion chez l'enfant: L'importance des données développementales typiques. *Enfance*, 53, 281-291. https://doi.org/10.3917/enf.533.0281
- Chaplin, T. M., & Aldao, A. (2013). Gender Differences in Emotion Expression in Children: A Meta-Analytic Review. *Psychological Bulletin, 139*, 735. <u>https://doi.org/10.1037/a0030737</u>
- Cockburn, J., Bartlett, M., Tanaka, J., Movellan, J., Pierce, M., & Schultz, R. (2008). Smilemaze: A Tutoring System in Real-Time Facial Expression Perception and Production in Children with Autism Spectrum Disorder. In ECAG 2008 Workshop Facial and Bodily Expressions for Control and Adaptation of Games (Vol. 3). Amsterdam.
- Dapogny, A., Bailly, K., & Dubuisson, S. (2015). Pairwise Conditional Random Forests for Facial Expression Recognition. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 3783-3791). https://doi.org/10.1109/ICCV.2015.431
- Ekman, P., Roper, G., & Hager, J. C. (1980). Deliberate Facial Movement. *Child Development*, 886-891. https://doi.org/10.2307/1129478
- Faso, D. J., Sasson, N. J., & Pinkham, A. E. (2015). Evaluating Posed and Evoked Facial Expressions of Emotion from Adults with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 45, 75-89.

#### https://doi.org/10.1007/s10803-014-2194-7

- Fernandes, T., Alves, S., Miranda, J., Queirós, C., & Orvalho, V. (2011). LIFEisGAME: A Facial Character Animation System to Help Recognize Facial Expressions. In *International Conference on Enterprise Information Systems* (pp. 423-432). Berlin, Heidelberg: Springer. <u>https://doi.org/10.1007/978-3-642-24352-3\_44</u>
- Grossard, C., Chaby, L., Hun, S., Pellerin, H., Bourgeois, J., Dapogny, A., Chen, L. et al. (2018). Children Facial Expression Production: Influence of Age, Gender, Emotion Subtype, Elicitation Condition and Culture. *Frontiers in Psychology*, *9*, 446. https://doi.org/10.3389/fpsyg.2018.00446
- Grossard, C., Grynspan, O., Serret, S., Jouen, A. L., Bailly, K., & Cohen, D. (2017). Serious Games to Teach Social Interactions and Emotions to Individuals with Autism Spectrum Disorders (ASD). *Computers & Education*, *113*, 195-211. <u>https://doi.org/10.1016/j.compedu.2017.05.002</u>
- Grossman, R. B., Edelson, L. R., & Tager-Flusberg, H. (2013). Emotional Facial and Vocal Expressions during Story Retelling by Children and Adolescents with High-Functioning Autism. *Journal of Speech, Language, and Hearing Research, 56*, 1035-1044. https://doi.org/10.1044/1092-4388(2012/12-0067)
- Grynszpan, O., Weiss, P. L., Perez-Diaz, F., & Gal, E. (2014). Innovative Technology-Based Interventions for Autism Spectrum Disorders: A Meta-Analysis. *Autism, 18,* 346-361. https://doi.org/10.1177/1362361313476767
- Jones, C. R., Pickles, A., Falcaro, M., Marsden, A. J., Happé, F., Scott, S. K., Simonoff, E. et al. (2011). A Multimodal Approach to Emotion Recognition Ability in Autism Spectrum Disorders. *Journal of Child Psychology and Psychiatry*, 52, 275-285. https://doi.org/10.1111/j.1469-7610.2010.02328.x
- Josman, N., Ben-Chaim, H. M., Friedrich, S., & Weiss, P. L. (2008). Effectiveness of Virtual Reality for Teaching Street-Crossing Skills to Children and Adolescents with Autism. *International Journal on Disability and Human Development*, 7, 49-56. https://doi.org/10.1515/IJDHD.2008.7.1.49
- Langdell, T. (1981). *Face Perception: An Approach to the Study of Autism.* Doctoral Dissertation, London: University of London. <u>http://discovery.ucl.ac.uk/1317720/</u>
- Loveland, K. A., Tunali-Kotoski, B., Pearson, D. A., Brelsford, K. A., Ortegon, J., & Chen, R. (1994). Imitation and Expression of Facial Affect in Autism. *Development and Psychopathology*, *6*, 433-444. <u>https://doi.org/10.1017/S0954579400006039</u>
- Mitchell, P., Parsons, S., & Leonard, A. (2007). Using Virtual Environments for Teaching Social Understanding to 6 Adolescents with Autistic Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 37, 589-600. https://doi.org/10.1007/s10803-006-0189-8
- Oster, H. (2005). The Repertoire of Infant Facial Expressions: An Ontogenetic Perspective. *Emotional Development*, 261-292.
- Park, J. H., Abirached, B., & Zhang, Y. (2012). A Framework for Designing Assistive Technologies for Teaching Children with ASDs Emotions. In *CHI'12 Extended Ab*stracts on Human Factors in Computing Systems (pp. 2423-2428). New York: ACM. https://doi.org/10.1145/2212776.2223813
- Senju, A. (2013). Atypical Development of Spontaneous Social Cognition in Autism Spectrum Disorders. *Brain and Development*, 35, 96-101. <u>https://doi.org/10.1016/j.braindev.2012.08.002</u>
- Serret, S., Hun, S., Iakimova, G., Lozada, J., Anastassova, M., Santos, A., Askenazy, F. et al. (2014). Facing the Challenge of Teaching Emotions to Individuals with Low- and

High-Functioning Autism Using a New Serious Game: A Pilot Study. *Molecular Autism, 5,* 37. https://doi.org/10.1186/2040-2392-5-37

- Sullivan, M. W., & Lewis, M. (2003). Emotional Expressions of Young Infants and Children: A Practitioner's Primer. *Infants & Young Children, 16*, 120-142. https://doi.org/10.1097/00001163-200304000-00005
- Tan, C. T., Harrold, N., & Rosser, D. (2013). Can you CopyMe?: An Expression Mimicking Serious Game. In SIGGRAPH Asia 2013 Symposium on Mobile Graphics and Interactive Applications (p. 73). ACM. https://doi.org/10.1145/2543651.2543687
- Trevisan, D. A., Hoskyn, M., & Birmingham, E. (2018). Facial Expression Production in Autism: A Meta-Analysis. *Autism Research*, 11, 1586-1601. https://doi.org/10.1002/aur.2037
- Vannetzel, L., Chaby, L., Cautru, F., Cohen, D., & Plaza, M. (2011). Neutral versus Emotional Human Stimuli Processing in Children with Pervasive Developmental Disorders Not Otherwise Specified. *Research in Autism Spectrum Disorders*, *5*, 775-783. https://doi.org/10.1016/j.rasd.2010.09.005
- Volker, M. A., Lopata, C., Smith, D. A., & Thomeer, M. L. (2009). Facial Encoding of Children with High-Functioning Autism Spectrum Disorders. *Focus on Autism and Other Developmental Disabilities, 24*, 195-204. https://doi.org/10.1177/1088357609347325
- Yin, L., Sun, X. C. Y., Worm, T., & Reale, M. (2008). A High-Resolution 3D Dynamic Facial Expression Database. In *IEEE International Conference on Automatic Face and Gesture Recognition* (Vol. 126). Amsterdam, The Netherlands. https://doi.org/10.1109/AFGR.2008.4813324
- Yirmiya, N., Kasari, C., Sigman, M., & Mundy, P. (1989). Facial Expressions of Affect in Autistic, Mentally Retarded and Normal Children. *Journal of Child Psychology and Psychiatry*, 30, 725-735. <u>https://doi.org/10.1111/j.1469-7610.1989.tb00785.x</u>
- Yusoff, A. (2010). A Conceptual Framework for Serious Games and Its Validation. Doctoral Dissertation, Southampton: University of Southampton. <u>https://eprints.soton.ac.uk/267690/1/3711a021.pdf</u> <u>https://doi.org/10.1109/ICALT.2009.19</u>

# Annex: Translation of the Questionnaire Given to Children to Assess the Serious Game JEMImE

This questionnaire will give you some affirmation sentences about the game JE-MImE you just played. You have to indicate how much you agree with each of these sentences, by choosing among the answer:



There is no good or bad answer: just give YOUR advice on the game.

#### When I created my avatar:

I like the visual aspect of my avatar when I created it.

Totally disagree	Rather disagree	Rather agree	Totally agree
	$\left(\begin{array}{c} \cdot \\ \end{array}\right)$	$(\cdot \cdot \cdot)$	$\overline{\bigcirc}$

It was easy to create my avatar.

Totally disagree	Rather disagree	Rather agree	Totally agree
$(\mathbf{\dot{\cdot}},\mathbf{\dot{\cdot}})$	$\left(\begin{array}{c} \cdot \\ \end{array}\right)$	$(\cdot \cdot)$	

I liked to have the possibility to create my avatar as I wanted it.

Totally disagree	Rather disagree	Rather agree	Totally agree	
$(\mathbf{x}, \mathbf{x})$	$\left(\begin{array}{c} \cdot \\ \hline \end{array}\right)$	$(\cdot \cdot)$		

## When I chose a companion (Robot/monster):

I liked the visual aspect of the companion I chose.

Totally disagree	Rather disagree	Rather agree	Totally agree
(	$\left(\begin{array}{c} \cdot \\ \hline \end{array}\right)$	$(\cdot \cdot)$	

It was easy to choose a companion.



I liked to have the possibility to choose the visual aspect of the companion.

Totally disagree	Rather disagree	Rather agree	Totally agree
	$(\cdot, \cdot)$	$(\cdot \cdot)$	$(\cdot, \cdot)$

# In the first part, when I had to imitate the avatars:

I liked the visual aspect of the avatars who showed the emotions.

Totally disagree	Rather disagree	Rather agree	Totally agree
$(\mathbf{x}, \mathbf{x})$	$\left(\begin{array}{c} \cdot \\ \end{array}\right)$	$(\cdot \cdot)$	

I liked the visual aspect of the images presented before some avatars.

Totally disagree	Rather disagree	Rather agree	Totally agree
× <	$\left(\begin{array}{c} \cdot \\ \end{array}\right)$	$(\cdot \cdot)$	

It was easy to understand the color gauges when I had to produce the emotion with my face.

Totally disagree	Rather disagree	Rather agree	Totally agree
$(\mathbf{x},\mathbf{x})$	$\left(\begin{array}{c} \cdot \\ \end{array}\right)$	$(\cdot \cdot)$	$(\cdot, \cdot)$

It helped me to see me in video to well produce the emotion with my face.

Totally disagree	Rather disagree	Rather agree	Totally agree
$(\mathbf{x}, \mathbf{x})$	$\left(\begin{array}{c} \cdot \\ \end{array}\right)$	$(\cdot \cdot)$	$\overline{\bigcirc}$

It was interesting to win coins.

Totally disagree	Rather disagree	Rather agree	Totally agree
$(\mathbf{x},\mathbf{x})$	$\left(\begin{array}{c} \cdot \\ \end{array}\right)$	$(\cdot \cdot)$	$\overline{\bigcirc}$

I liked when I had to imitate the avatars.

Totally disagree	Rather disagree	Rather agree	Totally agree
() · · ·	$\left(\begin{array}{c} \cdot \\ \end{array}\right)$	$(\cdot \cdot \cdot)$	$\overline{\bigcirc}$

I liked when I had to produce an emotion in function of the word displayed on the screen.

Totally disagree	Rather disagree	Rather agree	Totally agree
$\sim$	$\left(\begin{array}{c} \cdot \\ \end{array}\right)$	$(\cdot \cdot)$	

#### In the second part, when I was in the city:

I liked the visual aspect of the city.

\_

Totally disagree	Rather disagree	Rather agree	Totally agree
× <		$(\cdot \cdot)$	

#### I liked the visual aspect of the avtars in the town.

Totally disagree	Rather disagree	Rather agree	Totally agree
	$\left(\begin{array}{c} \cdot \\ \end{array}\right)$	$(\cdot \cdot)$	
was easy to move	in the city.		
Totally diagona	Dathan diaganag	Dath an agree	Totally agree

Totally disagree	Rather disagree	Rather agree	Totally agree
× *	$\left(\begin{array}{c} \cdot \\ \hline \end{array}\right)$	$(\cdot \cdot)$	

It was interesting to win gift.

Totally disagree	Rather disagree	Rather agree	Totally agree
$(\mathbf{x}, \mathbf{x})$	$(\cdot \cdot)$	$(\cdot \cdot)$	

It was interesting to seek for candies in the city.

Totally disagree	Rather disagree	Rather agree	Totally agree
() <	$(\cdot \cdot \cdot)$	$(\cdot \cdot)$	

I liked to move around the city.

Totally disagree	Rather disagree	Rather agree	Totally agree
	$(\cdot \cdot \cdot)$	$(\cdot \cdot)$	$(\cdot \cdot)$

I liked when I had to do an emotion in function of what happen with the other avatars.

Totally disagree	Rather disagree	Rather agree	Totally agree
() · · ·	$\left(\begin{array}{c} \cdot \\ \end{array}\right)$	$(\cdot \cdot)$	$\overline{\bigcirc}$

# About the game JEMImE in general:

I liked the visual aspect of the game JEMImE.

Totally disagree	Rather disagree	Rather agree	Totally agree
$(\mathbf{x}, \mathbf{x})$	$(\cdot \cdot)$	$(\cdot \cdot)$	$\overline{\bigcirc}$

The game JEMImE was easy to use.

Totally disagree	Rather disagree	Rather agree	Totally agree
() · · ·	$(\cdot \cdot)$	$(\cdot \cdot)$	$\overline{\bigcirc}$

# I liked the visual aspect of the menu and the buttons in the game JEMImE.

Totally disagree	Rather disagree	Rather agree	Totally agree
$\sim$	$\left(\begin{array}{c} \cdot \\ \end{array}\right)$	$(\cdot \cdot)$	

# The menu and the buttons were easy to use.

Totally disagree	Rather disagree	Rather agree	Totally agree
() <	$(\cdot \cdot)$	$(\cdot \cdot)$	

# I liked to use the game JEMImE.

Totally disagree	Rather disagree	Rather agree	Totally agree
× <	$(\cdot \cdot)$	$(\cdot \cdot)$	$\overline{\bigcirc}$

# I'd like to play the game JEMImE at home.

Totally disagree	Rather disagree	Rather agree	Totally agree
() · · ·	$(\cdot \cdot)$	$(\cdot \cdot)$	$\overline{\bigcirc}$