Impairments in decoding facial and vocal emotional expressions in high functioning autistic adults and adolescents

Anna Esposito[1], Italia Cirillo[1], Antonietta M. Esposito[2], Leopoldina Fortunati[3], Gian Luca Foresti[3], Sergio Escalera[4], Nikolaos Bourbakis[5]

[1]Università della Campania "L. Vanvitelli" Dept. Psychology and IIASS, IT, iiass.annaesp@tin.it, lialis12@libero.it

[2]INGV, Sez. di Napoli Osservatorio Vesuviano, Napoli, IT, antonella.esposito@ov.ingv.it

[3]University of Udine, IT, {gianluca.foresti, leopoldina.fortunati}@uniud.it

[4]CVC and University of Barcelona, ES, sergio.escalera.guerrero@gmail.com

[5]Wright State University, OH, US, nikolaos.bourbakis@wright.edu

***Abstract*—The present investigation shows that gender of stimuli, age, and emotional categories affects the ability of adults and adolescent with Autistic Spectrum Conditions (ASC) to decode facial and vocal emotional expressions. A total of 60 subjects participated to the research: 15 ASC and 15 control adolescents aged between 10-14 years; and 15 ASC and 15 control young adults aged between 20-24 years. Their tasks consisted in decoding: a) 24 adults and 24 children contemporary facial emotional expressions of happiness, sadness, anger, fear, surprise, and disgust; and b) 20 adult's vocal emotional expressions of the same abovementioned emotions (except disgust). Significant differences were observed between ASC and typically developed peers. The data suggest that gender, type (voices or faces) of stimuli, and participants' age affect the emotion recognition process making difficult the definition of a common and shared pattern of emotional expression's recognition compliance among autistic and control groups. These results suggest that efficient and effective e-health technologies need to be able to learn and adapt to user individual traits and subjective needs to offer personalized assistance and support.**

I. Introduction

The implementation of socially and emotionally believable ICT (Information Communication Technologies) interfaces brought the promise to enhance quality of life in society, particularly for vulnerable people, i.e., people suffering of physical and social disorders. However, to achieve this promise it is necessary to inform the implementation on the analytics of these disorders, in order to develop effective algorithms, models and computational paradigms. Autism is among these disorders and the current investigation aims to provide such analytics. Autism spectrum conditions (ASC) manifest themselves mainly with disorders of social interaction, communication and understanding of ideas and feelings (DSM V, 2013). The symptoms are associated, in an extremely variable way, to disorders in language, motor skills, cognition (generally mental retardation) and emotions' labeling and coding. According to the current version of the Diagnostic and Statistical Manual of Mental Disorders (DSM V, 2013), ASCs are neurodevelopmental disorders, whose seriousness can be severe, medium, or mild (high functioning ASCs) . Each of these levels manifests different behavioral, social, and emotional impairments.

Research on the ASC's ability to recognize emotions has been very extensive. It was powered by the desire to understand the basic mechanisms causing these individuals' failure to appropriately label and encode their and others facial, gestural and vocal emotional expressions, in order to successfully regulate interactional exchanges. The lack of a mutual sharing of emotional meanings, associated with compromised and atypical responses to other people's emotions, are perhaps the main factors preventing mild/high functioning ASC individuals from being involved in everyday shared livings.

II. Related Work

Research investigating whether ASC individuals (mainly high functioning ASCs) show deficits in decoding facial and vocal emotional expressions of surprise, disgust, fear, sadness, anger, and happiness, produced controversial results.
As for faces, studies had shown that ASC were less accurate than typically developed (TD) individuals in correctly decoding negative emotions of anger, disgust, fear, sadness (Xavier et al. 2015, Oerlemans et al. 2014, Humphreys et al. 2007) , and even surprise (Baron-Cohen et al. 1993) as well as some positive happy derived emotions (Humphreys et al. 2007). However, these findings were not replicated by other authors which reported for ASCs similar performances of their TD peers (Jones et al. 2011). These inconsistencies were found regardless of the following factors:

a) age – either no differences or substantial improvements in emotion decoding performance were observed between high functioning adults, adolescents, and school age children with ASC (Taylor et al. 2015);

b) verbal and nonverbal IQ status – either large or no differences were found in emotion decoding accuracy between ASC individuals matched controls (Taylor et al. 2015);

c) low and high functioning ASC status – either broad deficits or no differences in emotion decoding accuracy were reported between ASC individuals and matched controls (Sucksmith et al. 2013, Taylor et al. 2015).

d) type of stimuli – no matter the stimuli were static photographs, dynamic videos, morphed or blended facial expressions, some authors reported ASC related deficits, while others reported performances equal to controls matched for developmental level and/or verbal mental age (Kirchner et al. 2011, Lerner et al. 2013, Griffiths et al. 2017, Lot et al. 2018)

e) task demand – no matter the task was labeling or matching ASC individuals showed either deficits or no deficits in correctly decoding facial emotional expressions (Taylor et al. 2015, Lot et al. 2018).

In a comprehensive review, accounting of results from 44 papers published between 1986 and 2010 (Harms et al. 2010) and a late meta -analysis by Uljarevic and Hamilton (2012) it was concluded that neither age, nor task demand, type of stimuli, and IQ scores, play a role in ASCs emotion recognition accuracy. It was concluded that "*we should rule out theories that claim a global emotion recognition difficulty is primary and universal in autism. Rather, difficulties with emotion processing must be specific to particular emotions or stimuli*" (Uljarevic & Hamilton, 2012, p.7). Recently, Lot and colleagues (2018) debated these considerations reporting significant differences in ASC's recognition of static pictures of primary vs complex emotions: primary emotions were more accurately decoded than complex ones. However, these results were observed only for individuals with severe ASC deficits leaving unclear to establish whether "*emotion recognition difficulty is primary and universal in autism*" (Uljarevic & Hamilton, 2012, p.7).

The ability of ASC individuals to decode emotional states from voices has been less investigated than faces. Nevertheless, also for this task, inconsistent findings have been observed. Some studies showed that, compared to typically individuals, ASC patients did show impairments in the recognition of emotional vocal expressions (Charbonneau et al., 2013, Oerlemans et al. 2014). However, other authors did not find such evidences (Lerner et al. 2013, Javier et al 2015), independently from types of stimuli, age, and participant's IQ.

Considering the difficulties to identify patterns of evidence to relate ASC conditions to their performances in emotion labeling tasks, the present research proposes an investigation restricted to a narrowed age group (only adolescents and young adults) of high functioning ASC patients in order to eliminate possible confounders' effects (gender, IQ, task demand, and age) . To remove emotional category complexity effects (Lot et al. 2018), only happiness, fear, anger, surprise, and sadness were considered for voices, while for faces also disgust was added. Disgust was not considered due to the difficulty to associate a unique meaning to this emotion through speech (Rozin et al. 2000). In addition to test the emotion labeling ability of ASC patients, this investigation also aims to:

a) Shed light on the inconsistencies reported in literature. All the others being equal, would the restricted aged group substantiate ASCs' differences in decoding emotional expressions with respect their TD peers?

b) Identify developmental patterns measuring maturity (adolescents vs young adults) in ASCs and TD's' emotion labeling.

c)Identify effects of the gender of stimuli, neglected in literature. Does gender play a role on the ability of ASC and TD individuals to decode emotional expressions?

III. MATERIALS AND PROCEDURES

A. Participants

The recruitment took place at the association "La Forza del Silenzio ("The strength of Silence") at Casal di Principe, Caserta, Italy. The institutional ethical committee provided the permission to conduct the experiments, considering them at no risk for patients. Consent forms were collected, attesting the voluntary acceptation of adolescent's relatives and adults. Participants were debriefed on the aims and planned procedures of the proposed experiments. Measures to protect privacy were taken according to the European General Data Protection Regulation (GDPR, https://www.eugdpr.org/) and Italian laws on privacy safeguard (D.Lgs. 196/2003).

A total of 60 subjects were recruited, 15 ASC (2 females, mean age=11.8 years, SD=±0.7 years) and 15 control (2 females, mean age=11.8 years, SD=±1 year) adolescents, and 15 ASC (2 females, mean age= 21.7 years, SD=±1.6 years) and 15 control (2 females, mean age= 22.4 years, SD=±1.3 years) young adults. The ASC subjects received a diagnosis of mild autistic spectrum disorder by a psychiatric doctor, and their IQ was measured through the Vineland Adaptive Behavior Scales (VABS, Sparrow et al. 2016).

B. Materials

The stimuli consisted of 24 adult (12 females) and 24 children (12 females) faces portraying happiness, sadness, anger, fear, surprise, and disgust. The children stimuli were selected from a database (Esposito et al. 2018) of 84 contemporary children faces, where each emotion was portrayed by 14 different children (7 females) . The adult stimuli were selected from a database (Esposito et al. 2018) of 96 contemporary adult faces, where each emotion was portrayed by 16 different adults (8 females). Both databases were assessed by two different groups of 20 adults (10 females) aged between 20-30 years (mean age=26.22, SD=±2.87). The assessor's agreement percentage ranged between 90%-100% for children and 80%-100% for adult's stimuli - The agreement was obtained through a semi-forced choice task where assessors were presented emotional faces on a computer screen, and asked to select among the emotional labels (happiness, anger, surprise, fear, disgust, sadness, another emotion, and no emotion) that best described the proposed stimulus., 12 more stimuli were selected from the abovementioned databases (3 male and female children, and 3 male and female adults), for testing the subject's task comprehension. Fig. 1 shows examples of the exploited stimuli for surprise.

Fig. 1: Examples of adult and children emotional faces of surprise

The audio stimuli were selected from the COST 2102 multimodal database (Esposito & Riviello, 2010, Esposito 2009). The vocal emotional expressions in this database were produced by actors and actresses in the context of movie scenes and therefore, judged emotionally appropriate by the movie director. We consider these stimuli more naturalistic than those obtained by asking an actress/actor to reproduce the same sentence with different emotional intonations. The audio consisted of 20 stimuli, 4 for each emotion of happiness, disgust, surprise, sadness, and anger, each lasting from 3 to 5 sec and portrayed by four different protagonists (two females). The selected stimuli received The assessors' agreement ranged from 80% to 100% on a semi-forced choice task. Stimuli were presented through headphones. Care was taken in choosing vocal expressions whose semantic meaning did not refer to the portrayed emotional state. Three more stimuli were selected for the training phase.

C. Procedures

Two emotional facial labeling (one for children and one for adult faces) and one audio tests were administered to each group of 15 participants (ASC and TD adolescents, ASC and TD adults) through the Superlab software in a quiet room. Each participant listened to the emotional vocal expressions and saw the children and adult facial expressions. The stimuli were randomly presented on a computer screen. Participants were requested to select among the eight (or seven for audio) labels of happiness, anger, surprise, fear, disgust, sadness, another emotion, and no emotion the one that best described the seen or listened stimulus. No time limits or feedbacks were given. A series of repeated measure ANOVAs were performed on the labeling accuracy aimed to assess: a) the ability of (ASC and TD) adults and adolescents to decode the primary facial emotional expressions portrayed by children and adults;
b) The ability of (ASC and TD) adults and adolescents to correctly decode male and female emotional vocal expressions. The between factors were the group (ASC and TD individuals). The within factors the emotional categories (happiness, surprise, anger, sadness, disgust, and fear) and the gender of the stimuli. Bonferroni post hoc tests were performed in order to explain interaction effects. The confidence interval was set to alpha=.05.

IV. Results

The statistical results for adults and adolescents on faces and voices are discussed below.

A. ASC and TD adults' performance on children and adults' faces

Significant differences (F(1,28)=11.608, p=.002) were found between ASC (mean=1.256) and TD (mean=1.650) adults on their ability to correctly decode adults faces (see Fig. 2a). A significant interaction was found between emotional categories and the group of ASC and TD adults (F(5, 140)=9.672, p<.000). Bonferroni post hoc tests revealed that ASC's ability to recognize adults' anger (ASC mean=1.167, TD mean=1.733, p=.008) surprise (ASC mean=0.600, TD mean=1.787, p<.001) and disgust (ASC mean=1.000, TD mean=1.800, p<.001) was significantly poorer than TD subjects. Significant differences were found among emotional categories (F(5,140)=11.227, p<.001). Bonferroni post hoc tests revealed that ASC subjects were significantly more accurate in decoding adults' faces of happiness (ASC mean=1.967) with respect to fear (ASC mean=1.200, p=.002), anger (ASC mean=1.167, p<.001), surprise (ASC mean=.600, p<.001), and disgust (ASC mean=1.000, p<.001). ASCs were significantly less accurate in decoding adult' faces of surprise with respect to happiness (p<.001), fear (p=.01), anger (p=.013), sadness (ASC mean=1.600, p<.001). TD subjects, were significantly less accurate in decoding adults' faces of fear (TD mean=1.100) with respect to happiness (TD mean =1.867, p=.002), anger (TD mean=1.733, p=.018), surprise (TD mean=1.767, p=.004) and disgust (TD mean=1.800, p=.001).

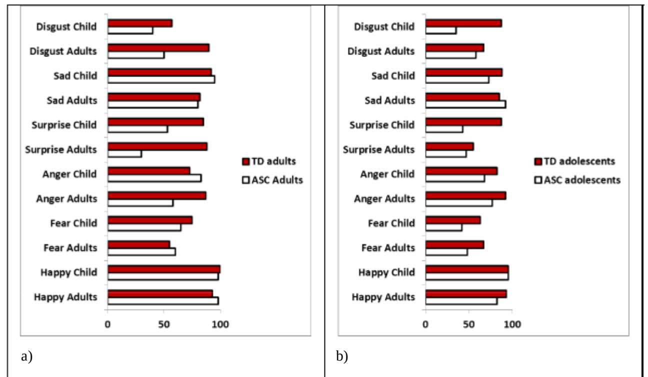

Fig. 2: 2a) Decoding accuracy for adults (ASC and TD) and 2b) adolescents on facial emotional expressions of children and adults

A significant interaction was found between emotional categories and gender of stimuli (F(5,140)=3.539, p=.005). Bonferroni post hoc tests revealed that ASC were significantly less accurate than TD adults in decoding adults' female faces of anger (ASC mean female =.800, TD mean female=1.733, p=.001), adults' male (ASC mean male =.667, TD mean male=1.733, p=.001) and female faces of surprise (ASC mean female =.533, TD mean female=1.800, p<.001) and adults' male (ASC mean male =.800, TD mean male =1.667, p=.002) and female faces of disgust (ASC mean female =1.200, TD mean female=1.933, p=.004).

As for children's stimuli, no significant differences (F(1, 28)=3.259, p=.082) were found between ASC (mean=1.450) and TD (mean=1.606) adults on their ability to correctly decode emotional children's faces, except for surprise (see Fig. 2a)), that explained the interaction found between emotional categories and the group of ASC and TD adults (F(5,140)=4.939, p<.000). Bonferroni post hoc tests revealed that ASC's ability to recognize children's surprise (ASC mean=1.067, TD mean=1.700, p=.001) was significantly poorer than TD subjects. Emotional categories significantly interacted with the gender of stimuli (F(5,140)=3.248, p=.008). Bonferroni post hoc tests revealed that ASC were significantly less accurate than TD adults in decoding children's male (ASC mean male =.933, TD mean male=1.600, p=.011) and female faces of surprise (ASC mean female =1.200, TD mean female=1.800, p=.013), and more accurate than TD adults in decoding children's female faces of sadness (ASC mean female=2.000, TD mean female=1.733, p=.032)

B. ASC's and TD adolescents' performance on children and adults' faces

Slightly significant differences (see Fig. 2b)) were found between ASC (mean=1.367) and TD (mean=1.528) adolescents on their ability to correctly decode adults' faces (F(1,28)=4.592, p=.041). Significant differences were found among emotional categories, (F(5,140)=15.489, p<.000). Bonferroni post hoc tests revealed that ASC adolescents were significantly less accurate in decoding adults' faces of fear (mean=1.033) with respect to happiness (mean=1.633, p=.008), anger (mean=1.533, p=.029), and sadness (mean=1.833, p=.005), and adults' faces of surprise (mean=1.000) with respect to happiness (p=.027), anger (p=.014) and sadness (p=.001). In addition, ASC adolescents were significantly less accurate in decoding adults' faces of disgust (mean=1.167) with respect to sadness (p<.001). TD adolescents were less accurate in decoding adults' faces of surprise (mean= 1.100) with respect to happiness (mean =1.867, p=.004), anger (mean= 1.833, p<.001), and sadness (mean=1.700, p=-029), and adults' faces of fear (mean=1.333) with respect to happiness (p=.025) and anger (p=.029), as well as in decoding adults' faces of disgust (mean=1.333) with respect to happiness (p=.016) and anger (p=.026).

A significant effect was found for the gender of stimuli (F(1,28)=16.282, p<.001) . Bonferroni post hoc tests revealed that TD adolescents were significantly (p=.032) more accurate on adult female (mean = 1.633) than adult male (mean= 1.422) faces. A significant interaction was found between emotions and gender of stimuli (F(5,140)=8.252, p<.001). Bonferroni post hoc tests revealed that ASC adolescents were significantly less accurate that their TD peers in decoding adults' male faces of happiness (ASC mean= 1.667, TD mean = 2.000, p=.013), adults' female faces of fear (ASC mean=1.267, TD mean= 1.800, p=.017), and adults' female faces of anger (ASC mean= 1.333, TD mean=1.800, p=.039)

As for children's stimuli, significant differences (F(1,28)= 30.399, p<.001) were found between ASC (mean= 1.161) and TD (mean=1.672) adolescents (see Fig. 2b)). A significant interaction was found between emotional categories and ASC and TD adolescents (F(5,140)= 6.207, p<.001). Bonferroni post hoc tests revealed that ASC adolescents' ability to recognize children's faces of fear (ASC mean=.833, TD mean=1.267, p=.005), anger (ASC mean=1.100, TD mean= 1.633, p=.003), surprise (ASC mean=.867, TD mean=1.733, p=.001) and disgust (ASC mean=.700, TD mean=1.733, p<.001) was significantly poorer than that of their TD peers.

Significant differences were found among emotional categories (F(5,140)=15.821, p<.001). Bonferroni post hoc tests revealed that ASC adolescents were significantly more accurate in decoding children's faces of happiness (ASC mean=1.900) with respect to fear (ASC mean=.833, p<.001), anger (ASC mean=1.100, p<.001), surprise (ASC mean=.867, p<.001), and disgust (ASC mean=.700, p<.001). ASC adolescents were significantly more accurate in decoding children's faces of sadness (mean=1.567) with respect to fear (p=.002), and disgust (p<.001). TD adolescents were significantly less accurate in decoding children's faces of fear (TD mean=1.267) with respect to children's faces of happiness (TD mean =1.900, p<.001), and disgust (TD mean=1.733, p=.041).

A significant interaction was found between emotional categories and gender of stimuli (F(5,140)=11.228, p<.001). Bonferroni post hoc tests revealed that ASC adolescents were significantly less accurate than TD adolescents in decoding children's male faces of fear (ASC male mean=.933, TD male mean=1.600, p=.006), children's female faces of anger (ASC female mean=.667, TD female mean=1.400, p=.003), children's male (ASC male mean=.733, TD male mean=1.600, p=.005) and female (ASC female mea =1.00, TD female mean =1.867, p=.002) faces of surprise, and children's male (ASC male mean=.600, TD male mean=1.667, p<.001) and female faces (ASC female mean=.800, TD female mean =1.800, p=.001) of disgust.

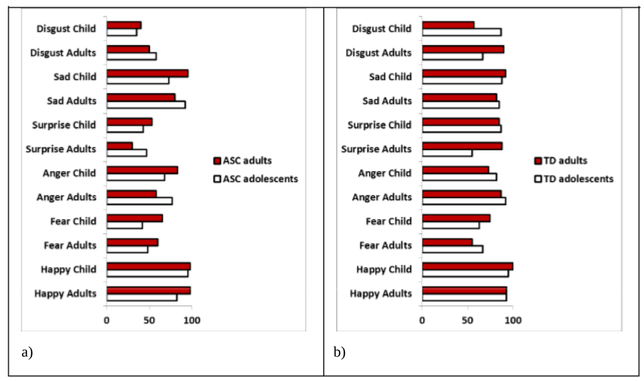

Fig. 3: 3a) Decoding accuracy of ASC (adults and adolescents) and 3b) TD (adults and adolescents) on facial emotional expressions of children and adults

C. ASC adults' and adolescents' performance on faces of children and adults

No significant differences (F(1,28)=.758, p=.391) were found between ASC adults (mean=1.256) and ASC adolescents (mean=1.367) on their ability to correctly decode adults' faces (see Fig. 3a). A significant interaction was found between emotional categories and ASC adults and adolescents (F(5,140)=3.309, p=.012). Bonferroni post hoc tests revealed that ASC adolescents were significantly poorer than ASC adults in decoding adults' faces of happiness (ASC adult mean =1.967, ASC adolescent mean= 1.633, p=.005). Significant differences were found among emotional categories, (F(5,140)=20.811, p<.001). Bonferroni post hoc tests revealed that ASC adults were significantly more accurate in decoding adults' faces of happiness (ASC adult mean=1.967) with respect to fear (ASC adult mean=1.200, p=.005), anger (ASC adult mean=1.167, p=.001), surprise (ASC adult mean=.600, p<.001), and disgust (ASC adult mean=1.000, p<.001). ASC adults were also significantly less accurate in decoding adults' faces of surprise with respect to happiness (p<.001), fear (p=.016), anger (p=.031), and sadness (ASC adult mean=1.600, p<.001). ASC adolescents were significantly more accurate in decoding adults' faces of sadness (ASC adolescent mean=1.833) with respect to fear (ASC adolescent mean =1.033, p=.005), surprise (ASC adolescent mean=1.000, p=.002), and disgust (ASC adolescent mean =1.167, p=.025).

A significant interaction was found between emotional categories and gender of stimuli (F(5,140)=7.071, p<.001). Bonferroni post hoc tests revealed that ASC adolescents were significantly less accurate than ASC adults in decoding adults' female faces of happiness (ASC adult female mean=2.000, ASC adolescent female mean =1.600, p=.005), and significantly more accurate in decoding adults' female faces of surprise (ASC adult female mean=.533, ASC adolescent female mean =1.133, p=.035).

As for children stimuli, significant differences (F(1,28)= 7.544, p=.010) were found between ASC adults (mean= 1.450) and ASC adolescents (mean=1.161) on their ability to correctly decode children's faces (see Fig. 3a)). Bonferroni post hoc tests revealed that ASC adults were significantly more accurate than ASC adolescents in decoding children's faces of fear (ASC adult mean=1.300, ASC adolescent mean =.833, p=.004), anger (ASC adult mean=1.667, ASC adolescent mean=.1.100, p=.001), and sadness (ASC adult mean=1.900, ASC adolescent mean=1.567, p=.023).

Significant differences were found among emotional categories (F(5,140)=31.284, p<.001). Bonferroni post hoc tests revealed that ASC adults were significantly more accurate in decoding children's faces of happiness (ASC adult mean=1.967) with respect to fear (ASC adult mean=1.300, p<.001), surprise (ASC adult mean=1.062, p=.002), and disgust (ASC adult mean=.800, p<.001). ASC adults were also significantly more accurate in decoding children's faces of sadness (ASC adult mean =1.900) with respect to fear (p=.007), surprise (p=.014), and disgust (p<.001), as well as, children's faces of anger (ASC adult mean=1.667) with respect to surprise (p=.014), and disgust (p<.001). ASC adolescents were significantly more accurate in decoding children's faces of happiness (ASC adolescent mean=1.900) with respect to fear (ASC adolescent mean=.833, p<.001), anger (ASC adolescent mean=1.100, p<.001), surprise (ASC adolescent mean=.867, p<.001), sadness (ASC adolescent mean =1.567, p=.039), and disgust (ASC adolescent mean=.700, p<.001), as well as, in decoding children's faces of sadness with respect to fear (p=.001), anger (p=.045), and disgust (p=.001).

A significant interaction was found between emotional categories and gender of stimuli (F(5,140)=7.859, p<.001). Bonferroni post hoc tests revealed that ASC adolescents were significantly less accurate than ASC adults in decoding children's male faces of fear (ASC adult male mean=1.533, ASC adolescent male mean=.933, p=.021), children's female faces of anger (ASC adult female mean=1.533, ASC adolescent female mean=.667, p<.001), and children's male faces of sadness (ASC adult male mean=1.800, ASC adolescent male mean=1.267, p=.02).

D. TD adults' and adolescents' performance on faces of children and adults

Significant differences (F(1,28)=.5.395, p=.028) were found between TD adults (mean=1.650) and adolescents (mean=1.528) on their ability to correctly decode adults' faces (see Fig. 3b)). A significant interaction was found between emotional categories and TD adults and adolescents (F(5,140)=4.798, p<.001). Bonferroni post hoc tests revealed that TD adolescents' ability to correctly decode adults' faces of surprise (TD adult mean=1.767, TD adolescent mean=1.100, p<.001) and disgust (TD adult mean=1.800, TD adolescent mean= 1.333, p=.006) was significantly poorer than that of TD adults.

Significant differences were found among emotional categories, (F(5,140)=8.504, p<.001). Bonferroni post hoc tests revealed that TD adults were significantly less accurate in decoding adults' faces of fear (TD adult mean=1.100) with respect to happiness (TD adult mean=1.867, p<.001), anger (TD adult mean=1.733, p=.009), surprise (TD adult mean=1.767, p<.015), and disgust (TD adult mean=1.800, p<.028). TD adolescents were significantly more accurate in decoding adults faces of happiness (TD adolescent mean=1.867) with respect to fear (TD adolescent mean =1.333, p=.010), surprise (TD adolescent mean=1.100, p<.001), and disgust (TD adolescent mean=1.333, p=.005), as well as, more accurate in decoding adults' faces of anger (TD adolescent mean=1.833) with respect to adults' faces of surprise (p<.001).

A significant effect was found for the gender of stimuli (F(1,28)=8.928, p=.006), with adults' female faces (mean =1.661) more accurately decoded than adults' male faces (mean=1.517). A significant interaction was found between emotional categories and gender of stimuli (F(5,140)=2.626, p=.027). Bonferroni post hoc tests revealed that TD adolescents were significantly more accurate than TD adults in decoding adults' male faces of happiness (TD adult male mean=1.733, TD adolescent male mean =2.000, p=.032), and adults' female faces of fear (TD adult female mean =1.067, TD adolescent female mean=1.800, p=.002). TD adults were significantly more accurate than TD adolescents in decoding adults' male (TD adult male mean=1.733, TD adolescent male mean =.867, p=.001) and female (TD adult female mean =1.800, TD adolescent female mean =1.333, p=.022) faces of surprise, and adults' male (TD adult male mean=1.667, TD adolescent male mean =1.200, p=.039) and female (TD adult female mean=1.933, TD adolescent female mean =1.467, p=.002) faces of disgust.

As for children's stimuli, no significant differences (F(1,28)=.897, p=.352) were found between TD adults (mean=1.606) and adolescents (mean=1.672) on correctly decoding children's faces (see Fig. 3b)). A significant interaction was found between emotions and TD adults and adolescents (F(5,140)=5.479, p<.001). Bonferroni post hoc tests revealed that TD adults were significantly less accurate than adolescents in decoding children's faces of disgust (TD adult mean=1.133, TD adolescent mean=1.733, p=.003).

Significant differences were found among emotional categories, (F(5,140)=12.410, p<.001). Bonferroni post hoc tests revealed that TD adults were significantly more accurate in decoding children's faces of happiness (TD adult mean =2.000) with respect to fear (TD adult mean=1.500, p<.001), anger (TD adult mean=1.467, p=.001), surprise (TD adult mean=1.700, p=.013), and disgust (TD adult mean=1.133, p<.001). TD adults were also significantly less accurate in decoding children's faces of disgust with respect to surprise (p=.003), and sadness (TD adult mean=1.833, p=.001). TD adolescents were significantly less accurate in decoding children's faces of fear (TD adolescent mean=1.267) with respect to happiness (TD adolescent mean=1.900, p<.001), surprise (TD adolescent mean=1.733, p<.001), and sadness (TD adolescent mean =1.767, p=.011).

A significant effect was found for the gender of stimuli (F(1,28)=6.731, p=.015), with children's female (mean =1.583) less accurately decoded than children's male faces (mean= 1.694). A significant interaction was found between emotional categories and gender of stimuli (F(5,140)=5.049, p<.001). Bonferroni post hoc tests revealed that TD adolescents were significantly more accurate than TD adults in decoding children's male faces of anger (TD adult male mean =1.467, TD adolescent male mean=1.867, p=.043), and children's male (TD adult male mean=1.333, TD adolescent male mean =1.667, p=.016), and female faces of disgust (ASC adult female mean=1.333, TD adolescent female mean =1.800, p=.016).

E. ASC and TD adults' and adolescents' performance on voices.

When analyzing voices, significant differences were found between ASC and TD both adults and adolescents (F(1,56) =11.031, p<.001). Bonferroni post hoc tests revealed that these significant differences were between ASC adults (ASC adult mean=1.020), TD adults (TD adult mean=1.567, p<.001), and TD adolescents (TD adolescent mean=1.453, p=.003). Significant differences were also between ASC adolescents (ASC adolescent mean=1.060), TD adults (p<.001), and TD adolescents (p=.009). No significant differences were found between ASC adults and ASC adolescents (p=1), and between TD adults and TD adolescents (p=1)

Significant differences were found among emotional categories, (F(4,224)=40.912, p<.001). Bonferroni post hoc tests revealed that ASC adults were significantly more accurate in decoding angry voices (ASC adult mean=1.633) with respect to happiness (ASC adult mean=.967, p<.001), fear (ASC adult mean=1.300, p=.038), surprise (ASC adult mean=.300, p<.001), and sadness (ASC adult mean=.900, p<.001). ASC adults were also significantly less accurate in decoding surprised voices with respect to happiness (p<.001), fear (p<.001), anger (p<.001) and sadness (p=.010). ASC adolescents were significantly more accurate in decoding angry voices (ASC adolescent mean=1.633) with respect to fear (ASC adolescent mean =1.067, p<.001), surprise (ASC adolescent mean=.333, p<.001), and sadness (ASC adolescent mean =.967, p<.001), and less accurate in decoding surprised voice with respect to happiness (ASC adolescent mean =1.300, p<.001), fear (p=.001), anger (p<.001), and sadness (p=.005)

TD adults were decoding equally well emotional voices of happiness (TD adult mean=1.567), fear (TD adult mean=1.600), anger (TD adult mean=1.733), surprise (TD adult mean=1.533), and sadness (TD adult mean=1.400). TD adolescents were significantly more accurate in decoding angry voices (TD adolescent mean=1.967) with respect to fear (TD adolescent mean=1.500, p=.001) surprise (TD adolescent mean=.967, p<.001), and sadness (TD adolescent mean=1.233, p<.001), and less accurate in decoding surprised voices with respect to fear (p=.034), and happiness (TD adolescent mean =1.600, p<.001).

A significant interaction was found between emotional categories and ASC and TD both adults and adolescents (F(12, 224)=3.790, p<.001). Bonferroni post hoc tests revealed that ASC adults' decoding of happy voices (ASC adult mean =.967) significantly differs from TD adults (TD adult mean =1.567, p=.018) and TD adolescents (TD adolescent mean=1.600, p=.011). ASC adults' decoding of surprised voices (ASC adult mean=.300) significantly differs from TD adults (TD adult mean=1.533, p<.001) and TD adolescents (TD adolescent mean= ,967, p=.008). ASC adults' decoding of sad voices (ASC mean=.900) significantly differs for TD adults (TD adult mean=1.400, p=.042). ASC adolescents' decoding of surprised voices (ASC adolescent mean=.333) significantly differs from TD adults (TD adult mean=1.533, p<.001) and TD adolescents (TD adolescent mean=.967, p=.013). ASC adolescents' decoding of fearful voices (ASC adolescent mean =1.067) significantly differs from TD adults (TD adult mean=1.600, p=.044).TD adolescents' decoding of surprised voices (TD adolescent mean=.967,) significantly differs from TD adults (TD adult mean=1.533, p=.033).

A significant interaction was found between emotional categories and gender of stimuli (F(4,224)=4.260, p=.002). Bonferroni post hoc tests revealed that ASC adults' decoding of male voices of happiness (ASC adult male mean=.800) significantly differs from TD adults (TD adult male mean =1.533, p=.046) and TD adolescents (TD adolescent male mean =.1.533, p=.046). TD adults' decoding of male voices of surprise (TD adult male mean=1.533) significantly differs from ASC adults (ASC adult male mean=.133, p<.001) and ASC adolescents (ASC adolescent male mean =.400, p<.001). TD adolescents' decoding of male voices of surprise (TD adolescent male mean =1.133) significantly differs from ASC adults' (p=.001) and ASC adolescents' (p=.032) decoding of male voices of surprise. TD adults' decoding of female voices of surprise (TD adult female mean =1.533) significantly differs from TD adolescents (TD adolescent female mean=.800, p=.016), ASC adults (ASC adult female mean=.467, p<.001), and ASC adolescents (ASC adolescent female mean =.267, p<.001). TD adults' decoding of female voices of fear (TD adult female mean =1.733) significantly differs from ASC adolescents (ASC adolescent female mean=1.067, p=.036). ASC adults' decoding of female voices of sadness (ASC adult female mean=.533) significantly differs from TD adults (TD adult female mean =1.267, p=.018) and TD adolescents (TD adolescent female mean =1.333, p=.008). TD adolescents' decoding of female voices of sadness significantly differ from ASC adolescents decoding (ASC adolescent female mean =.667, p=.039).

V. Discussion

The first aim of the present investigation was to compare very close aged ASC and TD adults and ASC and TD adolescents on their ability to decode correctly the six primary emotions of happiness, fear, anger, surprise, sadness, and disgust, portrayed by contemporary adults' and children' faces. The data showed that the ability of ASC adults to correctly decode contemporary adults' facial expressions of anger, surprise, and disgust, as well as, contemporary children's facial expressions of surprise was significantly poorer than that of TD adults. This was not the case for ASC and TD adolescents who show equal performances in this task. Conversely, the ability of ASC adolescents to decode correctly children's emotional facial expressions of fear, anger, surprise, and disgust, was significantly poorer than that of TD adolescents. There seems to be a common pattern among emotional categories, since ASC adolescents, as ASC adults, showed impairments for the same emotional categories of anger, surprise and disgust (fear is impaired only in ASC adolescents). However, they differed in the age of the stimuli since in ASC adults these failures are observed for adults' faces, whereas, in ASC adolescents they are observed for children's faces. Since no differences were observed between ASC and TD adolescents in their ability to decode adults' emotional faces, a different emotional learning process can be hypothesized for ASC with respect to TD individuals while aging.

These results suggest that the ability of ASCs to decode correctly emotional facial expressions is affected by participants' age, age of stimuli (adults' and children's faces), and emotional categories and may explain the controversial data reported in literature for faces.

For emotional voices, significant differences were observed between ASC adults and adolescents and TD adults and adolescents. In particular, ASC adults differed significantly from TD adults and adolescents in their ability to decode vocal expressions of happiness and surprise, and from TD adults in their ability to decode vocal expressions of sadness. ASC adolescents differed significantly from TD adults and adolescents in their ability to decode vocal expressions of surprise and, from TD adults, in their ability to decode vocal expressions of fear.

Therefore, when voices are considered, both ASC adults and adolescents showed difficulties in decoding surprise, while the ability to decode fear, happiness, and sadness is differently affected by ASC participants' age. These results, considering the very close aged groups of ASC and TD participants adopted in the present experiments support the idea that ASC's emotion processing impairments are specific and dependent on *particular emotions and/or stimuli* given the participants' age (Uljarevic &Hamilton, 2012).

The second goal of this investigation was to identify accuracy's patterns measuring developmental maturity in emotion labeling of ASC and TD individuals. To this aim, our data show that ASC adolescents' ability to decode correctly adults' happiness was significantly poorer than that of ASC adults. In addition, ASC adults were significantly more accurate than ASC adolescents in decoding children's faces of fear, anger, and sadness. Conversely, TD adolescents were significantly less accurate than TD adults in decoding adults' faces of surprise, and disgust, while TD adults were significantly less accurate than TD adolescents in decoding children's faces of disgust. These data suggest that ASC and TD subjects exploit a different developmental approach in learning emotional facial expressions, which is differently affected by contextual variables such as emotional categories, and is contingent on the age of the stimuli.

When emotional vocal expressions are considered, our data show that ASC adults' ability to decode sad voices was significantly poorer than that of TD adults, and their ability to decode happy and surprised voices was significantly poorer than that of TD adults and adolescents. On the other hand, ASC adolescents were significantly less accurate than TD adults and adolescents in decoding fearful ad surprised voices, whereas, TD adolescents differed significantly from TD adults only in their ability to decode surprised voices. Therefore, for voices, the emotional categories less accurately recognized are not the same observed for faces, suggesting that the communication modes (vocal or visual) convey a different amount of emotional information. In this context, surprise is the vocal emotional expression that still needs to be learned from adolescence to adulthood, but then happy, sad and fearful voices show for ASC and TD individuals a different developmental learning pattern.

Finally, this investigation aimed to identify effects of the gender of stimuli. In this respect, our data show that ASCs adults were significantly less accurate than TD adults in decoding adults' female faces of anger, adults' male and female faces of surprise and disgust, and children's male and female faces of surprise, and more accurate than TD adults in decoding children's female faces of sadness.

ASC adolescents were less accurate than TD adolescents in their ability to decode adults' male faces of happiness, adults' female faces of fear, and anger, children male faces of fear, children female faces of anger, and children male and female faces of surprise, and disgust. ASC adolescents were significantly less accurate than ASC adults in decoding adults' female faces of happiness, children's male faces of fear, and sadness, and children's female faces of anger, and significantly more accurate in decoding adults' female faces of surprise. A different pattern of significance for the gender of stimuli was observed between TD adolescents and adults. TD adolescents were significantly more accurate than TD adults in decoding adults' male faces of happiness, adults' female faces of fear, children's male faces of anger, and children's male and female faces of disgust. By contrast, TD adults were significantly more accurate than TD adolescents in decoding adults' male and female faces of surprise, and disgust.

As for voices, ASC adults were significantly less accurate than TD adults and adolescents in decoding happy male voices and female sad voices. TD adults and adolescents were significantly more accurate than ASC adults and adolescents in decoding male voices of surprise. TD adults were significantly more accurate than TD and ASC adolescents, and ASC adults, in decoding female voices of surprise. TD adults were significantly more accurate than ASC adolescents in decoding female voices of fear. TD adolescents were significantly more accurate than ASC adolescents in decoding of female voices of sadness.

These data suggest that the gender and type (voices or faces) of stimuli, participants' age, and communication modes, all together play a role in the global emotion recognition process of both ASC and TD individuals making difficult the definition of a common and shared pattern of emotion recognition compliance. Since individual personal traits, emotion regulation processes, personality, personal experience, and attachment style are also known to affect the human ability to decode emotional faces and voices (Esposito et al. 2014, Gross, 2002), the present data recommend care in generalizing and suggest personalized rehabilitation tools for these communication disorders.

VI. Conclusion

Although commonalities exist among vulnerable people in their requests of help and needs of support, ICT (Information Communication Technologies) assistive technologies for health care must take into account personal demands among the vulnerable persons to be cared for. Therefore, the design of user centered assistive technologies should not only take into account the symptoms of the illness but also individual differences. In other words, there is a need of technologies able to provide personalized care and support for rehabilitation and assistance. This research provides some knowledge on the different ability of ASC adults and adolescents to decode facial and vocal expressions of emotions. Incorporating this knowledge in health care technologies requires to develop algorithms able to learn user needs and adapt their services to individual requirements and expectations. Artificial intelligent algorithms are the route to provide an efficient and effective assistance.

Acknowledgment

The research leading to these results has received funding from a) the European Union Horizon 2020 research and innovation programme under grant agreement N. 769872 (EMPATHIC) and N. 823907 (MENHIR); b) the Italian project SIROBOTICS, MIUR, PNR 2015-2020, DD 1735 July 13 2017; c) the Italian project ANDROIDS, Università della Campania "Luigi Vanvitelli" programme V:ALERE 2019, funded with D.R. 906 del 4/10/2019, prot. n. 157264, Oct 17 2019; d) the Spanish project TIN2016-74946-P (MINECO /FEDER, UE) and CERCA Programme / Generalitat de Catalunya, e) and by ICREA under the ICREA Academia programme

References

[1] American Psychiatric Association (2013) Diagnostic and statistical manual of mental disorders, 5th ed. American Psychiatric Publishing.

[2] Baron-Cohen, S., Spitz, A., & Cross, P. (1993). Do children with autism recognise surprise? A research note. Cognition and Emotion, 7, 507–516.

[3] Charbonneau G, Bertone A, Lepore F, Nassim M, Lassonde M, Mottron L, Collignon O (2013) Multilevel deficits in the processing of audio-visual emotion expressions in autism spectrum disorders. Neuropsychologia.

[4] Esposito A, Palumbo D, Troncone A (2014) The influence of the attachment style on the decoding accuracy of emotional vocal expressions. Cognitive Computation, 6(4), 699-707.

[5] Esposito A, Esposito AM, Cordasco G, Maldonato MN, Vogel C, Bourbakis N. (2018) Emotional faces of children and adults: What changes in their perception. 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom 2018), August 22-24, 2018, Budapest, Hungary.

[6] Esposito A, Riviello MT (2010): The new Italian audio and video emotional database. In A. Esposito et al. (Eds.), LNCS 5967, 406–422, Springer-Verlag Berlin Heidelberg

[7] Esposito A (2009) The perceptual and cognitive role of visual and auditory channels in conveying emotional information. Cognitive Computation Journal, , 1 (2) 268-278

[8] Harms MB, Martin A, Wallace G (2010) Facial emotion recognition in autism spectrum disorders: a review of behavioral and neuroimaging studies. Neuropsychol., Rev 20:290–322.

[9] Humphreys K, Minshew N, Leonard GL, Behrmann,M (2007) A fine-grained analysis of facial expression processing in high-functioning adults with autism. Neuropsychologia, 45, 685–695.

[10] Griffiths S, Jarrold C, Penton Voak IS, Woods AT, ·Skinner AL, Munafò MR (2017) Impaired recognition of basic emotions from facial expressions in young people with autism spectrum disorder: Assessing the importance of expression intensity. J Autism Dev Disord,1-11.

[11] Gross JJ (2002) Emotion regulation: affective, cognitive, and social consequences. Psychophysiology,39(3):281-91.

[12] Jones CRG, Pickles A, Falcaro M, Marsden AJS, Happe´ F, Scott SK, Sauter D, Tregay J, Phillips RJ, Baird G, Simonoff E, Charman T (2011) A multimodal approach to emotion recognition ability in autism spectrum disorders. J Psychol Psychiatry 52:275–285

[13] Kirchner JC, Hatri A, Heekeren HR, Dziobek I (2011) Autistic symptomatology, face processing abilities, and eye fixation patterns. J Autism Dev Disord., 41(2):158-67.

[14] Lerner MD, McPartland JC, Morris JP (2013) Multimodal emotion processing in autism spectrum disorders: An event-related potential study. Developmental Cognitive Neuroscience 3 (2013) 11– 21.

[15] Loth E, Garrido L, Ahmad J, Watson E, Duff A, Duchaine B (2018) Facial expression recognition as a candidate marker for autism spectrum disorder: how frequent and severe are deficits? Molecular Autism (2018) 9:7.

[16] Oerlemans AM, van der Meer JM, van Steijn DJ, de Ruiter SW, de Bruijn YG, de Sonneville LM, Buitelaar JK, Rommelse NN (2014) Recognition of facial emotion and affective prosody in children with ASD (+ADHD) and their unaffected siblings. Eur Child Adolesc Psychiatry, 23(5):257-71.

[17] Rozin P, Haidt J, McCauley CR (2000). Disgust. In M. Lewis & J. Havliand (Eds.), *Handbook of emotions* (2nd ed., Vol. 1, pp. 637--653). New York: Guilford Press

[18] Sparrow SS, Cicchetti DV, Saulnier CA (2016) Vineland Adaptive Behavior Scales, Third Edition (Vineland-3).

[19] Sucksmith E, Allison C, Baron-Cohen S, Chakrabarti B, Hoekstra RA (2013) Empathy and emotion recognition in people with autism, first-degree relatives, and controls. Neuropsychologia, 51:98–105.

[20] Taylor LJ, Maybery MT, Grayndler L, Whitehouse AJO (2015) Evidence for shared deficits in identifying emotions from faces and from voices in autism spectrum disorders and specific language impairment. Int J of Language and Communication Disord, 50 (4), 452–466

[21] Uljarevic M, Hamilton A (2012) Recognition of emotions in autism: a formal meta-analysis. J Autism Dev Disord., 43(7):1517-26.

[22] Xavier J, Vignaud V, Ruggiero R, Bodeau N, Cohen D, Chaby L (2015) A multidimensional approach to the study of emotion recognition in autism spectrum disorders. Frontiers in Psychology, doi: 10.3389/fpsyg.2015.0195