Impact of sensory preferences of individuals with autism on the recognition of emotions expressed by two robots, an avatar, and a human

Pauline Chevalier, Jean-Claude Martin, Brice Isableu, Christophe Bazile & Adriana Tapus

Autonomous Robots

ISSN 0929-5593

Auton Robot
DOI 10.1007/s10514-016-9575-z

Springer

Auton Robot
DOI 10.1007/s10514-016-9575-z

Impact of sensory preferences of individuals with autism on the recognition of emotions expressed by two robots, an avatar, and a human

Pauline Chevalier[1,2,3,4] · **Jean-Claude Martin**[2] · **Brice Isableu**[3,4] · **Christophe Bazile**[5] · **Adriana Tapus**[1]

Received: 30 March 2015 / Accepted: 13 May 2016

Abstract We design a personalized human-robot environment for social learning for individuals with autism spectrum disorders (ASD). In order to define an individual's profile, we posit that the individual's reliance on proprioceptive and kinematic visual cues should affect the way the individual suffering from ASD interacts with a social agent (human/robot/virtual agent). In this paper, we assess the potential link between recognition performances of body/facial expressions of emotion of increasing complexity, emotion recognition on platforms with different visual features (two mini-humanoid robots, a virtual agent, and a human), and proprioceptive and visual cues integration of an individual. First, we describe the design of the EMBODI-EMO database containing videos of controlled body/facial expressions of emotions from various platforms. We explain how we validated this database with typically developed (TD) individuals. Then, we investigate the relationship between emotion recognition and proprioceptive and visual profiles of TD individuals and individuals with ASD. For TD individuals, our results indicate a relationship between profiles and emotion recognition. As expected, we show that TD individuals that rely more heavily on visual cues yield better recognition scores. However, we found that TD individuals relying on proprioception have better recognition scores, going against our hypothesis. Finally, participants with ASD relying more heavily on proprioceptive cues have lower emotion recognition scores on all conditions than participants relying on visual cues.

Keywords Autism · Socially assistive robotics · Proprioception · Vision · Kinematics · Emotions

This is one of several papers published in *Autonomous Robots* comprising the "Special Issue on Assistive and Rehabilitation Robotics".

✉ Pauline Chevalier
pauline.chevalier@ensta-paristech.fr

Jean-Claude Martin
martin@limsi.fr

Brice Isableu
brice.isableu@u-psud.fr

Christophe Bazile
cbazile@gapas.org

Adriana Tapus
adriana.tapus@ensta-paristech.fr

1 Robotics and Computer Vision Lab, U2IS ENSTA ParisTech, Université Paris-Saclay, 828 bd des Maréchaux, 91762 Palaiseau cedex, France

2 LIMSI, CNRS, Université Paris-Saclay, Bât 508, Campus Universitaire, 91405 Orsay, France

3 CIAMS, Univ. Paris-Sud, Université Paris-Saclay, 91405 Orsay Cedex, France

4 CIAMS, Université d'Orléans, 45067 Orléans, France

5 I.M.E. Notre Ecole, 78955 Carrière sous Poissy, France

1 Introduction

Currently, research on Socially Assistive Robotics (SAR) is expanding (Feil-Seifer and Mataric 2005; Tapus et al. 2007). One of the target populations is people suffering from autism spectrum disorders (ASD). Individuals with ASD have impaired skills in communication, interaction, emotion recognition, joint attention, and imitation (Charman et al. 1997). Many studies have showed that children with ASD have a great affinity for robots, computers, and mechanical components (Hart 2005). In the field of SAR, robots are used as tools in socialization therapies for children with ASD in order to enhance social engagement, imitation, and joint

Published online: 26 May 2016

attention skills (Tapus et al. 2012; Kim et al. 2013; Wainer et al. 2010; Kozima et al. 2005).

An overreliance on proprioceptive information in autism has been suggested. *Proprioception* can be defined as the ability of an individual to determine body segment positions (i.e., joint position sense) and detect limb movements (kinesthesia) in space. It is derived from complex somatosensory signals provided to the brain by multiple muscles (Goodwin et al. 1972; Burke et al. 1976; Roll and Vedel 1982), joints (Ferrell et al. 1987), and skin receptors (Edin 2001). Individuals with ASD show normal to exacerbated integration of proprioceptive cues compared to typically developed (TD) individuals (Gowen and Hamilton 2013). Proprioceptive integration in ASD is studied so as to better understand how the contribution of these cues influences interactive and social capacities. In (Haswell et al. 2009), the authors observed that the greater the reliance on proprioception, the more children with ASD exhibit impairments in social functions and imitation. Moreover, limited visual processing skills lead to difficulties in managing social interactions. Vision is an important component for communication and social behaviors. An impairment in visual processing and the gaze function of individuals with ASD may lead to unusual eye contact, difficulty in following the gaze of others or supporting joint attention, and difficulty in interpreting facial and bodily expressions of emotions (Simmons et al. 2009). In addition, visual field dependent individuals are considered to be more group-oriented and cooperative and less competitive than field independent individuals (Liu and Chepyator-Thomson 2009). Visual field dependent individuals are strongly interested in people and get closer to the person with whom they are interacting, while visual field independent individuals appear to be cold and detached; they are socially isolated but have good analytic skills (Saracho 2003).

The long-term goal of our research is to develop a new personalized robot/virtual-agent-based social interaction model, in addition to standard therapy, for individuals suffering from ASD. Such personalized interaction is advantageous for enhancing social skills. There are strong interindividual differences among individuals with ASD. Thus, such an interaction need to be personalized, and the interindividual differences that are relevant to the components of social interaction, such as recognition of facial and bodily expressions of emotions, must be considered. To define relevant individual profiles, we posit that an individual's reliance on proprioceptive and kinematic visual cues will affect the way they interact with a humanoid robot (Haswell et al. 2009; Liu and Chepyator-Thomson 2009; Saracho 2003; Coates et al. 1975). We hypothesize that *a mitigated behavioral response (i.e., hyporeactivity) to visual motion and an overreliance on proprioceptive information are linked in individuals with ASD to their difficulties in integrating social cues and engaging in successful interactions* (**H0**).

To the best of our knowledge, no study has examined the sensorimotor and visual profiles of individuals suffering from ASD in order to determine and elaborate individualized and personalized scenarios for Human-Robot social interaction therapy.

As a first step, we define each participant's perceptivo-cognitive and sensorimotor profiles with respect to the integration of proprioceptive and visual inputs. Next, we develop the EMBODI-EMO database,[1] a video database of expressions of emotions. We use this database to evaluate individuals' skills to recognize body and facial expressions of emotions on different supports with increasing visual difficulties and to analyze relationships between these skills (recognition scores) and a user's profiles. With this view, a user's sensory profile, more specifically, their visual and proprioceptive dependencies, is assessed to analyze whether they modulate the recognition of nonverbal expressions of emotions.

We use and compare different platforms (two humanoid robots, one avatar, and one real human) to control the complexity of bodily expressions of emotions for the recognition task: Nao robot from Aldebaran Robotics, R25 from Hanson Robotics, female virtual agent from Multimodal Affective and Reactive Character (MARC) (Courgeon and Clavel 2013), and a female human (Fig. 1). We observe and compare behaviors toward different platforms and the influence of their embodiment and facial features on the recognition of emotions.

We make three hypotheses regarding the impact of proprioceptive and visual integration profiles on an individual's emotion recognition performance.

- **H1.1** Individuals with high reliance on proprioceptive cues should more easily recognize emotions expressed by less visually complex platforms than individuals with low reliance on proprioceptive cues.
- **H1.2** Individuals with high reliance on proprioceptive cues should less easily recognize expressions of emotions than individuals with low reliance on proprioceptive cues.
- **H2** Individuals with high reliance on visual cues should more easily recognize emotions than individuals with low reliance on visual cues, whatever the platform or the channels (i.e., expressed by the face and/or body).

This paper is structured as follows: Sect. 2 presents related work. Section 3.1 shortly presents the participants. Section 3.1 describes the assessment of each participant's perceptivo-cognitive and sensorimotor profiles. It also presents the results of a first time greeting interaction with the Nao robot and the participants in light of their profiles. Section 4 focuses on (1) the collection of the EMBODI-EMO database con-

[1] Database link: http://perso.ensta-paristech.fr/~tapus/eng/media/EMBODI-EMO.zip.

Fig. 1 Different platforms in the neutral pose, from *top left* to *bottom right* Nao-robot, Zeno-robot, Mary-virtual agent, and Pauline-human

taining combinations of facial and bodily expressions of emotions with increasing complexities displayed on different platforms and (2) its validation by a TD population. Section 5 presents a correlation analysis between emotion recognition scores and visual and movement sensitivity scores in TD individuals. Section 6 analyzes the relationships between emotion recognition scores and profiles of our participants. Section 7 concludes our work.

2 Related work

2.1 Atypical integration of proprioceptive and visual integration of cues in ASD

Motor, sensory, and visual processing impairments are present in autism but are not taken into account in the ASD diagnostic (Simmons et al. 2009; Gowen and Hamilton 2013). However, these deficits have an influence on the quality of life of individuals suffering from ASD and on their social development. In (Bar-Haim and Bart 2006), the authors observed the impact of motor impairment in children on their social participation in kindergarten. Their work showed that children with motor impairments are more likely to have solitary play and less interaction with peers, and therefore, do not explore their physical and social environment leading to social and emotional difficulties. Moreover, visual impairment leads to difficulties in social behaviors, a fact that has been largely documented in the literature. Vision is an important component of communication and social behaviors. An impairment in visual processing and the gaze function in individuals with ASD may lead to unusual eye contact, difficulty in following the gaze of others or supporting joint attention, and difficulty in interpreting facial and bodily expressions of emotions (Simmons et al. 2009). In (Qin et al. 2014), the authors hypothesized that the differences between ASD and TD individuals' perceptions caused communication problems. The authors proposed a simulator to enable TD individuals experience the perceptual over- and under-responsivity of individuals with ASD.

An overreliance on proprioceptive information in autism is also suggested (Gowen and Hamilton 2013; Greffou et al. 2012; Kohen-Raz et al. 1992; Gepner et al. 1995; Gepner and Mestre 2002; Haswell et al. 2009). Individuals with autism show normal to exacerbated integration of proprioceptive cues compared to TD individuals (Gowen and Hamilton 2013). More specifically, in (Greffou et al. 2012) and (Kohen-Raz et al. 1992), the authors observed abnormal postural behavior in autism. They found that individuals with ASD show less age-related postural behaviors and are less stable than TD individuals. Results in (Greffou et al. 2012) suggest that postural hyporeactivity to visual information is present in the tested individuals with autism (individuals suffering from ASD with IQs comparable to those of TD individuals). Furthermore, Gepner et al. (1995) pointed out that individuals with ASD show very poor postural response to visual motion and have movement perception impairments. This result was also observed in (Gepner and Mestre 2002). Proprioceptive integration in ASD has been studied to better understand how the contribution of these cues influences interactive and social capacities. In Haswell et al. (2009), the authors asked 14 children with ASD and 13 TD children to perform a reaching task with their arm while holding a robotic arm that applied a force constraining the movement. The authors observed that the autistic brain built a stronger than normal association between self-generated motor commands and proprioceptive feedback, confirming an overreliance on proprioceptive cues in individuals with ASD. Furthermore, they established that the greater the reliance on proprioception, the greater the children with ASD exhibit impairments in social function and imitation. These findings are consistent with the anatomical miswiring in the ASD brain, i.e., the observed overgrowth of short-range white matter connections between somatosensory regions and motor brain regions (these short-range connections favor the movement coding in the intrinsic coordinates of joints and muscles) and an underdevelopment of more distant connections between distant brain regions (e.g., visual-motor and premotor/posterior parietal cortex processing engaged during the planning phase of complex action sequences both favor a coding in extrinsic coordinates of the task) (Herbert et al. 2004).

2.2 Robots in ASD therapy

The use of robots in therapy for individuals with ASD has been a great topic of interest over the past few years. Indeed, robots have been found to be great partners for facilitating learning, social interaction, and development imitation skills. In addition, robots are attractive mechanical systems to people with ASD (Hart 2005; Kim et al. 2013). They are appealing and engaging for individuals with ASD because their nonverbal behaviors are quite simple and predictable. Robots allow a user to experience endless social interactive scenarios in which complexities of social situations, of bodily expression of emotions, can be controlled to facilitate the detection of regularities and predictable events. Research on robots for therapy for individuals with ASD is rapidly growing and focuses on the following (Feil-Seifer and Mataric 2005): (1) designing sociable robots especially modeled for ASD, (2) designing Human-Robot Interaction (HRI), i.e., knowledge, methods, and algorithms for natural, transparent HRI that enables humans and robots to interact effectively and cooperatively, and (3) evaluating these robots in therapies.

Positive effects of robots on ASD participants have been observed. Dautenhahn et al. (2009) set up dyadic interactions through play between autistic children and a robot or human partner. Each child had four interaction sessions, starting with human partners and alternating between human and robot partners. An increased collaborative behavior was observed with a human partner after a session with the robot. Kim et al. (2013) used a robot, human, and computer program as a bridge between a human partner and a child with autism in a triadic interactions. Results suggest that children with autism display more interest in human partners during the triadic interaction with the robot. In therapy designed for children with autism, Kozima et al. (2005) found that children with ASD engaged spontaneously in dyadic play with the robot Keepon. This study was also expanded to a triadic interaction between a robot, an adult, and a child. In light of these encouraging findings, many challenges in SAR for individuals with ASD must be addressed. Because of small subject pools and/or short-term experiments, generalized results in the improved skills are often questionable (Scassellati et al. 2012). Salter et al. (2010) discussed how variability (i.e., the environment: experimental or real-life and the autonomy of the robot and/or the participant) of the HRI setup may bias the results in this area. Considering these limitations, the new challenge of assistive robot therapies is to identify how to reduce the bias in such findings. In particular, Thill et al. (2012) proposes a new step in robot-assisted therapy; robotic-assisted therapeutic scenarios should develop more substantial levels of autonomy, which would allow the robot to adapt to the individual needs of children over longer periods of time.

2.3 Emotion recognition in ASD therapy

Emotion recognition plays an important role in interaction and social communication and has often been used in ASD therapies; examples include a computer program (Silver and Oakes 2001), cartoon using animated vehicles with real emotional faces for young children (Golan et al. 2010), and with expressive robots (Costa et al. 2014). Emotions are expressed through different channels such as facial expressions (Ekman and Friesen 1984) body movements/postures (Mehrabian 1972), and speech (Scherer 1995). In (Simmons et al. 2009), the authors observed that people with ASD have difficulties in recognizing emotions expressed with facial and body movements. Moreover, the participants with ASD could not easily describe and recognize emotions from a body in motion and interpreted movement in a social context (Simmons et al. 2009); the face was interpreted as a complex object, without social interest in autistic individuals (Simmons et al. 2009). The recognition of facial expression of emotions by individuals with autism and TD individuals has been studied using human pictures (Baron-Cohen et al. 1997), virtual agents (Courgeon et al. 2012), and robots (Costa et al. 2014). Overall, TD individuals recognize facial expressions better than ASD individuals. However, in (Begeer et al. 2006), the authors observed that children with ASD have greater difficulty recognizing emotions without context, whereas TD children performed identically, with or without context. Additionally, Celani et al. (1999) observe that individuals with ASD have good observation strategies and can indeed detect emotional cues.

3 Clustering participants with ASD using sensory preferences

3.1 Description of participants

Our work was done in collaboration with two care facilities for people suffering from ASD: MAIA Autisme (France), an association for children and adolescents with ASD, and FAM-La Lendemaine (France), a residence for adults with ASD. Informed consent for participation was obtained from the parents or by the participants themselves when able. The experimental protocol was approved by the EA 4532 local ethics committee. Our subject pool was composed of six autistic children (10.9 ± 1.8 years) and seven autistic adults (26.1 ± 7.9 years) from these two care facilities. There were nine male and four female participants. A description of each participant's information is given in Table 1. For confidentiality purposes, we encoded participant identities

Table 1 Information on the adults and children participating in the study

ID#	Gender	Age	Comments
AD1	F	24	–
AD2	F	24	Creatine Transporter Deficiency
AD3	M	43	Asperger Syndrome
AD4	F	20	Followed an imitation therapy program
AD5	M	26	Suffers from epilepsy
AD6	M	24	Suffers from echolalia
AD7	M	20	–
CH1	M	10	–
CH2	M	8	Suffers from echolalia
CH3	M	11	–
CH4	M	12	Asked to be part of the program so as to meet Nao.
CH5	F	10	West Syndrome (uncommon-to-rare epileptic disorder), Non-verbal
CH6	M	13	–

as follows: AD# for the participants from the adults care center and CH# for the participants from the children care center. The symbols M and F were used to denote male and female, respectively.

3.2 Defining proprioceptive and kinematic profiles

In our previous work (Chevalier et al. 2016), we described a first experiment that defined each participant's perceptivo-cognitive and sensorimotor profiles. We succeeded to form three groups with significant different behavioral responses, thanks to the Adolescents/Adults Sensory Profile questionnaire (Brown and Dunn 2002) and an experimental setup.

The Adolescents/Adults Sensory Profiles (AASP) developed by Dunn (Brown and Dunn 2002) were completed by all participants. This questionnaire enabled us to collect an individual's movement, visual, touch, and auditory sensory processing preferences described in terms of the quadrants in Dunn's model of sensory processing (Dunn 1999): low registration, sensation seeking, sensory sensitivity, and sensation avoiding. These quadrants are formed by the junctions between individual differences in neurological thresholds for stimulation (high-low) and self-regulation strategies (active-passive).

We designed an experimental setup to assess the effect of a moving virtual visual scene on postural control and an individual's capability to use proprioceptive inputs provided in dynamics of balance to reduce visual dependency (Isableu et al. 2011; Bray et al. 2004). In an unstable posture, the integration of proprioceptive feedback differs among individuals. An individual who integrates proprioceptive cues less than other individuals is called visual dependent. For example, while exposed to visual motion in an unstable posture, his/her body sway follows the visual stimulus.

Participants were asked to stand quietly in various postural conditions of increasing balance difficulty (normal vs. tandem Romberg) in front of a virtual scene rolling at 0.25 Hz with an inclination of $\pm 10°$. They were asked to stand on a force platform in front of the virtual visual scene, static or rolling.

For all sessions, a force platform was used to record the displacement of the center of pressure (CoP). We computed the root mean square (RMS) of the CoP in both mediolateral and anteroposterior directions as an indicator of an individual's stability. Indeed, the RMS provided us the information about the variability of the CoP in space (Chiari et al. 2002). In order to assess the correlation between sensory preferences and postural responses to our visual stimulus, we computed the correlation between the scores of different items of the AASP and the data obtained from the CoP recording. We also computed the correlation between the CoP behaviors during different conditions to evaluate if a behavior in a condition induced another behavior in a different condition. Then, we used clustering analysis (dendrogram, Ward method) to form significantly different behavioral groups.

Results of the AASP and experimental setup allowed us to identify three groups with significant different behavioral responses to the moving virtual visual scene:

- Group **G1** (participants: *AD2; AD4; AD6; CH3; CH5*) exhibited high scores in Movement Sensory Sensitivity (MSS) and visual sensation seeking, low scores in Visual Sensory Sensitivity (VSS), and strong visual independence to the moving virtual visual scene, suggesting an overreliance on proprioceptive input and hyporeactivity to visual cues.
- Group **G2** (participants: *AD1; AD3; CH1; CH4*) exhibited moderate scores in MSS and low scores in visual sensation seeking, low scores in VSS, and moderate reactivity to the moving virtual visual scene, suggesting that they rely evenly on both visual and proprioceptive inputs.

– Group **G3** (participants: *AD5; AD7; CH2; CH7*) exhibited high scores in VSS and low scores in MSS and in visual sensory seeking, and hypereactivity to the moving virtual visual scene, suggesting weak proprioceptive integration and strong visual dependency.

3.3 Analysis of the first interactions with Nao

A first analysis of the behavior of the participants towards the Nao robot was conducted, as reported by Chevalier et al. (2016). The purpose of the interaction was to present the Nao robot to children and adults with ASD for a short duration, up to two minutes. In general, some individuals with ASD are reluctant to unusual events and changes in their daily routine. By introducing them to the Nao robot in the context of a short greeting task, we softly introduced the robot as a social partner. Indeed, in (Meltzoff et al. 2010), researchers observed that children who saw the robot act in a social-communicative way were more likely to follow its gaze than those who did not.

We presented participants with the following scenario: after being seated in front of Nao, the robot waved to the participant and said "Hello, I am Nao. You and I, we are going to be friends." If the participant was verbal, the robot asked for the participant's name. Afterwards, it continued with "Hello" followed by the name of the participant, and asked if he/she wanted him to dance for him/her and then danced. For the duration of the experiment, the participant was with his/her caretaker. The caretaker encouraged the participant to look at and answer the robot.

We video-analyzed the participant's gaze direction and gestures towards the robot, caretaker, and other directions. We analyzed the behavior of each participant to determine a match to the three groups described in Sect. 3.2. Our goal was to evaluate our hypothesis **H0**, which states that the behavioral response to visual and proprioceptive cues influences the success of the interaction between the participant and robot.

Overall, most of the participants showed a strong interest in the robot. The caretakers were amazed and surprised by the attention and reactions of the participants to the robot. We also found a relationship between the social behavior displayed by the participants and their groups as discussed in Sect. 3.2. Individuals from **G3** showed more social behaviors than the other groups, with a gaze more focused towards Nao, more social gestures, and by answering directly to Nao. Individuals from **G2** showed more moderate social behavior than individuals from group **G3** by switching more of their gaze in directions other than the robot. They were also less inclined to answer directly to the robot and to perform social gestures. Finally, individuals from group **G1** showed less social behaviors than the other two groups. These results are encouraging with respect to our hypothesis **H0**.

4 The EMBODI-EMO database of facial and bodily expressions of emotions with different embodiments

4.1 Design

We wanted to test the recognition of emotions expressed by different embodiments and also on different channels such as facial expressions and body expressions. We were unable to find a suitable database, so we decided to create our own database.[2] We recorded different types of expressions with respect to: (1) monochannel: the emotion is expressed only through one communication channel (e.g., on the face or on the body), and (2) multichannel: the body and face express the same emotion in a synchronous manner.

In this work, we used four basic emotions: anger, happiness, fear, and sadness. We selected these four emotions for several reasons. In particular, these emotions are part of the six basic emotions proposed by (Ekman and Friesen 1984) and are widely documented in the literature. This permits us to have reliable sources to elaborate animations and behaviors. Moreover, we can also compare our work with previous studies on how these emotions are perceived by TD and ASD individuals. Furthermore, these emotions have been chosen because of their interest and accessibility to individuals with ASD. Indeed, while designing the database, we asked caretakers in charge of individuals with ASD which emotions to choose, and they indicated to us that emotions such as "disgust" and "contempt" are too complex to understand for their participants.

We used four platforms to create our emotional video database. They can display different levels of embodiment and visual complexity, as shown in Fig. 1 and detailed below:

– Nao (Aldebaran Robotics): a mini-humanoid robot with a static face.
– Zeno (R25, Hanson Robotics): a mini-humanoid robot with a silicon-made actuated skin face, which enables it to express emotions on its face.
– Mary (MARC, LIMSI-CNRS): a female humanoid virtual agent.
– Pauline: a female human.

Videos began in a neutral posture and ended with the climax of the expression of emotion. All of the platforms occupied the same space in the video (feet and top of the head were at the same place for each platform). The physical platforms (Nao, Zeno, and Pauline) were placed in front of a white wall, and the virtual platform (Mary) was in a white environment. All of the videos were one to two seconds long.

[2] Link to the database: http://perso.ensta-paristech.fr/~tapus/eng/media/EMBODI-EMO.zip.

Table 2 Emotion representation by action units (AU), and the ability of Zeno to perform them. Mary (virtual agent) and Pauline (human) were able to perform all of these AU

Emotion	AU combination	Zeno
Anger	AU4—Brow Lowerer	Yes
	AU7—Lid Tightener	Yes
	AU10—Upper Lip Raiser	No
	AU17—Chin Raiser	No
	AU23—Lip Tightener	No
	AU24—Lip Pressor	Yes
Happiness	AU6—Cheek Raiser	No
	AU12—Lip Corner Puller	Yes
	AU25—Lips Part	Yes
Fear	AU1—Inner Brow Raiser	Yes
	AU2—Outer Brow Raiser	Yes
	AU4—Brow Lowerer	No*
	AU5—Upper Lid Raiser	Yes
	AU20—Lip Stretcher	No
	AU25—Lips Part	Yes
	AU26—Jaw Drop	Yes
Sadness	AU1—Inner Brow Raiser	Yes
	AU4—Brow Lowerer	No*
	AU15—Lip Corner Depressor	Yes

* Zeno was only able to have its brow up or down and could not display AU1 and AU4 at the same time

All of the platforms performed the same movements. As such, we adapted the behaviors of the platforms to have similar movements between different embodiments. This was also done in similar studies (Coulson 2004; Erden 2013). However, several difficulties in movement adaptation were encountered, especially due to different numbers of joints, dynamics of animation, and different balances of platforms. We did our best to adapt movements through different platforms to achieve homogeneity.

4.1.1 Facial animation

The platforms (except for the Nao robot) have facial features, which enable them to express the four emotions present in our database (happiness, anger, fear, and sadness). The animation of the emotions was designed based on a combination of action units (AU) (Ekman and Friesen 1984) as a function of the capabilities of the platforms (see Table 2). One video was recorded for each emotion for Zeno, Mary, and Pauline.

4.1.2 Body animation

To elaborate on body animations, we were inspired by the Bodily Expressive Action Stimulus Test (BEAST) database (De Gelder and Van den Stock 2011). This database contains 254 photos of whole body expressions from 46 actors expressing four emotions (anger, fear, happiness, and sadness). We selected three images for each emotion to create our animations (anger: *F03AN, F11AN, F22AN* were chosen for animations *AN1, AN2, AN3*, respectively; fear: *F04FE, F16FE, F32FE* were chosen for *FE1, FE2, FE3*, respectively; happiness: *F02HA, F06FA, F22HA2* were chosen for *HA1, HA2, HA3*, respectively; sadness: *F02SA, F04SA, F24SA2* were chosen for *SA1, SA2, SA3*, respectively). To animate movement from a neutral posture to the posture depicted by BEAST images, we performed linear interpolation. The velocity of the movement was based on observations of body emotion expression in humans (Wallbott 1998); specifically, a greater velocity in movement for anger and fear, and low dynamics for happiness and sadness. Three animations were built for each emotion in Nao, Zeno, Mary, and Pauline.

4.1.3 Body and facial animation

We combined facial animations with corresponding body expression of emotion animations, resulting in three animations for each emotion for Zeno, Mary, and Pauline.

4.1.4 Content of our database

Our database contained 96 videos formed by 12 videos of facial expressions (four emotions × three platforms), 48 videos of body expressions (three variations × four emotions × four platforms), and 36 videos of body and facial expressions (three variations × four emotions × three platforms).

4.2 Validation of the EMBODI-EMO database

4.2.1 Protocol

We evaluated our database with a TD population. To evaluate the videos, we set up an online questionnaire. For each video, participants were asked to evaluate the recognized emotion (anger, happiness, fear, or sadness) using a five-point Likert scale (1 = totally disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, and 5 = totally agree). First, participants evaluated the facial-only (FO) videos, then the body-only (BO) videos, and finally, the body and facial (BF) videos. Within conditions, the videos were in a randomized order and combined all platforms. Due to the characteristics of the system used for the questionnaire, we were not able to randomize the videos for each participant. Therefore, each participant had the same sequence of videos to evaluate. All participants were also asked to fill out the AASP questionnaire (Brown and Dunn 2002). Participants were recruited through mailing lists from our universities

(ENSTA-ParisTech, LIMSI-CNRS, and Université Paris-Sud - UFR STAPS). It was an internet form, and no compensation was given.

4.2.2 Participants

Sixty-four participants answered our online questionnaire (31 females, age 28.23 ± 8.31, 62.5 % with a technological background).

4.2.3 Results

Data analysis We let LSS denote the Likert-scale score, which is the mean value of the scores of the Likert-scale for an emotion for a given condition/test platform. It provided a score on a five-point scale. We used the LSS to evaluate our videos. We obtained four scores for participants by video, corresponding to the four emotions. We considered that an emotion was recognized when it obtained a score over three. An emotion was correctly identified when the correct emotion was identified in the video. An emotion was confused with another emotion when the emotion identified in the video was not the emotion expressed. We worked with ordinal data from the Likest-scale, on three or four independent groups, i.e., the conditions, emotions, and platforms. Therefore, we used the Kruskal-Wallis test for multiple group comparisons. A post-hoc test was performed with the Wilcoxon Rank Sum test for pairwise comparisons.

Facial-only (FO) emotion recognition As shown in Fig. 2, in condition FO, all of the emotions are correctly identified for all of the platforms. Fear expressed only through facial features was confused with sadness on the Mary platform. Regardless of this fact, the recognition score for fear (M = 3.42, $SD = 1.36$) was greater than the recognition score for sadness ($M = 3.20$, $SD = 1.48$) on the Mary platform.

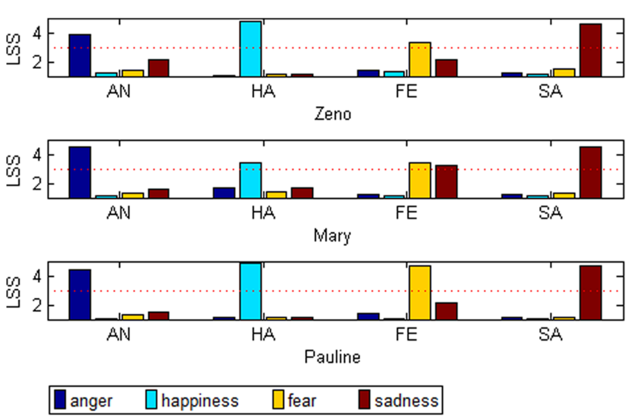


Fig. 2 Emotion recognition and confusion Likert-scale score (LSS) for each facial animation (facial-only (FO) condition). For each animation (AN#; HA#; FE#; SA#), the mean recognition score for each emotion (anger, happiness, fear, sadness) is displayed

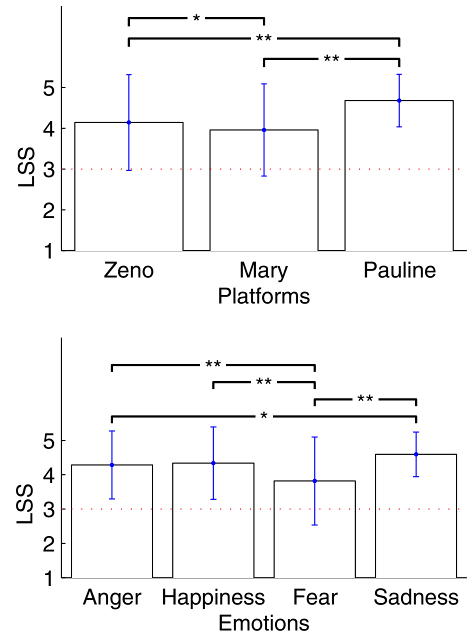


Fig. 3 Correct recognition scores of the platforms (*top*) and emotions (*bottom*) in the facial-only condition. $*p < 0.01$; $**p < 0.001$

We found a significant difference in the correct recognition scores between the platforms in the FO condition ($\chi^2(2) = 82.27$; $p < 0.001$), as shown in Fig. 3. The correct recognition scores were significantly different between Zeno and Mary ($Z = 2.86$; $p < 0.01$), Zeno and Pauline ($Z = -6.09$; $p < 0.001$), and Mary and Pauline ($Z = -9.08$; $p < 0.001$). Emotions were more well recognized on Pauline ($M = 4.68$, $SD = 0.64$), than on Zeno ($M = 4.14$, $SD = 1.17$), than on Mary ($M = 3.96$, $SD = 1.13$).

We found a significant difference in the correct recognition scores between the emotions in the FO condition ($\chi^2(3) = 49.01$; $p < 0.001$), as shown in Fig. 3. Fear was significantly different than anger ($Z = 3.80$; $p < 0.001$), happiness ($Z = 4.71$; $p < 0.001$), and sadness ($Z = -6.69$; $p < 0.001$). This emotion was the least recognized in the FO condition, with a correct recognition score of $M = 3.82$, $SD = 1.3$ (anger: $M = 4.29$, $SD = 0.99$; happiness: $M = 4.34$, $SD = 1.06$; sadness: $M = 4.59$, $SD = 0.65$). Anger was significantly different than sadness ($Z = -3.17$; $p < 0.01$).

Body-only (BO) emotion recognition The recognition and confusion scores of the videos in the BO condition can be seen in Fig. 4. All of the emotions were recognized on Nao, and no confusion was present, contrary to the other platforms. For Zeno, only sadness was recognized and not confused with the other emotions. For example, the AN3 (anger) behavior was confused with the happiness emotion. Mary's and

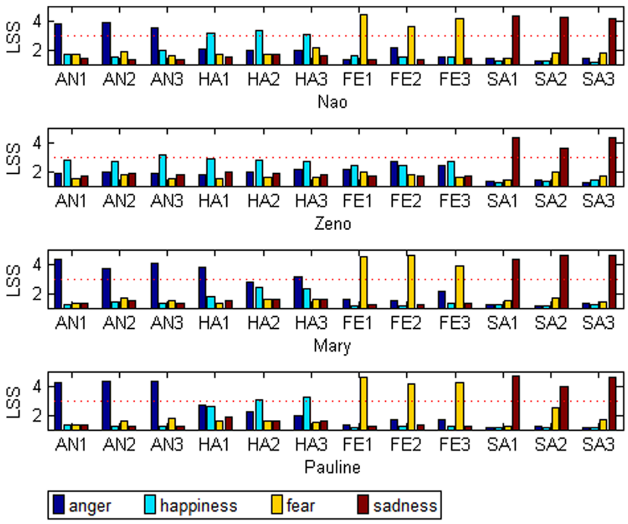


Fig. 4 Emotion recognition and confusion LSS for each body animation (body-only (BO) condition). For each animation (AN#; HA#; FE#; SA#), the mean recognition score for each emotion (anger, happiness, fear, sadness) is displayed

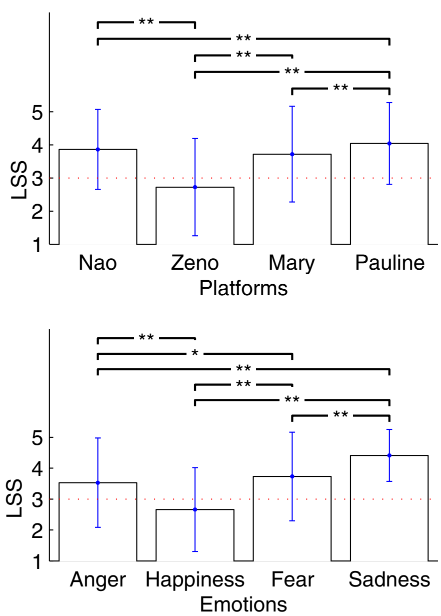


Fig. 5 Correct recognition scores of the platforms (*top*) and emotions (*bottom*) in body-only condition. $*p < 0.01$; $**p < 0.001$

Pauline's anger, fear, and sadness emotions were correctly identified. However, on Mary, happiness was confused with anger for HA1 (happiness), and not identify for HA2 (happiness) and HA3 (happiness) (the score for anger was greater than the score for happiness for these two animations). For Pauline, HA2 and HA3 were correctly identified, and HA1 was identified as showing no emotion.

We found a significant difference in the correct recognition scores between the platforms in the BF condition ($\chi^2(3) = 377$; $p < 0.001$), as shown in Fig. 5. Zeno's correct recognition score was significantly different than those of Nao ($Z = -15.19$; $p < 0.001$), Mary ($Z = -13.22$; $p < 0.001$), and Pauline ($Z = -17.46$; $p < 0.001$). The emotions on this platform were recognized less in the BO condition, with a correct recognition score of $M = 2.72$, $SD = 1.46$ (Nao: $M = 3.86$, $SD = 1.21$; Mary: $M = 3.72$, $SD = 1.44$; Pauline: $M = 4.04$, $SD = 1.23$). Pauline's correct recognition score was significantly different than those of Nao ($Z = 4.17$; $p < 0.001$), Zeno ($Z = 17.46$; $p < 0.001$), and Mary ($Z = 4.16$; $p < 0.001$). The emotions on this platform were the most recognized in the BO condition.

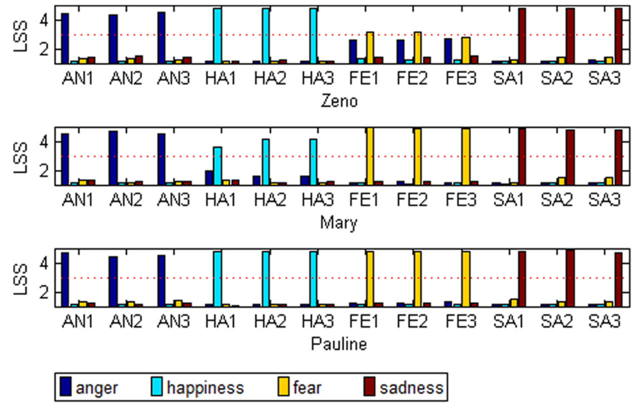


Fig. 6 Emotion recognition and confusion LSS for each combination of body and facial animation (body and facial (BF) condition). For each animation (AN#; HA#; FE#; SA#), the mean recognition score for each emotion (anger, happiness, fear, sadness) is displayed

We found a significant difference in the correct recognition scores between the emotions in the BF condition ($\chi^2(3) = 625$; $p < 0.001$), as shown in Fig. 5. Anger was significantly different than happiness ($Z = 12.32$; $p < 0.001$), fear ($Z = -3.20$; $p < 0.01$), and sadness ($Z = -12.75$; $p < 0.001$). Happiness was significantly different than anger, fear ($Z = -24.81$; $p < 0.001$), and sadness ($Z = -24.81$; $p < 0.001$). Fear was significantly different than anger, happiness, and sadness ($Z = -9.22$; $p < 0.001$). Sadness was significantly different than anger, happiness, and fear. Sadness was the most recognized emotion ($M = 4.41$, $SD = 0.84$), followed by fear ($M = 3.73$, $SD = 1.43$), then by anger ($M = 3.53$, $SD = 1.44$); happiness was the least recognized emotion ($M = 2.66$, $SD = 1.36$).

Body and facial (BF) emotion recognition The recognition and confusion scores of the videos in the BF condition are shown in Fig. 6. Out of the 36 videos presented in the BF condition, 35 videos were correctly identified and no confusion was present. The animation FE3 for Zeno was not well identified. Overall, fear animations of Zeno were weakly recognized (recognition score of fear in fear animation was $M = 2.99$, $SD = 1.49$ on Zeno), and it was confused with anger (recognition score of anger in fear animation was $M = 2.60$, $SD = 1.53$ on Zeno).

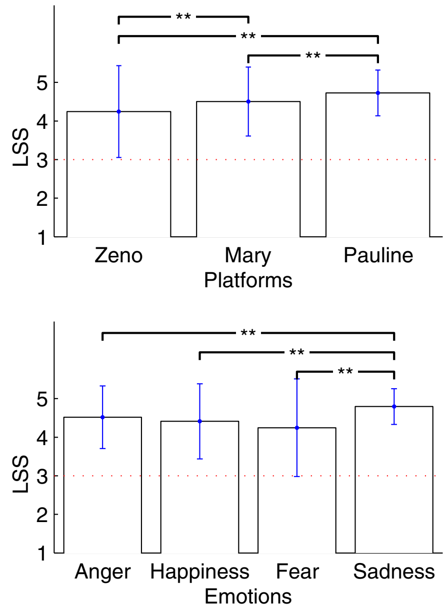


Fig. 7 Correct recognition scores of the platforms (*top*) and emotions (*bottom*) in the body and facial condition. $*p < 0.01$; $**p < 0.001$

We found a significant difference in the correct recognition scores between the platforms in the BF condition ($\chi^2(2) = 73.53$; $p < 0.001$), as shown in Fig. 7. The correct recognition scores were significantly different between Zeno and Mary ($Z = -3.75$; $p < 0.001$), Zeno and Pauline ($Z = -8.53$; $p < 0.001$), and Mary and Pauline ($Z = -5.01$; $p < 0.001$). Emotions were more well recognized on Pauline ($M = 4.73$, $SD = 0.59$), than on Mary ($M = 4.50$, $SD = 0.89$), than on Zeno ($M = 4.24$, $SD = 1.2$).

We found a significant difference in the correct recognition scores between the emotions in the BF condition ($\chi^2(3) = 74.24$; $p < 0.001$), as shown in Fig. 7. Sadness was significantly different than anger ($Z = 6.47$; $p < 0.001$), happiness ($Z = 7.72$; $p < 0.001$), and fear ($Z = 7.48$; $p < 0.001$). This emotion was least recognized in the BF condition, with a correct recognition score of $M = 4.79$, $SD = 0.46$ (anger: $M = 4.52$, $SD = 0.81$; happiness: $M = 4.41$, $SD = 0.97$; sadness: $M = 4.24$, $SD = 1.27$).

Comparison of conditions, platforms, and emotions. We found a significant difference in the correct recognition scores between the conditions (FO, BO, BF) ($\chi^2(2) = 733$; $p < 0.001$), as shown in Fig. 8. The correct recognition scores were significantly different between conditions BO and BF ($Z = 26.42$; $p < 0.001$), BO and FO ($Z = -12.14$; $p < 0.001$), and FO and BF ($Z = -7.07$; $p < 0.001$). Emotions in the BF condition were the most recognized ($M = 4.49$, $SD = 0.95$), followed by emotions in FO

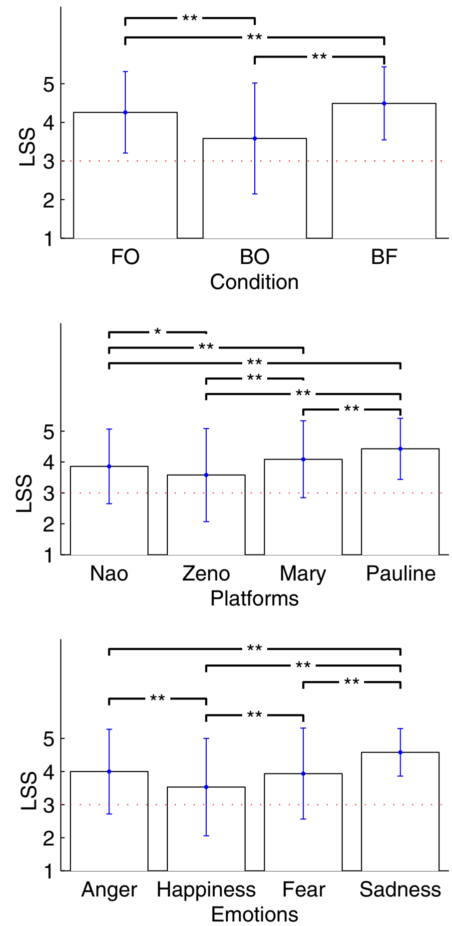


Fig. 8 Correct recognition scores of the conditions (*top*), platforms (*middle*) and emotions (*bottom*). $*p < 0.01$; $**p < 0.001$

($M = 4.30$, $SD = 1.06$); emotions in the BO condition were the least recognized ($M = 3.58$, $SD = 1.43$).

We found a significant difference in the correct recognition scores between the platforms (Nao, Zeno, Mary, Pauline) ($\chi^2(3) = 389$; $p < 0.001$), as shown in Fig. 8. The correct recognition scores were significantly different between Nao and Zeno ($Z = 2.58$; $p < 0.01$), Nao and Mary ($Z = -6.46$; $p < 0.001$), and Nao and Pauline ($Z = -13.88$; $p < 0.001$). The correct recognition scores were significantly different between Zeno and Mary ($Z = -10.08$; $p < 0.001$) and Zeno and Pauline ($Z = -18.19$; $p < 0.001$). The correct recognition scores were significantly different between Mary and Pauline ($Z = -8.92$; $p < 0.001$). Emotions were more well recognized on Pauline ($M = 4.43$, $SD = 0.99$), than on Mary ($M = 4.09$, $SD = 1.25$), than on Nao ($M = 3.86$, $SD = 1.21$). Emotions on Zeno were recognized the least ($M = 3.58$, $SD = 1.51$).

We found a significant difference in the correct recognition scores between emotions (anger, happiness, fear, sadness) ($\chi^2(3) = 483$; $p < 0.001$), as shown in Fig. 8. Sadness was significantly different than anger ($Z = -13.48$; $p < 0.001$), happiness ($Z = -21.89$; $p < 0.001$), and fear ($Z = -13.11$; $p < 0.001$). Happiness was significantly different than anger ($Z = 9.16$; $p < 0.001$), fear ($Z = -8.59$; $p < 0.001$), and sadness. Sadness was the most recognized emotion ($M = 4.58$, $SD = 0.72$), followed by fear ($M = 3.93$, $SD = 1.38$) and anger ($M = 3.99$, $SD = 1.28$). Happiness was the least recognized emotion ($M = 3.53$, $SD = 1.47$).

4.3 Conclusion

For the 96 videos tested, emotions were correctly identified in 82 videos (12 out of 12 in the FO condition, 35 out of 48 in the BO condition, and 35 out of 36 in the BF condition). Pauline's animations were correctly identified most of the time, followed by Mary's, Nao's, and Zeno's animations. Zeno showed the lowest recognition score, even though it was expressing facial features. Sadness was the best recognized emotion, and happiness was the least recognized emotion. Emotion recognition results were higher in conditions in which the face expressed emotion (FO and BF conditions) than in the condition where only the body expressed the emotion (BO condition). This highlights the importance of facial expressions for recognizing the category of emotion, as observed by Buisine et al. (2014) and Meeren et al. (2005). Sadness was the easiest emotion to recognize, followed by fear and anger. Happiness was the most difficult. This same pattern was reported by De Gelder and Van den Stock (2011).

5 Correspondence between AASP quotient and emotion recognition on multiple test platforms in the TD population

5.1 Objectives

As explained in Sect. 3.1, a correspondence between the AASP patterns and postural behavior was observed. We searched for a correspondence between the sensory preferences of a TD individual and his/her abilities to recognize emotions in order to validate our **H1.1**, **H1.2**, and **H2** hypotheses.

5.2 Method

As mentioned in Sect. 4.2.1, all of the individuals that participated in the emotion recognition survey completed the AASP questionnaire. We searched for a relationship between the emotion rating scores of the participants and their sensory profiles. We selected the MSS and VSS as markers of the proprioceptive and visual dependency of an individual, as described in Sect. 3.1. For MSS, we divided our participants into two groups: *MSSLow* (29 participants), grouping individuals with low MSS scores, suggesting a weak reliance on proprioceptive cues, and *MSSHigh* (35 participants), grouping individuals with high MSS scores, suggesting a strong reliance on proprioceptive cues. Furthermore, we divided our participants symmetrically in two groups in VSS: *VSSLow* (45 participants), grouping individuals with low VSS scores, suggesting a weak reliance on visual cues, and *VSSHigh* (19 participants), grouping individuals with high VSS scores, suggesting a strong reliance on visual cues.

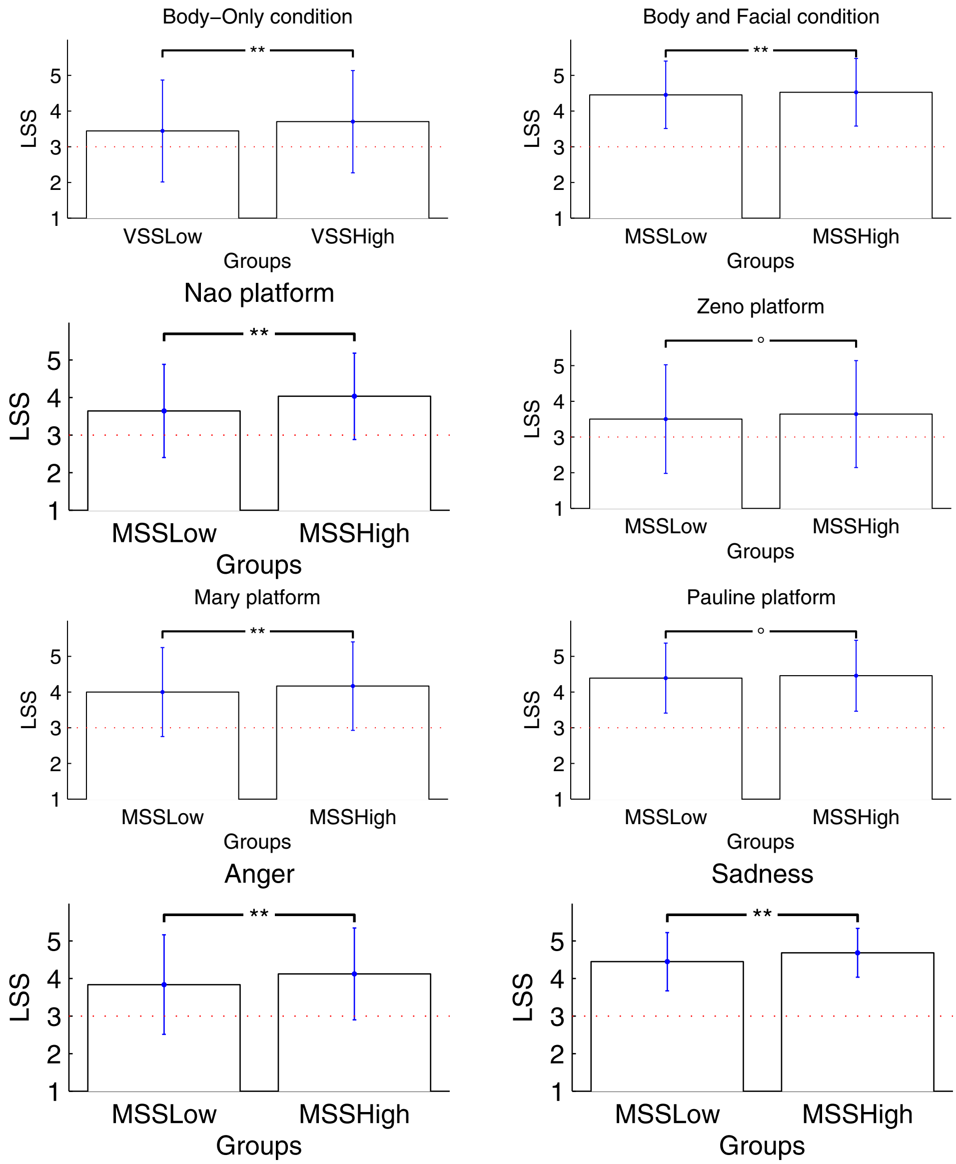
5.3 Data analysis

We worked with ordinal data from the Likest-scale, on two independent groups (MSSLow vs. MSSHigh and VSSLow vs. VSSHigh). Therefore, we used the Wilcoxon Rank Sum test for pairwise comparisons.

5.4 Results

Movement sensory sensitivity (MSS) We found significant differences in recognition scores between MSSLow and MSSHigh in the BO condition ($Z = -6.04$; $p < 0.001$). MSSHigh had higher results ($M = 3.70$, $SD = 1.43$) than MSSLow ($M = 3.44$, $SD = 1.43$). We found significant differences in recognition scores between MSSLow and MSSHigh in the BF condition ($Z = -3.21$; $p < 0.01$). MSSHigh had higher results ($M = 4.52$, $SD = 0.95$) than MSSLow ($M = 4.45$, $SD = 0.94$). We found significant differences in recognition scores between MSSLow and MSSHigh for Nao ($Z = -4.98$; $p < 0.001$). MSSHigh had higher results ($M = 4.03$, $SD = 1.15$) than MSSLow ($M = 3.64$, $SD = 1.24$). Moderate evidence of a difference in recognition scores between MSSLow and MSSHigh for Zeno was found ($Z = -2.20$; $0.01 < p < 0.05$). MSSHigh had higher results ($M = 3.64$, $SD = 1.15$) than MSSLow ($M = 3.5$, $SD = 1.52$). We found significant differences in recognition scores between MSSLow and MSSHigh for Mary ($Z = -4.28$; $p < 0.001$). MSSHigh had higher results ($M = 4.16$, $SD = 1.24$) than MSSLow ($M = 4.00$, $SD = 1.25$). Moderate evidence of a difference in recognition scores between MSSLow and MSSHigh for Pauline was found ($Z = -2.50$; $0.01 < p < 0.05$). MSSHigh had higher results ($M = 4.46$, $SD = 0.99$) than MSSLow ($M = 4.39$, $SD = 0.98$). We found significant differences in recognition scores between MSSLow and MSSHigh for anger ($Z = -5.01$; $p < 0.001$). MSSHigh had higher results ($M = 4.12$, $SD = 1.22$) than MSSLow ($M = 3.84$, $SD = 1.33$). We found significant differences in recognition scores between MSSLow and MSSHigh for sadness ($Z = -7.56$; $p < 0.001$). MSSHigh had higher results

Fig. 9 Correct recognition scores for the MSSLow and MSSHigh groups. ° $0.01 < p < 0.05$; $^{*}p < 0.01$; $^{**}p < 0.001$



($M = 4.68$, $SD = 0.65$) than MSSLow ($M = 4.45$, $SD = 0.78$).
All these results are reported in Fig. 9.

Vision sensory sensitivity (VSS) We found significant differences in recognition scores between VSSLow and VSSHigh in the BO condition ($Z = 4.66$; $p < 0.001$). VSSHigh had higher results ($M = 3.76$, $SD = 1.39$) than VSSLow ($M = 3.51$, $SD = 1.45$). We found significant differences in recognition scores between VSSLow and VSSHigh for Nao ($Z = 2.78$; $p < 0.01$). VSSHigh had higher results ($M = 4.02$, $SD = 1.16$) than VSSLow ($M = 3.79$, $SD = 1.22$). We found significant differences in recognition scores between VSSLow and VSSHigh for Pauline ($Z = 3.40$; $p < 0.001$). VSSHigh had higher results ($M = 4.55$, $SD = 0.86$) than VSSLow ($M = 4.37$, $SD = 1.03$). We found significant differences in recognition scores between VSSLow and VSSHigh for happiness ($Z = 4.44$; $p < 0.001$). VSSHigh had higher results ($M = 3.78$, $SD = 1.37$) than VSSLow ($M = 3.42$, $SD = 1.50$).
All these results are reported in Fig. 10.

5.5 Conclusions and discussion

Our results emphasize significantly different behaviors for the tested individuals as a function of their MSS and VSS scores.

The MSSHigh group had higher abilities in recognizing the expression of emotions compared to the MSSLow group in the BO and BF conditions. We also found significant differences in the recognition level between platforms by comparing the MSS groups. The MSSHigh group performed

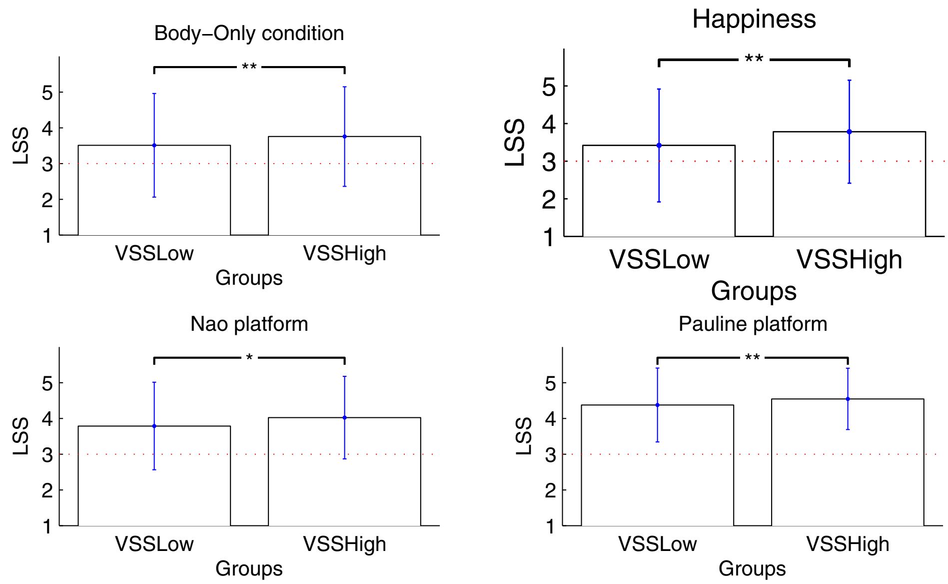


Fig. 10 Correct recognition scores of the body-only condition (*top-left*), the emotion happiness (*top-right*), and the platforms Nao (*bottom-left*) and Pauline (*bottom-right*) for the VSSLow and VSSHigh groups. $*p < 0.01$; $**p < 0.001$

better than MSSLow for the Nao and Mary platforms. The MSSHigh group had a higher recognition score for anger and sadness than the MSSLow group. Our results go against our hypothesis **H1.1**, as individuals with a high reliance on proprioceptive cues had better recognition scores.

Results for the VSS groups showed that VSSHigh participants had higher abilities than VSSLow participants in recognizing expression of emotions compared to the MSSLow group in the BO condition. We also found by comparing the VSS groups' significant differences in the recognition level between platforms that the VSSHigh group performed better than the VSSLow group for the Nao and Pauline platforms. The VSSHigh group had higher recognition scores in happiness than in VSSLow. As discussed in Sect. 4, happiness was the most difficult emotion to recognize. These results validate our **H2** hypothesis on a TD population.

6 Emotion recognition on multiple platforms by individuals with ASD

6.1 Objectives

In this section, we investigate the relationship between the individual sensory preferences defined in Sect. 3.1 of individuals with ASD and their abilities to recognize emotions (validation of the **H1.1**, **H1.2**, and **H2** hypotheses for the ASD population).

6.2 Subset selection of the best videos for ASD

The evaluation of emotion recognition on individuals with ASD was done on a subset of videos of the database. We agreed with the caretakers that evaluating 96 videos would be a difficult task, especially for individuals with ASD.

We selected the videos with the highest recognition scores obtained among TD individuals (see Sect. 4.2) for each emotion in each condition for each platform, resulting in a 40 videos subset. The selected videos were as follows:

- FO: all of the videos were selected
- BO:
  - Nao: AN3, FE3, HA2, SA1
  - Zeno: AN2, FE1, HA1, SA1
  - Mary: AN2, FE3, HA3, SA1
  - Pauline: AN1, FE2, HA3, SA1
- BF:
  - Zeno: AN3, FE1, HA2, SA1
  - Mary: AN2, FE3, HA2, SA1
  - Pauline: AN2, FE3, HA2, SA2

6.3 Method

We developed software to evaluate individuals' skills to recognize bodily and facial expressions of emotion on different supports with increasing visual complexity. The software consisted of a tactile computer game, as shown in Fig. 11. In the game, a video was shown to the participants. Then, the participants were asked to categorize the emotion recognized in the video with a forced choice. The choices were displayed at the bottom of the screen, with Picture Exchange Communication System (PECS) images (Bondy and Frost 1994) and labels for anger, happiness, fear, and sadness. If desired, the participant could replay the video. The possibility to skip to the next trial without answering the question was not allowed because we feared that some of the participants would have skipped all of the videos without trying to recognize any

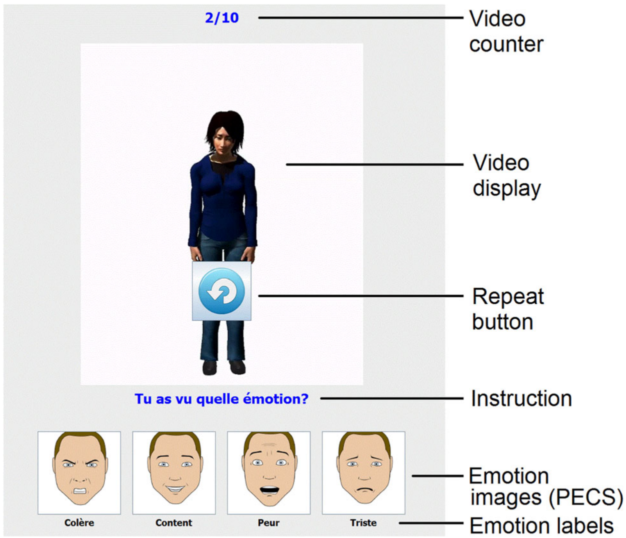


Fig. 11 Screen shot of the game at the end of a video. Emotion labels (written in French) are, from *left* to *right* anger, happiness, fear, and sadness. The instruction, Tu as vu quelle émotion ?" means "Which emotion did you see?"

emotion at all. However, if we noticed that the participant was lost or verbalized his/her inability to answer, the experimenter could skip the current video and move on to the next video. The software recorded the time between the launch of a video and the choice of an emotion. We divided our 40 videos (see Subsect. 6.2) into four sessions of ten videos to be able to take breaks between sessions. The task was quite demanding and required high cognitive levels of concentration from the participants (in our case, people suffering from ASD); hence, the breaks were necessary. We divided our 40 videos into sessions of ten videos to stay consistent across the sessions. A counter was displayed at the top of the screen to inform the participant of his/her progress within the current subset of videos. Before performing the game, the procedure was explained to each participant to ensure that the instructions were correctly understood. During the tests, the experimenter and the caretaker were as neutral as possible so as not to interfere with participants' recognition skills. To do so, they did not help the participants or correct them in the choice of their answers.

First, participants evaluated the FO condition, followed by the BO condition, and finally, the BF condition. Within conditions, the videos were presented in a randomized order and combined all platforms, i.e., the Nao, Zeno, Mary, and Pauline, animations were randomized. The order of the videos was randomized for each participant.

6.4 Participants

All participants with ASD (Sect. 3.1) performed the experiment, except for participant AD3; he was not present in the care facility when the experiment was conducted.

6.5 Data analysis

We used Fisher's Exact Test analysis on the recognition scores of our participants since we worked on small samples (groups of three to five individuals).

6.6 Results

6.6.1 Emotion recognition by individuals with ASD

We computed the scores for each participant as a percentage. Each answer is listed in Table 3 in annex. Overall, eight of the participants obtained a score below 50 %. By looking more closely at the score for each emotion (see Table 3), notice that some participants recognized one emotion more than the other emotions (e.g., participant AD2 and participant AD1 recognized happiness better than the other emotions). The forced choice may be the cause of this phenomenon because the participant chose a "default emotion" before moving on to the next video.

We found significant differences ($p < 0.01$) in recognition scores between FO and BO and between BO and BF, as shown in Fig. 12. The FO and BF conditions exhibited better performance than the BO condition (FO: $M = 50\%$, $SD = 29.30\%$; BO: $M = 34.90\%$, $SD = 22.37\%$; BF: $M = 51.39\%$, $SD = 35.50\%$). We also found this pattern in TD individuals' results.

The recognition score of fear was significantly different than that of happiness ($p < 0.01$) and sadness ($p < 0.01$). Fear was recognized less ($M = 32.5\%$, $SD = 29.27\%$) than happiness ($M = 54.17\%$, $SD = 29.37\%$) and sadness ($M = 49.17\%$, $SD = 30.29\%$). We found moderate evidence of a difference between happiness and anger ($0.01 < p < 0.05$), as shown in Fig. 12. Anger ($M = 41.67\%$, $SD = 30.40\%$) was recognized less than happiness.

No significant difference in recognition scores for all participants was found between the platforms.

We found a significant difference between G1 and G2 ($p < 0.01$) and moderate evidence of a difference between G1 and G3 ($0.01 < p < 0.05$) on the recognition score of the whole database, as shown in Fig. 13. Participants from G1 had lower recognition scores ($M = 36\%$, $SD = 12.07\%$) than participants from G2 and G3 (51.7%, $SD = 39.71\%$ and $M = 49.38\%$, $SD = 25.53\%$, respectively).

Significant differences ($p < 0.01$) between G1 and G2 and between G1 and G3 were found on the recognition scores for Mary, as shown in Fig. 13. Participants from G1 had lower recognition scores for Mary ($M = 26.7\%$, $SD = 16.02\%$) than participants from G2 and G3 ($M = 55.56\%$, $SD = 38.49\%$ and $M = 56.25\%$, $SD = 17.18\%$, respectively).

Moderate evidence of a difference ($0.01 < p < 0.05$) between G1 and G2 and between G1 and G3 was found on the recognition scores of the fear emotion,, as shown in Fig.

Table 3 Participants' answers for all videos and correct percentage scores for each platform in each condition

	Platform	Emotion	AD2	AD4	AD6	CH3	CH5	AD1	CH1	CH4	AD5	AD7	CH2	CH6
			G1	G1	G1	G1	G1	G2	G2	G2	G3	G3	G3	G3
Condition FO	Zeno	an	ha	ha	sa	sa	fe	sa	sa	an	ha	sa	an	an
		ha	ha	ha	sa	ha	ha	fe	sa	ha	an	ha	ha	ha
		fe	ha	an	sa	ha	fe	an	an	fe	an	ha	fe	fe
		sa	ha	fe	sa	an	sa	ha	sa	sa	–	ha	fe	sa
		correct (%)	25	25	25	25	75	0	25	100	0	25	75	100
	Mary	an	ha	fe	an	ha	an	an	an	an	an	an	an	an
		ha	ha	sa	sa	ha	an	sa	sa	ha	fe	sa	sa	ha
		fe	ha	an	ha	sa	an	sa	fe	fe	ha	ha	fe	an
		sa	ha	an	ha	sa	ha	sa	sa	sa	ha	an	sa	sa
		correct (%)	25	0	25	50	25	50	75	100	25	25	75	75
	Pauline	an	ha	sa	sa	an	an	sa	an	an	fe	fe	an	an
		ha	ha	sa	sa	ha	fe	ha	ha	ha	an	sa	ha	ha
		fe	ha	ha	an	fe	fe	fe	sa	fe	fe	sa	fe	fe
		sa	ha	sa	sa	sa	sa	ha	sa	sa	sa	sa	sa	sa
		correct (%)	25	25	25	100	75	50	75	100	50	25	100	100
Condition BO	Nao	an	ha	sa	fe	fe	an	ha	ha	an	ha	an	sa	an
		ha	ha	ha	sa	sa	sa	ha	an	ha	sa	sa	sa	ha
		fe	ha	sa	sa	an	an	ha	sa	fe	ha	ha	sa	fe
		sa	ha	sa	sa	sa	sa	ha	sa	sa	sa	sa	sa	an
		correct (%)	25	50	25	25	50	25	25	100	25	50	25	75
	Zeno	an	ha	an	ha	sa	ha	ha	ha	an	fe	–	sa	ha
		ha	ha	ha	sa	ha	fe	ha	fe	sa	an	ha	sa	ha
		fe	ha	sa	ha	ha	sa	ha	sa	fe	ha	fe	sa	an
		sa	ha	fe	sa	ha	sa	ha	ha	sa	fe	ha	sa	ha
		correct (%)	25	50	25	25	25	25	0	75	0	50	25	25
	Mary	an	ha	an	sa	ha	fe	ha	sa	an	fe	sa	sa	an
		ha	ha	sa	sa	sa	sa	ha	sa	ha	ha	ha	sa	fe
		fe	ha	sa	sa	ha	an	ha	an	fe	an	fe	sa	ha
		sa	ha	sa	an	fe	an	ha	ha	sa	sa	ha	sa	sa
		correct (%)	25	50	0	0	0	25	0	100	50	50	25	50
	Pauline	an	ha	an	an	fe	sa	ha	an	an	sa	an	sa	fe
		ha	ha	sa	ha	sa	sa	ha	–	ha	ha	an	sa	an
		fe	ha	fe	sa	fe	an	ha	sa	fe	an	an	fe	–
		sa	ha	sa	an	ha	an	ha	fe	sa	ha	ha	sa	sa
		correct (%)	25	75	50	25	0	25	25	100	25	25	50	25
Condition BF	Zeno	an	ha	an	ha	an	an	an	ha	an	fe	ha	an	an
		ha	ha	ha	sa	ha	an	ha	ha	ha	sa	an	ha	ha
		fe	ha	sa	ha	ha	sa	ha	an	fe	an	–	fe	fe
		sa	sa	sa	sa	an	sa	an	an	sa	fe	ha	sa	sa
		correct (%)	50	75	25	50	50	50	25	100	0	0	100	100
	Mary	an	ha	an	ha	ha	an	ha	ha	an	sa	an	an	an
		ha	ha	ha	sa	an	an	ha	sa	ha	ha	ha	ha	ha
		fe	fe	fe	sa	sa	an	ha	fe	fe	fe	sa	fe	fe
		sa	ha	sa	ha	sa	fe	ha	ha	sa	ha	ha	sa	sa
		correct (%)	50	100	0	25	25	25	25	100	50	50	100	100
	Pauline	an	ha	an	ha	sa	an	ha	fe	an	ha	–	an	an

Table 3 continued

Platform	Emotion	AD2	AD4	AD6	CH3	CH5	AD1	CH1	CH4	AD5	AD7	CH2	CH6
		G1	G1	G1	G1	G1	G2	G2	G2	G3	G3	G3	G3
	ha	ha	ha	ha	ha	fe	ha	sa	ha	ha	fe	ha	ha
	fe	ha	fe	ha	fe	an	ha	an	fe	an	–	fe	fe
	sa	ha	sa	ha	an	an	ha	an	sa	ha	ha	sa	sa
	correct (%)	25	100	25	50	25	25	0	100	25	0	100	100

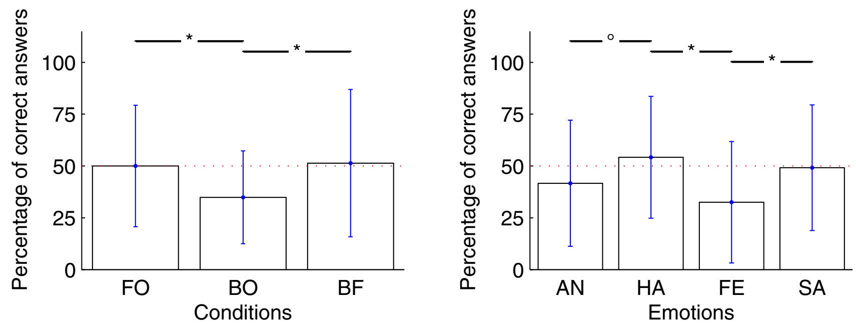


Fig. 12 Percentage of correct recognition scores for conditions Facial-Only (FO), Body-Only (BO), and Body and Facial (BF) (left) and for each emotion (rigth) for the whole database. °0.01 $< p <$ 0.05; *$p <$ 0.01

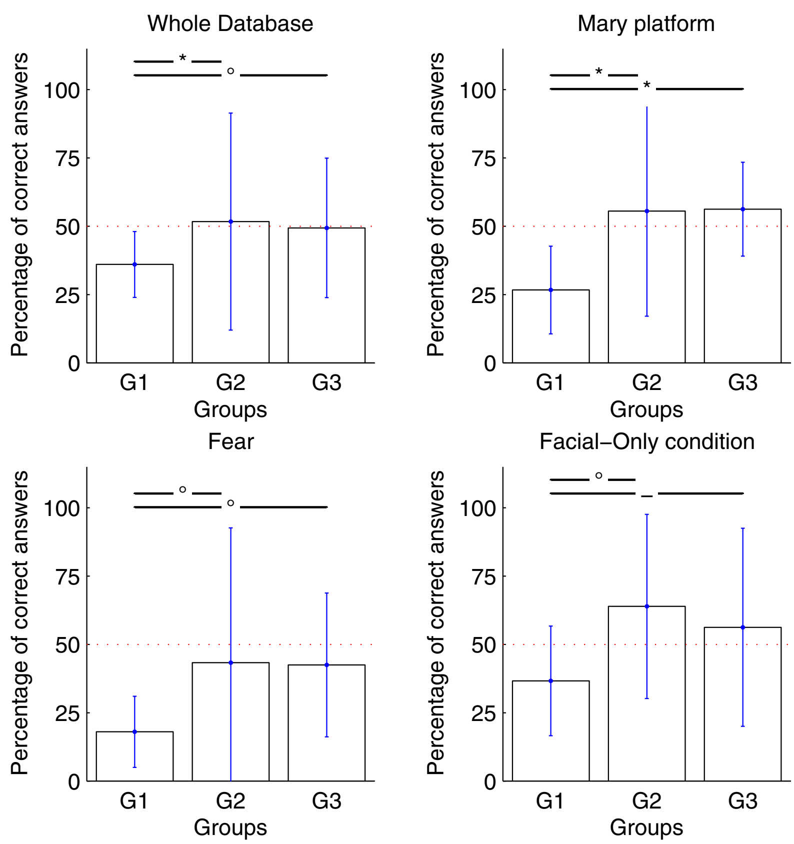


Fig. 13 Percentage of correct recognition scores for the whole database (*top-left*), on the platform Mary (*top-right*), on the emotion Fear (bottom-left) and for the Facial-Only condition (*bottom-right*) for groups G1, G2 and G3. °0.05 $< p <$ 0.1; - 0.01 $< p <$ 0.05; *$p <$ 0.01

13. Participants from G1 had lower recognition scores for fear (M = 18%, SD = 13.03%) than participants from G2 and G3 (M = 43.33%, SD = 49.33% and M = 42.50%, SD = 26.30%, respectively).

Moderate evidence of a difference (0.01 $< p <$ 0.05) between G1 and G2 and weak evidence of a difference (0.05 $< p <$ 0.1) between G1 and G3 was found on the recognition scores of the FO condition, as shown in Fig. 13. Participants from G1 had lower recognition scores in the FO condition (M = 36.67%, SD = 20.07%) than participants from G2 and G3 (M = 63.89%, SD = 33.68% and M = 56.25%, SD = 36.24%, respectively).

We found a significant difference between child and adult participants ($p <$ 0.001), as shown in Fig. 14. Children per-

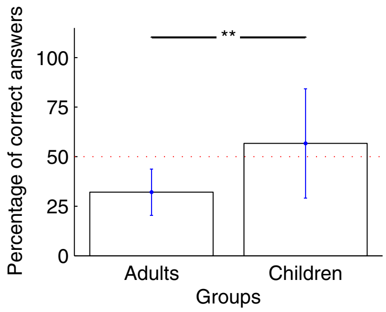


Fig. 14 Percentage of correct recognition scores for child participants and adult participants for the whole database. $^{**}p < 0.001$

formed better than adults (Children: $M = 32.08\%$, $SD = 11.66\%$; Adults: $M = 56.67\%$, $SD = 27.60\%$). Three out of six children (CH4, CH2, and CH6) obtained a global score of more than 50 %, whereas only one adult achieved this result (AD4).

6.6.2 Participants' scores and behaviors

Below, we describe the behavior and scores of each participant with ASD. Participant AD3 was not in the care facility during the setup of the task. A brief description of the participants is provided in Sect. 3.1. Most of the participants found that the task was difficult.

AD1 (G2 group): In the BO and BF conditions, she chose the answer "Happy" most of the time. Nevertheless, she seemed to understand the game's instructions.

AD2 (G1 group): She selected happiness for 38 of the 40 videos. When we noticed this trend during the break after the first session, we checked to see if she understood the task and asked her to describe to us the movements she saw on the screen. She was able to describe them and understood the task. Nevertheless, she always chose the happy emotion, except for the last two videos where she chose the correct emotions, which were not happiness. For the duration of the experiment, she showed great enthusiasm, but started to be more quiet at the end, which is when she chose the correct emotions. After careful discussion with the caretakers, we determined that she was really happy to perform the task with us, as she liked to be the center of attention. Moreover, she always smiled or said statements with a happy connotation. Her answers may thus have been influenced by her own affective state (happy).

AD4 (G1 group): She showed great enthusiasm for performing the task. She had difficulties with the FO condition (16.7 % good recognition), but displayed higher scores in the BO and BF conditions (BO: 56.3 %; BF: 91.7 %). She obtained the best scores among the G1 group.

AD5 (G3 group): He was recovering from the flu, but performed the task with care. However, it was clear that he was really tired at the end of the task. He showed great difficulty in understanding Zeno's emotions (0 % of good recognition).

AD6 (G1 group): He had difficulties understanding the task. The caretaker had to help him touch the screen. Since he suffers from echolalia, we think that he only repeated what we were asking him when he did not know what to choose (i.e., when he did not seem to answer, we asked him, "What do you think the character shows? Anger? Happiness? Fear? Sadness?" and he repeated, "Sadness," as it was the last word of our sentence). However, he answered some of the items without hesitation, showing that he nevertheless partially understood the instructions.

AD7 (G3 group): He was always struggling while choosing an emotion to evaluate the videos (a common symptom in ASD). It took him a long time to perform the task.

CH1 (G2 group): He was really reluctant to perform the task. This can be explained by the facts that (1) he was coming down with a flu, (2) a recent change in his referent caretakers occurred, which made him reluctant to many tasks.

CH2 (G3 group): He understood the task and performed it calmly. Overall, he had a good recognition score (67.5 %). He had more difficulties with the BO condition than with the other conditions (31.3 %) and had more difficulty with the Nao platform than with the other platforms (25 %).

CH3 (G1 group): He understood the task. However, at the end, he seemed to "accelerate" the procedure by not properly choosing the correct image. He showed good recognition scores in the FO condition and with the real human platform.

CH4 (G3 group): He succeeded in correctly recognizing the emotions from 39 of the videos. However, his process in understanding the emotion in the BO condition is noteworthy: after looking at the videos, he did not directly recognize the emotions. He imitated the movement in front of the computer saying, "How do I feel if I do that" and succeeded to understanding them. He presented the best cognition capacity among all of the participants.

CH5 (G1 group): We were unsure of her capacities to perform the task, as she is non-verbal and unfamiliar to this kind of task. However, she understood it and showed an interest in the task (she replayed most of the videos). She had a good emotion recognition rate in the FO condition and for the Nao and Zeno platforms.

CH6 (G3 group): He easily understood the game. This can be explained by the fact that a similar game is performed as part of his weekly therapy. However, the task put him in a state of frustration when he was not able to provide the correct answer, especially in the BO condition.

6.7 Conclusion

A direct relationship between the groups we defined in Sect. 3.1 and the emotion recognition scores was found. We observed that overall, groups G2 and G3 had higher recognition scores than group G1. These results validate our hypotheses **H1.2** and **H2**. Recognition scores for the Mary platform were lower for group G1 than for G2 and G3. We did not find this relationship with the other platforms, therefore, we had no evidence to validate our hypothesis **H1.1**. Recognition scores for the FO condition were lower for group G1 than for G2 and G3. We did not find this relationship with the other conditions.

The results of the task seemed to be led by the motivation and/or condition of the participant. Moreover, the task was described as difficult by the participants, showing their involvement and efforts in performing the task. There seems to be a relationship between the difficulty reported and the inability to properly recognize emotions. This may also explain the overall low emotion recognition scores. Moreover, the groups were small (three participants in group G2, four in group G3, and five in group G1), which made the classification difficult. Overall children performed better than the adults. This can be explained by the fact that they often perform similar tasks with their caretakers (which is not the case for the adult participants).

The recognition scores of happiness were higher than those of other emotions. This can be explain by the behavior of the participants during the experiment. Note that some participants repeatedly chose "Happiness". In addition, the emotion of happiness was well identified by the participants, as it is often used in their daily life. On the other hand, fear, which had the lowest recognition score, was used less by the caretakers in their environment.

Overall, participants with ASD recognized emotions better in the FO and BF conditions than in the BO condition.

7 Conclusion and discussion

The long-term goal of our work is to define individual profiles in order to develop a new personalized robot/virtual-agent-based social interaction for individual with ASD. We worked with 13 participants with ASD from two care facilities. In our previous work (Chevalier et al. 2016), we defined the proprioceptive and visual profiles of our 13 participants. In this work, we established a way to evaluate the capacity of our participants to recognize bodily and/or facial (in isolation or combined) expressions of emotions with respect to their capacity of integration of visual and proprioceptive cues.

First, we developed a database (EMBODI-EMO) of 96 videos expressing four emotions (anger, happiness, fear, and sadness) on four test platforms (human, virtual, robots) of increased visual complexity displayed with different types of embodiment (with or without facial expressions, combined or not with body movements).

Second, we evaluated the EMBODI-EMO database with TD participants. For the 96 videos, the emotions were correctly identified by the TD participants in 82 videos. A smaller subset of the database, composed of 40 selected videos with the best recognition scores for each emotion, for each platform, and for each condition, was built to evaluate individuals with ASD. Indeed, the caretakers stated that evaluating 96 videos would have been a difficult task for individuals with ASD. As in (Meeren et al. 2005; Buisine et al. 2014), we found that the combination of body and facial expressions of emotions enhanced emotion recognition for all participants.

For TD participants, we tested the relationship between one's capacity of emotion recognition with respect to their capacity of integration of visual and proprioceptive cues. We used two items (VSS and MSS) of the AASP questionnaire to describe an individual's response to visual and proprioceptive cues. Results showed a relationship between the profiles and emotion recognition. Results showed that TD individuals that relied more on visual cues had better recognition scores, validating our hypothesis **H2**. However, we found that TD individuals relying on proprioception had better recognition scores, going against our hypothesis **H1.2**.

For participants with ASD, we used AASP scores and reactions to an experimental setup of a Moving Virtual Room to cluster three groups with significantly different behaviors with respect to the way participants with ASD integrated visual and proprioception cues, as described in our previous paper (Chevalier et al. 2016). We validated hypotheses **H1.2** and **H2** on our population with ASD. Indeed, participants with ASD relying more heavily on proprioceptive cues had lower emotion recognition scores on all conditions than participants relying on visual cues.

Mary was the only platform that induced a significant difference in emotion recognition between the groups for participants with ASD. Recognition scores for Mary were lower for participants with ASD relying more heavily on proprioceptive cues than for participants relying on visual cues. This result suggests the importance of embodiment (real vs. virtual) for individuals relying on proprioceptive information.

Zeno showed the lowest recognition level for TD participants. Body animation was difficult on this platform, especially because of the joints degrees of freedom limitation of the arm of this model (R25). This impacted emotion recognition in BO condition on this platform. Excepted sadness, emotions were not correctly recognized on Zeno in BO condition. Emotions were correctly recognized on this platform in FO and BF conditions, suggesting that Zeno can express successfully emotion through Facial features.

For TD participants, happiness was the least recognized emotion, as found in (De Gelder and Van den Stock 2011). But this emotion was the most recognized for participants with ASD. TD individuals relying more on visual cues had more facilities to recognized this emotion, which goes with our hypothesis **H2**. We can try to explain the high recognition score of happiness in individual with ASD by their environment. Indeed, the use of "overacted" happiness often occurs in their daily-life. Very expressive congratulations are used as reinforcement by caretaker and parents when the child with ASD accomplish something.

For TD participants, emotions were recognized more in BF condition than in FO and BO conditions. For participants with ASD, the emotions were recognized more in FO and BF conditions than in BO condition. We did not find difference between FO and BF conditions for this population. In TD participants, visual cues integration from body and face seems to follow a maximum-likelihood estimation. Participants with ASD did not improve their results with the addition of channels (Facial and Body). Participants with ASD would have difficulties to take into account the reliability of bodily and face signals to disambiguate emotional expressions and produce more robust emotions recognition.

Furthermore, the FO condition was the only condition that induced significant differences in emotion recognition between the groups for participants with ASD. Recognition scores for the FO condition were lower for participants with ASD relying more heavily on proprioceptive cues than for participants relying on visual cues. This suggests deficient use of emotional cues expressed by the face of individuals relying on proprioceptive information.

These results informed us that integration of visual and proprioceptive cues of an individual with ASD influence his/her abilities in recognizing emotions. Participants with ASD relying on proprioceptive information had lower recognition scores than the other participants with ASD. We also found that they had more difficulties than the other participants with ASD to recognize emotion when expressed only through facial features, and when expressed by the virtual agent. In future work, these findings can help us to develop personalized interaction. Designers of Human-Machine Interaction for users with ASD may thus select the channels (e.g. facial expression) or the type of embodiment (e.g. robots vs. virtual agent) depending on the user's visual and proprioceptive profile.

Participants with ASD found the task difficult. As observed in their behaviors, AD2 and CH4 had good observation strategies, but had impairments in understanding emotions, similar to the observations in another study (Celani et al. 1999). They showed good observation of the movements performed in the videos. CH4 succeeded in understanding the emotions using imitation, but AD2 was overwhelmed by her own internal state.

We could have chosen to include a context for our videos, as in (Begeer et al. 2006). Here, the authors observed that children with ASD had more difficulties recognizing emotions when they were out of context, but they recognized emotions better when contextual information was provided. In their study, children were shown four pictures of faces and had to match two of them. In the condition with context, the experimenter prompted the child through different sentences such as *"Which two would be most likely to give you a sweet?"* or *"Imagine all of the men in the pictures are teachers. Which two teachers are most likely to tell you off?"* However, in our setup, the inclusion of context for the emotions would have been more difficult. We did not want to induce cues regarding the answer by adding a context or verifying that each participant understood the background story. Several studies pertaining to the recognition of expressions of emotions do not involve any context at all (Baron-Cohen et al. 2001).

In our software, we displayed images for the selection of emotions; however, this possibly had an impact on our participants. They may have tried to match them to the specific facial expression images shown in the video, rather than the more general concept of the emotion. During the design process of the system, we researched the best way for participants to choose one of the labels (happiness, angry, fear, and sadness). We searched for a simplified, easy to understand method for all of our participants. Since only a few of our participants were able to read, we could not simply use written labels; hence, we chose images. The use of pictograms is widely used by individuals with ASD (PECS, for example Bondy and Frost 1994) to describe the planning of the day or their emotional state.

We also imagined showing several types of pictograms of the face, body, and face and body combined. However, we did not find a validated set of this kind of pictograms. The use of faces to describe emotion is also more common and already used by the families and caretakers of the participants in our study. As such, we decided to use PECS (Bondy and Frost 1994) images of facial expressions of emotions.

One limitation of our work is that the emotions were acted. Natural and acted emotions lead to different expressions and underlying processes (De Gelder and Van den Stock 2011; Dael et al. 2012). However, for the postural expressions of emotions, we were inspired by the BEAST validated database (De Gelder and Van den Stock 2011), which also contains expressions of acted emotions. Furthermore, acted emotions are used for emotion recognition in therapy for individuals with ASD, such as for the *Mind Reading* software (Baron-Cohen et al. 2001).

We also used videos to assess emotion recognition. Several studies have suggested that interacting with a physical robot is different from interacting with a virtual or video version of the same robot. However, we chose to use videos for the following reasons: (1) for TD participants, the evaluation was done using an online questionnaire. In a setup with real robots, each participant would have needed to come to our laboratory to perform the test, one at a time. Under these conditions, we would have had fewer participants; (2) it might have been difficult to achieve the same controlled conditions for every participant. We did not want to induce variability into the setup between the participants (i.e., changes of lights, changes of background, unexpected behavior of the robots, distance to the robots, variability in the human's behavior while acting, impact of the different sizes); (3) moreover, a real robot produces a lot of sounds due to its motors and joints. It would have added another sensory input not already present in the virtual agent or human. Individuals with ASD can be very sensitive to sound; indeed, this was true for some of our participants.

We also noticed that the task (i.e., choosing a label for the video) was too complex for some of the participants with ASD. However, by observing their behavior during the session, we are confident that the instructions were understood. Some of the participants had difficulties understanding or linking an emotion to the video they were watching. In addition, in the results, we observed that children with ASD participated better than adults with ASD. This may be caused by the fact that children perform this kind of task more often than adults in their care facility, in the form of games. Prior to the task, we tried to avoid this effect by explaining and showing to the participants how to perform the task before doing it. However, in future studies, it might be more effective to train subjects with the same system but without the emotions involved. For example, we could ask a participant to perform a complete session of recognizing actions (such as walking, running, jumping, etc.), to make sure they are familiar with the task, and then proceed to the emotion recognition task. Since we wanted to assess a participant's capacity to recognize emotions, we did not want to overexpose them to the emotion videos in order to avoid any learning effects.

We also noticed that the experimenter and caretaker behaviors had an impact on the participants. Indeed, when the subjects played or worked with their caretaker, they generally receive feedback to correct or be encouraged. In our sessions, we did not exhibit this behavior in an effort to not induce any learning effects. Some of the participants showed frustration during the task, as no feedback was provided.

Another limitation is that TD participants ans participants with ASD may have learned the emotions while doing the experiment. Indeed, the first presentation of emotions during the FO and BO conditions could have been a preamble for the participants to recognize the emotions for the BF condition. However, a learning effect in a database of this size is difficult to avoid.

Future work includes interactions between the individuals with ASD and various platforms over longer periods of time. This will allow the robot/agent to better adapt and also permit us to investigate various behaviors that would facilitate an understanding of the impact of each parameter on the success rate of the adaptation and interaction processes for individuals with ASD.

Acknowledgments This work is supported by IdF Doctoral Fellowship 2013, HANDICAP theme and is done in collaboration with MAIA Autisme and FAM-La Lendemaine thanks to G. Lerigoleur and D. Mercier. We would like to thank our participants, their families, and their caretakers. We also thank Virginie Demulier for her help.

References

Bar-Haim, Y., & Bart, O. (2006). Motor function and social participation in kindergarten children. *Social Development*, *15*(2), 296–310.

Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is there a" language of the eyes"? evidence from normal adults, and adults with autism or asperger syndrome. *Visual Cognition*, *4*(3), 311–331.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "reading the mind in the eyes" test revised version: A study with normal adults, and adults with asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, *42*(2), 241–251.

Begeer, S., Rieffe, C., Terwogt, M. M., & Stockmann, L. (2006). Attention to facial emotion expressions in children with autism. *Autism*, *10*(1), 37–51.

Bondy, A. S., & Frost, L. A. (1994). The picture exchange communication system. *Focus on Autism and Other Developmental Disabilities*, *9*(3), 1–19.

Bray, A., Subanandan, A., Isableu, B., Ohlmann, T., Golding, J. F., & Gresty, M. A. (2004). We are most aware of our place in the world when about to fall. *Current Biology*, *14*(15), R609–R610.

Brown, C., & Dunn, W. (2002). *Adolescent-adult sensory profile: User's manual*. San Antonio: Therapy Skill Builders.

Buisine, S., Courgeon, M., Charles, A., Clavel, C., Martin, J. C., Tan, N., et al. (2014). The role of body postures in the recognition of emotions in contextually rich scenarios. *International Journal of Human-Computer Interaction*, *30*(1), 52–62.

Burke, D., Hagbarth, K. E., Löfstedt, L., & Wallin, B. G. (1976). The responses of human muscle spindle endings to vibration of non-contracting muscles. *The Journal of physiology*, *261*(3), 673–693.

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

Charman, T., Swettenham, J., Baron-Cohen, S., Cox, A., Baird, G., & Drew, A. (1997). Infants with autism: An investigation of empathy, pretend play, joint attention, and imitation. *Developmental Psychology*, *33*(5), 781.

Chevalier, P., Isableu, B., Martin, J. C., & Tapus, A. (2016). Individuals with autism: Analysis of the first interaction with nao robot based on their proprioceptive and kinematic profiles. In T. Borangiu (Ed.), *Advances in robot design and intelligent control* (pp. 225–233). Berlin: Springer.

Chiari, L., Rocchi, L., & Cappello, A. (2002). Stabilometric parameters are affected by anthropometry and foot placement. *Clinical Biomechanics*, *17*(9), 666–677.

Coates, S., Lord, M., & Jakabovics, E. (1975). Field dependence-independence, social-non-social play and sex differences in preschool children. *Perceptual and Motor Skills*, *40*(1), 195–202.

Costa, S., Soares, F., Pereira, A. P., Santos, C., Hiolle, A. (2014). Building a game scenario to encourage children with autism to recognize and label emotions using a humanoid robot. In *Robot and human interactive communication, 2014 RO-MAN: The 23rd IEEE international symposium on*, IEEE, pp 820–825

Coulson, M. (2004). Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of Nonverbal Behavior*, *28*(2), 117–139.

Courgeon, M., & Clavel, C. (2013). Marc: A framework that features emotion models for facial animation during human-computer interaction. *Journal on Multimodal User Interfaces*, *7*(4), 311–319.

Courgeon, M., Grynszpan, O., Buisine, S., Martin, J. (2012). Interactive expressive virtual characters: Challenges for conducting experimental studies about multimodal social interaction. In *Proceedings of the 9th International Conference on Disability, Virtual Reality & Associated Technologies Laval*, France, 10–12 Sept.

Dael, N., Mortillaro, M., & Scherer, K. R. (2012). Emotion expression in body action and posture. *Emotion*, *12*(5), 1085.

Dautenhahn, K., Nehaniv, C. L., Walters, M. L., Robins, B., Kose-Bagci, H., Mirza, N. A., et al. (2009). Kaspar—A minimally expressive humanoid robot for human-robot interaction research. *Applied Bionics and Biomechanics*, *6*(3–4), 369–397.

De Gelder, B., & Van den Stock, J. (2011). The bodily expressive action stimulus test (beast). construction and validation of a stimulus basis for measuring perception of whole body expression of emotions. *Frontiers in Psychology*, *2*, 419.

Dunn W (1999) The sensory profile: Users manual (psychological corporation, San Antonio, TX)

Edin, B. B. (2001). Cutaneous afferents provide information about knee joint movements in humans. *The Journal of Physiology*, *531*(1), 289–297.

Ekman, P., & Friesen, W. (1984). *Emotion facial action coding system (em-facs)*. San Francisco: University of California.

Erden, M. S. (2013). Emotional postures for the humanoid-robot nao. *International Journal of Social Robotics*, *5*(4), 441–456.

Feil-Seifer, D., Mataric, M. J. (2005) Defining socially assistive robotics. In *Rehabilitation robotics, 2005. ICORR 2005. 9th International conference on, IEEE*, pp 465–468

Ferrell, W., Gandevia, S., & McCloskey, D. (1987). The role of joint receptors in human kinaesthesia when intramuscular receptors cannot contribute. *The Journal of physiology*, *386*(1), 63–71.

Gepner, B., & Mestre, D. R. (2002). Brief report: Postural reactivity to fast visual motion differentiates autistic from children with asperger syndrome. *Journal of Autism and Developmental disorders*, *32*(3), 231–238.

Gepner, B., Mestre, D., Masson, G., & de Schonen, S. (1995). Postural effects of motion vision in young autistic children. *NeuroReport*, *6*(8), 1211–1214.

Golan, O., Ashwin, E., Granader, Y., McClintock, S., Day, K., Leggett, V., et al. (2010). Enhancing emotion recognition in children with autism spectrum conditions: An intervention using animated vehicles with real emotional faces. *Journal of Autism and Developmental Disorders*, *40*(3), 269–279.

Goodwin, G. M., McCloskey, D. I., & Matthews, P. B. (1972). Proprioceptive illusions induced by muscle vibration: Contribution by muscle spindles to perception? *Science*, *175*(4028), 1382–1384.

Gowen, E., & Hamilton, A. (2013). Motor abilities in autism: A review using a computational context. *Journal of Autism and Developmental Disorders*, *43*(2), 323–344.

Greffou, S., Bertone, A., Hahler, E. M., Hanssens, J. M., Mottron, L., & Faubert, J. (2012). Postural hypo-reactivity in autism is contingent on development and visual environment: A fully immersive virtual reality study. *Journal of Autism and Developmental Disorders*, *42*(6), 961–970.

Hart, M. (2005). Autism/excel study. In *Proceedings of the 7th international ACM SIGACCESS conference on computers and accessibility*, ACM, pp 136–141

Haswell, C. C., Izawa, J., Dowell, L. R., Mostofsky, S. H., & Shadmehr, R. (2009). Representation of internal models of action in the autistic brain. *Nature Neuroscience*, *12*(8), 970–972.

Herbert, M. R., Ziegler, D. A., Makris, N., Filipek, P. A., Kemper, T. L., Normandin, J. J., et al. (2004). Localization of white matter volume increase in autism and developmental language disorder. *Annals of Neurology*, *55*(4), 530–540.

Isableu, B., Fourre, B., Vuillerme, N., Giraudet, G., & Amorim, M. A. (2011). Differential integration of visual and kinaesthetic signals to upright stance. *Experimental Brain Research*, *212*(1), 33–46.

Kim, E. S., Berkovits, L. D., Bernier, E. P., Leyzberg, D., Shic, F., Paul, R., et al. (2013). Social robots as embedded reinforcers of social behavior in children with autism. *Journal of Autism and Developmental Disorders*, *43*(5), 1038–1049.

Kohen-Raz, R., Volkman, F. R., & Cohen, D. J. (1992). Postural control in children with autism. *Journal of Autism and Developmental Disorders*, *22*(3), 419–432.

Kozima, H., Nakagawa, C., Yasuda, Y. (2005) Interactive robots for communication-care: A case-study in autism therapy. In *Robot and human interactive communication, 2005. ROMAN 2005. IEEE International Workshop on*, IEEE, pp. 341–346

Liu, W., & Chepyator-Thomson, J. R. (2009). Field dependence-independence and physical activity engagement among middle school students. *Physical Education and Sport Pedagogy*, *14*(2), 125–136.

Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(45), 16,518–16,523.

Mehrabian, A. (1972). *Nonverbal communication*. Transaction Publishers.

Meltzoff, A. N., Brooks, R., Shon, A. P., & Rao, R. P. (2010). Social robots are psychological agents for infants: A test of gaze following. *Neural Networks*, *23*(8), 966–972.

Qin, S., Nagai, Y., Kumagaya, S., Ayaya, S., Asada, M. (2014). Autism simulator employing augmented reality: A prototype. In *Development and Learning and Epigenetic Robotics (ICDL-Epirob), 2014 Joint IEEE international conferences on*, IEEE, pp 155–156

Roll, J., & Vedel, J. (1982). Kinaesthetic role of muscle afferents in man, studied by tendon vibration and microneurography. *Experimental Brain Research*, *47*(2), 177–190.

Salter, T., Michaud, F., & Larouche, H. (2010). How wild is wild? a taxonomy to characterize the wildness of child-robot interaction. *International Journal of Social Robotics*, *2*(4), 405–415.

Saracho, O. (2003). Matching teachers' and students' cognitive styles. *Early Child Development and Care*, *173*(2–3), 161–173.

Scassellati, B., Admoni, H., & Mataric, M. (2012). Robots for use in autism research. *Annual Review of Biomedical Engineering*, *14*, 275–294.

Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice*, *9*(3), 235–248.
Silver, M., & Oakes, P. (2001). Evaluation of a new computer intervention to teach people with autism or asperger syndrome to recognize and predict emotions in others. *Autism*, *5*(3), 299–316.
Simmons, D. R., Robertson, A. E., McKay, L. S., Toal, E., McAleer, P., & Pollick, F. E. (2009). Vision in autism spectrum disorders. *Vision Research*, *49*(22), 2705–2739.
Tapus, A., Mataric, M., Scassellati, B. (2007). Attention to facial emotion expressions in children with autism. In *IEEE robotics and automation magazine (RAM), Special issue on Grand Challenges in Robotics*, Vol. 14, No. 1
Tapus, A., Peca, A., Aly, A., Pop, C., Jisa, L., Pintea, S., et al. (2012). Children with autism social engagement in interaction with nao, an imitative robot-a series of single case experiments. *Interaction Studies*, *13*(3), 315–347.
Thill, S., Pop, C. A., Belpaeme, T., Ziemke, T., & Vanderborght, B. (2012). Robot-assisted therapy for autism spectrum disorders with (partially) autonomous control: Challenges and outlook. *Paladyn, Journal of Behavioral Robotics*, *3*(4), 209–217.
Wainer, J., Dautenhahn, K., Robins, B., Amirabdollahian, F. (2010). Collaborating with kaspar: Using an autonomous humanoid robot to foster cooperative dyadic play among children with autism (pp 631–638)
Wallbott, H. G. (1998). Bodily expression of emotion. *European Journal of Social Psychology*, *28*(6), 879–896.

Pauline Chevalier received her B.S. Degree with Honors in Electronic Engineering, followed by a M.S. Degree in Electronic Engineering, Intelligent System and Robotics in Université Pierre et Marie Curie, Paris, France. She is currently a Ph.D. student in ENSTA-ParisTech, in the Robotics and Computer Vision Lab, and her Ph.D. subject is "Proprioceptive, Kinematic, Socially and Individualized Human-Machine Interaction" for individuals with Autistic Spectrum Disorder since December 2013. Further details about her research can be found at http://perso.ensta-paristech.fr/~chevalier/.

Jean-Claude Martin is full professor of Computer Science at Paris-South University. He is head of the pluridisciplinary group "Cognition Perception Use" at LIMSI-CNRS to which belong 7 permanent researchers (2 full professor in Computer Science, 4 Associate Professors in Psychology, 1 Engineer). He is the Editor-in-Chief of the Springer Journal on Multimodal Interfaces. He is an elected member of the Association for the Advancement of Affective Computing (AAAC). He recently collaborated with MIT Media Lab and our joint paper won a best paper award at the UBICOMP 2013 Conference. He has participated in several European projects (HUMAINE, ATRACO, NICE, ISLE, Magic Lounge). He is the co-head of the Computer Science department at the Institute where he teaches. He is Program Chair of the International Conference on Social Robotics to be held in Paris in 2015.

Brice Isableu is associate professor in Human Movement Science at Paris-South University. He is member of a pluridisciplinary team "Human Movement, Adaptation & Sports Performance" of the CIAMS laboratory (Complexity, Innovation in Sport and Motor Activities) (EA 4532) to which belong 30 permanent researchers (11 full professor, 28 associate professor, 1 Engineer). His main research interest is focused on identifying the main factors and processes responsible of interindividual differences in motor control, sensory integration and bodily expression of emotion in social interaction. His studies addressed the issue of preferential modes of spatial referencing (visual and proprioceptive-based ones) and how they influence the way sensory signals can be integrated in an optimal manner to minimize problematic sources of uncertainties. These questions are addressed using psychophysical methods and biomechanical modelling to assess the effects of developmental constraints (from children to seniors), pathological deficits and athletic achievement in expertise and learning. He is co-head of the Sport, Science Innovation R&D team of UR CIAMS. He has participated in several national academic projects (ANR Comparse, ANR Ingredible) and industrial projects. He is the co-head of the Master 1 Motor Control; Psychological perspectives at the UFR STAPS where he teaches.

Christophe Bazile received is Ph.D. at Laboratoire CIAMS (UFR STAPS Paris Sud). His thesis subject was to both identify the information-movement couplings involved in the regulation of rhythmical hand-eye coordination (ball-bouncing task) and their development in children aged from 5 to 12 years old. He is now project chief in "Institut médico-éducatif Notre Ecole", an institution for individuals with autism.

Adriana Tapus is a full Professor at ENSTA-ParisTech, France. She received her Ph.D. in Computer Science from Swiss Federal Institute of Technology, Lausanne (EPFL) in 2005 and her degree of Engineer in Computer Science and Engineering from "Politehnica" University of Bucharest, Romania, in 2001. She worked as an Associate Researcher at the University of Southern California (USC), where she pioneered the development of socially assistive robotics, also participating to activity n machine learning, human sensing, and human-robot interaction. Her main interest is on long-term learning (i.e. in particular in interaction with humans) and on-line robot behavior adaptation to external environmental factors. Prof. Tapus is General Chair of the International Conference on Social Robotics (ICSR) 2015 and is in the steering committee of major robotics conferences. She received the Romanian Academy Award for her contributions in assistive robotics in 2010. Further details about her research can be found at http://perso.ensta-paristech.fr/~tapus/.