Mixed reality or LEGO game play? Fostering social interaction in children with Autism

Batuhan Sayis¹ · Rafael Ramirez¹ · Narcis Pares¹

Received: 24 November 2020 / Accepted: 10 September 2021 / Published online: 23 September 2021
© The Author(s) 2021

Abstract

This study extends the previous research in which it has been shown that a mixed reality (MR) system fosters social interaction behaviours (SIBs) in children with Autism Spectrum Condition (ASC). When comparing this system to a LEGO-based non-digital intervention, it has been observed that an MR system effectively mediates a face-to-face play session between a child with ASC and a child without ASC providing new specific advantageous properties (e.g. not being a passive tool, not needing to be guided by the therapist). Considering the newly collected multimodal data totaling to 72 children (36 trials of dyads, child with ASC/child without ASC), a first goal of the present study is to apply detailed statistical inference and machine learning techniques to extensively evaluate the overall effect of this MR system, when compared to the LEGO condition. This goal also includes the analysis of psychophysiological data and allows the context-driven triangulation of the multimodal data which is operationalized by (i) video-coding of SIBs, (ii) psychophysiological data, and (iii) system logs of user-system events. A second goal is to show how SIBs, taking place in these experiences, are influenced by the internal states of the users and the system. SIBs were measured by video-coding overt behaviours (Initiation, Response and Externalization) and with self-reports. Internal states were measured using a wearable device designed by the FuBIntLab (Full-Body Interaction Lab) to acquire: Electrocardiogram (ECG) and Electrodermal Activity (EDA) data. Affective sliders and State Trait Anxiety Scale questionnaires were used as self-reports. Repeated-measures design was chosen with two conditions, the MR environment and the traditional therapy LEGO. The results show that the MR system has a positive effect on SIBs when compared to the LEGO condition, with an added advantage of being more flexible.

Keywords Children with Autism · Mixed reality · Embodied interaction · Social initiation · Psychophysiology · Multimodal evaluation

1 Introduction

Current improvement of daily lives of people with Autism Spectrum Condition (ASC) can benefit from well-structured interventions at a behavioral level. Interventions that provide practice on specific skills may help people with ASC overcome some of the challenges that our society imposes on them. For example, interventions that help them acquire

better social skills can make a huge difference in their daily wellbeing, in their social integration, and ultimately in their autonomy as adults.

Social initiation is the primordial social act to improve social skills in general and lead a more autonomous life. Providing tools and intervention methods to foster and practice social interaction behaviors (SIBs) can help people with ASC in their social life. Early intervention has the greatest possibilities of having an impact on their behaviors and on improving their adult lives; hence, addressing children is essential. Typical interventions to foster SIBs mediated by a therapist are affected by subjectivity of the expert and suffer the interference of having a human social agent in the process; i.e., experts are often unsure whether the behavior of the child with ASC is indeed an initiation or merely a response to a previous action of the expert (Paul 2008; Kasari and Patterson 2012; Srinivasan et al. 2016).

✉ Batuhan Sayis
batuhan.sayis@upf.edu
Rafael Ramirez
rafael.ramirez@upf.edu
Narcis Pares
narcis.pares@upf.edu

¹ Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain

Computer-mediated intervention has the advantage of reducing bias, subjectivity, and very importantly, allows the mediation of social experiences without the human interference, as the therapist can observe the session externally (Golan and Baron-Cohen 2006; Ramdoss et al. 2012). In addition, computer-mediated interventions are stable and predictable environments which allow reducing the anxiety that people with ASC may present during real social interactions. FuBIntLab (Full-Body Interaction Lab) has developed in the last five years a large-scale mixed reality (MR), full-body interaction environment, which allows two children to play face-to-face using: exploration of the physical and virtual worlds simultaneously; body gestures and nonverbal communication; joint attention; and collaborative activities. This environment allows a child with ASC to play with a child without ASC and fosters SIBs to allow the child with ASC to understand the mechanisms and benefits of SIBs. It is an ecologically valid context as it resembles the encounters that children with ASC may find in a public park or in the school playground with a technology that is non-invasive and unencumbered. It also allows the child without ASC to see the child with ASC as a valid play partner and enhances the integration of children with ASC in society.

The MR system was designed in a two-year feasibility study that allowed us to prove that it fosters SIBs in children with ASC (Mora-Guiard et al. 2016). We then undertook a three-year phase to compare its potential with a typical intervention used by therapists, based on construction toys (e.g. LEGO bricks) (LeGoff 2004). Through an initial set of 18 trials from the second phase, it has been observed that the MR system generates at least as many SIBs in children with ASC as an intervention using LEGO bricks, while having the advantage of not being a passive tool as LEGO, and not needing to be guided by the therapist who could introduce interference (Crowell et al. 2020). Moreover, the MR system has great flexibility by providing a broad range of different experiences. It is an adaptable tool that can provide an ecologically valid experience for different types of children.

Since this project provides multi-modal data, it has been evaluated by using as many data sources as possible. During the first phase, the effects of the MR system has been evaluated by using: video-coding (coding behavioral data from video recordings) of overt behaviors using an adapted video-coding scheme to acquire fine-grained detail; system log files detailing system triggering of events and decisions, spatial position of children, etc; and questionnaires to the children and to the parents (Mora-Guiard et al. 2016). To try to achieve more objective information, during the second phase psychophysiological measures were added: heart rate variability (through ECG) and electrodermal activity (EDA) (Crowell et al. 2020).

Based on our previous observations, in this study, the following research question is being answered: Is it possible to

develop a full-body interactive ICT system, and in particular an MR system, to affect the frequency of SIBs as much as a traditional therapy setting? To this end, in the current study, the data from 36 children have been added and multi-modal data models have been obtained to evaluate the overall effect of the MR system, when compared to the LEGO condition. Moreover, context-driven approaches operationalized by the video-coding of social interaction behaviors, psychophysiological data and system logs have been used. Additionally, it has been shown that how the SIBs taking place in these experiences may be influenced by internal states and system logs.

2 Background

Social interaction includes person to person contact in the form of a changing sequence of initiations and responses between individuals or groups (Goffman 1955). A considerable amount of literature has been published on high functioning children with Autism Spectrum Condition (ASC) and their capability on continuing social interactions when an initiation is addressed to them. However, they are significantly less successful in generating social initiations by themselves (Sigman et al. 1999).

The emergence of technology and promising results on the affinity that children with ASC have toward the Information and Communication Technologies (ICT) has allowed the development of more engaging and dynamic interventions and learning experiences. There is a large volume of published studies describing the important role of ICT applications in ASC therapy because it triggers less socially threatening situations (Brown and Murray 2001; Bernard-Opitz et al. 2001; Moore and Calvert 2000). Moreover, much of the current literature pays particular attention to the virtual reality (VR) interventions since they create an immersive experience so that children remain focused while using them (Trepagnier 1999).

There are several VR-based technologies that have been explicitly developed and introduced for users with ASC (Strickland et al. 2007; Ke and Im 2013; Parsons et al. 2006). Parsons et al. (2006) developed two types of VEs (Virtual Environments)—a cafe' and a bus to investigate whether participants with ASC relate their use of the VE to experiences in the real world. The VEs presented to participants on a laptop. Strickland et al. (2007) presented a series of VR programs spanning to their 15 years of research that teach real world actions to children with ASC. These programs included immersive street and fire safety VR (first-person interaction) delivered through VR headsets, PC and web based programs. Ke and Im (Ke and Im, 2013), examined the implementation and potential effect of a VR based social interaction program on the

interaction and communication performance of children with ASC. This program was using web based VR and participants practiced communication in the Second Life game. Moreover, there is an additional benefit of using virtual environments in terms of becoming an adaptable tool that can provide the most adequate type of experience for different types of children. The study by Bernardini et al. (2014) described how the actions of a child were continuously interpreted by a system that tailored a virtual character's responses to interact with the child spontaneously. However, there are issues in these studies. On the one hand they do not provide face-to-face interaction with other users without encumbering or invasive technologies. This makes them less ecologically valid since they miss the potential of exchanging non-verbal communication, facial expressions and proximity cues. Moreover, they miss the coupling between the virtual environment and the user's body proprioceptive abilities, kinesthetic sensation, and navigational potential of physical space. Embodied cognition theories emphasize the formative role of proprioceptive and kinesthetic cues when people make sense of the world through their bodies (Wilson 2002). According to the Embodied Cognition, human awareness is based on the active interplay between our body and the environment (Borghini and Cimatti 2010). Along with the unique viewpoint, each individual's cognition is affected and linked with its respective body dynamics and social context (Roussos et al. 1999), defining an embodied state, embodiment. In our view, the lack of embodiment in the interaction design is a missed opportunity in fully utilizing the virtual environments. This view is backed up by Di Paolo, where the critical role of the embodiment for the promotion of better social understanding is emphasized (Di Paolo et al. 2010). Already in 2001–2004 a pioneering project in multi-modal interactive spaces for children with ASC, known as MEDIATE, integrated the concepts of full-body interaction where children were able to involve themselves within the environment through body cues, touch and movement (Pares et al. 2005). In the ASC context, this study was one of the first to discuss ways to encourage and improve exploration in interactive settings in both communication and social scenarios. Other studies have also found that an increase in body movements contributed to an increase in the player's engagement and created a playful social context (Bianchi-Berthouze et al. 2007), making Mixed Reality (MR) full-body interactive systems beneficial for intervention. In Pico's Adventure, for example, children with ASC were called upon to cooperate with others by synchronized gestures and motions, where game play was controlled by Kinect motion tracking sensors (Malinverni et al. 2017). Furthermore, large scale interactive full-body environments also facilitate the ability to effectively navigate a physical and virtual

space simultaneously. They also make face-to-face communication possible between participants, where interaction is not limited to just verbal exchanges as they include also nonverbal body cues such as interactional synchrony, joint attention, and proximity (Mora-Guiard et al. 2016). Although the advantages are apparent, this can also be problematic for children with ASC, in cases where the embodiment is needed. While there is no consensus on the question around "Do people with ASC have distinct embodiment?", difficulties in coordination and/or sensory responsiveness is highly common. Moreover, people with ASC may be exhibiting overt behaviors which do not correspond with their internal state. Besides, self-assessment questionnaires for many people with ASC can be misleading too (Picard 2009); e.g. people with ASC can appear relatively calm despite being in a hyperaroused state which can be detected with physiological cues such as high resting heart rate (Goodwin et al. 2006). Likewise, an individual with ASC can have an electrodermal activity (EDA) between the hyperarousal and the hypoarousal levels without showing obvious behavioral changes (Hirstein et al. 2001). Thus psychophysiology may be useful for studies on social practices for individuals with ASC. The impact of anxiety can be quantified through physical cues such as heart and muscle activity. It has been shown that the EDA and electromyogram (EMG) measurements provide valuable information to assess the shifts in anxious behavior when a training is conducted with a virtual teacher (Prendinger et al. 2005). In this regard, the project Virtual Environment System for Social Interaction (VESSI) (Welch et al. 2010) undertook an evaluation which brought together the ratings from a clinical observer and participants' physiological signals to uncover patterns in anxiety with respect to specific social factors (e.g. eye contacts and proximity). Their observations were aligned with the previous social anxiety studies in the context of adults without ASC in real-world experiences, as well as children with ASC and children without ASC in a virtual environment. In another study, two computer-based cognitive tasks were used to investigate anxiety in children with ASC (Liu et al. 2008). Using Machine Learning (ML) techniques they were able to classify anxiety (with the tagged instances from a behavioral therapist) at an average accuracy of 79.5%, using different physiological modalities together, including ECG and EDA.

Nonetheless, psychophysiological signals are prone to contamination by motion artifacts (Boucsein 2012). Particular consideration should be paid to using both ECG and EDA, since they are susceptible to motion artifacts due to movement of the peripheral body. Furthermore, an association between physical activity and the measured physiological signals can also exist (Picard and Healey 1997). Thus, an ambulatory nature of the experiences (especially

in ecologically valid conditions) and the design of wearable sensing devices should be considered carefully in terms of how physical activity of participants can affect the assessment of physiological measurements. How this is addressed will be described in the following section.

3 Materials and methods

3.1 Wearable multichannel psychophysiology

To acquire physiological signals, ECG and EDA sensors were embedded in a comfortable wearable designed by FuBIntLab (Full-Body Interaction Lab). The hardware multi-modal platform developed by PLUX was used as a main processing unit which possesses sufficient signal fidelity for recordings of physiological signals (PLUX Wireless Biosignals 2020). This plux platform is a ISO 13485 class medical device for biosignals' acquisition. The designed wearable is primarily designed and built to capture accurate physiological data (robust to variations in contact during motion) from children (with and without ASC) in the context of full-body interaction. In addition, the sensors and cables were integrated into the wearable that was kid friendly by giving it a form factor that is comfortable and acceptable for them. It is quite unobtrusive and does not interfere with the user's experience or create additional stress or burden (hands-free). Different choices of sensor and electrode types and their placement on the body were analyzed and tested as they play a significant role in the accuracy of the measured physiological signals. Gelled self-adhesive disposable Ag/AgCl electrodes were preferred, as adhesive pads provide sufficient contact so as to not move around the attached location. In EDA measurements, it has been shown that the most responsive measurement sites are the feet, fingers and shoulders compared to other evaluated body parts (van Dooren and Janssen 2012). Therefore, in order to get a stable signal that is least disturbed by movement artifacts, we verified the best position for EDA electrodes that was on the shoulder. As accurate ECG signal analysis can also be confounded by movement artifacts, proper placement of electrodes on locations far from high muscle activity can help reduce interference artifacts. Thus, to avoid these, two of the three lead ECG electrodes were placed in the upper chest area, and the ground electrode was placed on the neck. Further, details will be disclosed in a future article.

3.2 The MR system (experimental condition)

The MR system is an installation based on a virtual environment, specifically developed as a space to encourage social initiations in children with ASC while playing with children without ASC, where exploration and discovery of hidden

Fig. 1 Configuration of the MR system during an experimental trial

Fig. 2 Manipulation of an object by two players when their place-holders brought together (left) / Children were motivated to work together in building constructions in LEGO condition (right)

virtual objects and surprises is possible. It is an ecologically valid face-to-face full-body interaction environment in which a natural dynamic of collaboration and exploration is fostered.

Users experiment and share their experiences immersed in a floor-based, six-meter diameter, circular visual interface generated by two full HD projectors. The interaction with the VE occurs through a physical object acting as a “placeholder” (Fig. 2). This object is designed in the shape of a firefly net to allow the children to have a better sense of control over the VE and focus their attention. The LED lights placed on these nets are tracked on the playing field with the help of a multi-camera system.

The range of sounds and visuals, the changes in the characters and objects, as well as the surprising elements, all generated by the system, provide a positive reinforcement to players and encourage creative exploration, which would have otherwise been provided by the therapist, having to mediate, guide and support the progress of the children. The MR system provides structure and assistance to therapists to observe and mediate the session externally, while ensuring a stable and predictable context for children to put social interactions into practice.

The game is based on an imaginary world designed through Participatory Design Workshops together with four male children with ASC (10–12 years old). It is a world covered by a layer of virtual fog which opens wherever the

children place their nets. This way they can locate and collect virtual fireflies that transform into companion characters. The environment employs sophisticated techniques, inviting and motivating the children to interact with the game in terms of modifying objects and merging characters in a collaborative way. For a detailed explanation of the interaction design elements in the game, refer to (Mora-Guiard et al. 2016). All the features extracted from the system logs will be hereinafter referred to as game activity measurements (Fig. 1).

3.3 The traditional play therapy: LEGO (control condition)

The traditional play activity was created based on Daniel Legoff's well-known therapy. Legoff obtained promising results in improving the acquisition of social skills in children with ASC especially in peer interaction scenarios (LeGoff 2004). We have adopted this strategy and created a customized LEGO condition, with the advice from the team of experts that we collaborated with in the Hospital Sant Joan de Déu. It matches the game dynamics defined in the MR environment such as the ambulatory nature of the activity, the exertion which children have to go through, standing position with downward looking attitude, scenarios around proximity between children. This is achieved by an especially designed LEGO play environment in the form of a hexagonal-shaped table that allows free movement around it, as opposed to playing by sitting on the floor. The children start the game by looking for pirate-themed figures in buckets that contain the LEGO blocks, placed on the six vertices of the table. Following the discovery of these figures they continue by collaboratively building a ship for the figures in the center of the table. A notable difference with the MR system is that this one is a passive system in which the therapist mediates the activities (Fig. 2).

4 Experimental design and procedure

The user trials were developed based on dyads of children: a child with ASC (high-functioning) and a child without ASC. Seventy-two children (36 child with ASC/child without ASC dyads) participated from the city of Barcelona, with ages between 8–12 years old ($N = 12$ female, $N = 60$ male). With the help of the collaborator hospital, children with ASC who had been diagnosed with high-functioning ASC through the scale of Observation for the Diagnosis of Autism (ADOS) module 3 having a minimum severity diagnosis of 4 (Lord et al. 1989) were recruited. Children without a diagnosis for ASC were recruited through dissemination on social media and in schools. Additionally, both children with ASC and children without ASC had to score a minimum IQ level of 70,

as determined by the Wechsler Intelligence Scale for Children (WISC) (Wechsler 1949). People scoring below 70 are considered to have mental retardation Gordon and Fleisher (2010). It was decided that these exclusion/inclusion criteria mentioned above would allow accomplishing the minimum degree of collaboration needed to go through the game play specified in the experimental conditions, so the child with ASC and child without ASC could play without the assistance of a therapist or parent.

All procedures performed were aligned with the 1964 Helsinki declaration, and had ethical approval granted by the ethics committee of the collaborator hospital for conducting the sessions. Consents to participate were gathered from parents through a signed informed consent form, which detailed the goals and procedures of the project. Assents from the children were also gathered on the day of the trial to make sure they were willing to participate. All data have been kept anonymous and under Universitat Pompeu Fabra's approved data safety protocols.

At the beginning of the session, the psychologist welcomed the children and created the first contact with them. In a relaxed environment, the psychologist explained the session and used the visual support tool called "Jumby is Calm" (Gallo and Annenberg 2009), as used in social skills therapy, to anticipate possible reactions from the children during the session. Having gained the children's attention and trust by the psychologist, a researcher fitted a kid-friendly wearable, designed by FuBIntLab for recording the psychophysiological measurements. Open Signals software, developed by PLUX Wireless Biosignals S.A., was used to verify the correct placement of the electrodes (EDA & ECG) and be ready for acquisition to start. The child with ASC and the child without ASC dyads were exposed to two conditions in a single experimental session: the MR system (experimental condition) and the LEGO play activity (control condition). Each pair of children played each of the conditions once for 15 minutes, with a 5 minutes break between them. The children went through the MR system (experimental condition) and the LEGO activity (control condition) just one time. Before each condition, there were 1 min baseline recordings, in a standing position, looking at black screen. Before and after each condition, questionnaires were administrated. The assignment of the order of the conditions was carried out randomly to counterbalance the order effect of the tasks. Before the children played in the MR system, they were introduced to the placeholder they needed to interact with the VE, and watched an introductory video of how they should use it.

5 Data analysis

5.1 Preprocessing and feature extraction

5.1.1 Video-coding

Aligned with the previous research (Bauminger 2002; LeGoff 2004; McMahon et al. 2013; Owens et al. 2008; Ruble et al. 2008), the SIBs from the overt behaviors were coded using an observational grid with the categories derived from Bauminger's Social Emotional Intervention study (Bauminger 2002). With the help of the collaborator hospital and the project's psychologist, Bauminger's scheme was adapted to fit the real-time specificities of the data set being analyzed. This scheme determines whether each SIBs is: (1) initiation: the child with ASC begins a new social sequence directed toward the other participant, distinguished from a continuation of a previous sequence by a change in activity; (2) response: the child with ASC responds verbally and/or nonverbally to social stimuli directed toward him/her by the other participant; (3) externalization: when the behavior of the child with ASC was not clearly directed to anyone in particular, or was addressed at a game element, it was coded as an "externalization" (e.g. self-talk, shouting, dancing). Externalizations are also an adaptation that have been included in the coding process. They are important and relevant because, despite them not being formal initiations addressed to the child without ASC, they do have the potential to call his/her attention, as if it were addressed to him/her, and spark a response. The intercoder reliability of the video-coding scheme was calculated for each category (occurrence of social initiations, responses and externalization), both through percentage agreement and Cohen's Kappa. Kappa scores were between 0.60 to 0.69 and the percentage level of agreements were between 0.71 to 0.78 with three coders, providing reliability of the data. See the Supplementary Materials for the cross-tables that the Kappa scores calculated from.

5.1.2 ECG signal processing and feature extraction

Heart Rate Variability (HRV) indicates the parasympathetic nervous system (PNS) index, which is of great interest since greater PNS activity of the vagal tone is associated with better social functioning (Laborde et al. 2017). HRV was computed from the raw ECG signal obtained through the electrodes of the wearable; and the signal data was imported into Kubios HRV software (Tarvainen et al. 2014). Kubios HRV software is a device independent software and supports data from heart rate monitors. Kubios

HRV provides detailed HRV analysis for short- and long-term measurements and for different kinds of study protocols. The software computes commonly used time-domain, frequency-domain and nonlinear HRV indices. Using Kubios HRV, first, the samples were analyzed manually for possible artifacts. The processing of the signal features was performed in a time window of 15 minutes following the start of each experimental condition in which there were series of R-R intervals. The R-R interval refers to the time between two R peaks of a traditional ECG signal. All features were extracted in the time-domain, the frequency-domain, and the nonlinear indices that the spectrum of Kubios includes. Each feature was separately normalized by subtracting the corresponding baseline values from experimental condition values. In the statistical inference analyses (see Sect. 6.1 for the details), RMSSD (root mean square of successive difference of R–R Intervals) was used which is one of the main variables reflected in HRV (Shaffer and Ginsberg 2017) also associated with social functioning. In the ML analyses (see Sect. 6.2 for the details), all the extracted features were used.

5.1.3 EDA signal processing and feature extraction

The EDA signal is made up of the superposition of two signals: the tonic level of skin conductance (SCL), representing the baseline signal, and the superimposed phasic increases in conductance. The phasic components reflect a unitary skin response (SCR). In turn, the responses are given by the activity of the eccrine sweat glands in response to external stimuli (Fowles et al. 1981). The decomposition of both components of the EDA signal was done through a software package for MATLAB called Ledalab 3.4.9 (Benedek and Kaernbach 2010). The Automatic EDA Artifacts Identification library EDA Explorer (www.edaexplorer.media.mit.edu) was used to identify possible artifacts in the samples. Following this step, the identified artifacts were corrected by interpolation. The movements of the body did not generate strong artifacts, so the features of the signal could be correctly extracted. Subsequently, the deconvoluted signal was analyzed by the default peak detection algorithm. To detect significant peaks, the local maximum must have a difference greater than 0.01 microsiemens compared to the previous peak or must follow a local minimum (Braithwaite et al. 2013). The phasic features were calculated within a response window (rw) up to 15 minutes during each experimental condition. All the features included in the Ledalab's analysis spectrum were extracted. Each feature was separately normalized by subtracting the corresponding baseline values from experimental condition values. In the statistical inference analyses (see Sect. 6.1 for the details), CDAmpSum (the sum of amplitudes of all reconvolved SCR with onset in response window) was used which is one of the main

variables reflecting the phasic activity. In the ML analyses (see Sect. 6.2 for the details), all the extracted feature were used.

5.1.4 Questionnaires

To gather data on the child's changes in anxiety level after each experimental condition, a standardized questionnaire called STAIC (State-Trait Anxiety Inventory for Children) was used (Spielberger et al. 1973). This part of the questionnaire (STAIC_state) was read aloud to the children from a tablet by the psychologist of the team, although children could also read the questions themselves; they could then mark their responses. It has been seen in the previous studies that interviewing children with ASC with sensory objects (Kirby et al. 2015), and the potential of multi-touch apps (Hourcade et al. 2013), have been seen successful. Children answered inventory items such as “Right now, I feel calm” with one of three responses: Not at all, Somewhat, A lot. All questions were translated into Spanish and Catalan, as the children were belong to mix culture and they were non-English speakers. Moreover, the Affective Slider scales were used, which were judged related well to the questions on the children's affective state (Betella and Verschure 2016). The Affective Slider consists of continuous Likert scales with icon images on either ends to measure the levels of arousal (arousal_level) and valence (valence_level). Together with these sliders, questions such as “How well do you know your partner right now?” (social_status) and “Would you like to get to know them more?” (desire_to_know_more) (Fig. 3) were asked. See also the Supplementary Materials for the details of the questionnaires.

5.2 Hypotheses

This study has analyzed how a full-body interactive ICT system, specifically an MR system, can be used to foster the SIBs in at least a similar degree as can be achieved with a traditional therapy setting (e.g. with LEGO blocks). Within this big picture, the study also analyzed how SIBs taking place in these experiences are related to the internal states of the participants as well as the events generated and sensed

by the system. More in detail, the hypotheses of the work were:

Hypothesis 1: The MR system generates a similar frequency of SIBs (operationalized by the video-coding of their overt behaviors and questionnaires) in children with ASC as LEGO

Hypothesis 2: The level of internal state activity during MR sessions (operationalized as the psychophysiological measurements and questionnaires) will be related to the frequency of SIBs shown by the children with ASC in at least the same level as they are related for the LEGO sessions.

Hypothesis 3: The count of Game Activity during the MR sessions (operationalized as the amount of events triggered by the system) will show a significant relation to the amount of SIBs shown by the children with ASC.

6 Results

6.1 Statistical inference analyses

The analyses took place in two phases. Firstly, paired sample T tests were used to evaluate the efficacy of the MR system in affecting the SIBs as much as a traditional therapy setting (LEGO). Secondly, the link between SIBs taking place in these experiences with internal state activity and game activity was analyzed. In this regard, correlation analysis and multilevel linear regression analysis were used to evaluate their reciprocal influence. The statistical analyses used the SPSS 23 software (Spss 2015).

6.1.1 Paired sample T tests

When the paired sample T tests were conducted with the target variable MR/LEGO, results showed that there were no statistically significant differences in the operationalized measurements between the MR and LEGO conditions (Table 1).

6.1.2 Correlation analysis

To better explore the possible link between SIBs with the internal state activity and game activity, in both setups, the correlations between SIBs and internal state activity; and the correlations between SIBs and game activity were analyzed.

In LEGO, positive correlations emerged between the internal state measurement RMSSD and Initiation, Response and Externalization. Moreover, a negative correlation emerged between the internal state measurement arousal_level and Externalization. Correlation analyses were not applicable between LEGO game activity measurements

Fig. 3 Tablet questionnaire included sliding bars and icons for arousal, valence, knowing the other peer, and motivation for knowing other peer levels

Table 1 Paired Sample T tests with the target variable LEGO/MR

Modality	Item	LEGO	MR	Significance
Video-Coding of Overt Behaviors	Initiations	M = 9.39 (SD = 6.78)	M = 7.97 (SD = 7.32)	0.225
	Responses	M = 7.06 (SD = 5.06)	M = 5.48 (SD = 4.93)	0.084
	Externalizations	M = 5.61 (SD = 5.38)	M = 6.35 (SD = 6.01)	0.475
Psychophysiological Measurements	HRV (RMSSD)	M = -9.23 (SD = 23.82)	M = -10.81 (SD = 21.5)	0.705
	EDA (CDAmpsum)	M = 70.29 (SD = 92.38)	M = 47.59 (SD = 65.97)	0.12
Questionnaires	social_status	M = 0.17 (SD = 0.33)	M = 0.27 (SD = 0.39)	0.32
	desire_to_know_more	M = 0.01 (SD = 0.1)	M = 0.02 (SD = 0.08)	0.609
	arousal_level	M = 0.05 (SD = 0.32)	M = -0.04 (SD = 0.36)	0.256
	valence_level	M = -0.01 (SD = 0.25)	M = -0.06 (SD = 0.22)	0.298
	STAIC_state	M = 39.34 (SD = 1.45)	M = 39.63 (SD = 2.25)	0.5

Fig. 4 Left: Results for the correlation analyses in LEGO. Right: Results for the correlation analyses in MR. **Correlation is significant at the 0.01 level (2-tailed). *Correlation is significant at the 0.05 level (2-tailed)

and other measurements as the game events in LEGO have not been logged (Fig. 4).

In MR, there were no correlations between SIBs and internal state measurements. However, there was a positive correlation between social_status and number of collected fireflies during the MR experience (hunted_fireflies). Moreover, a positive correlation emerged between internal state measurement HRV (RMSSD) and game activity measurement distance between participants (distance_bt看_participants) and a negative correlation between RMSSD and number of manipulated props (manipulated_props). Additionally, a positive correlation emerged between valence_level and distance_bt看_participants. On the other hand,

negative correlations emerged between valence_level and game activity measurements such as number of pointing at action of the virtual character toward other virtual objects (point_at_action) and number of virtual character merges between participants (merge_char) (Fig. 4). As there were no correlations between SIBs and internal state measurements, in order to investigate any indirect relationship, the relationship between hunted_fireflies and other game activity measurements which are associated with the internal state measurements was further investigated. It was found that there was a positive correlation (0.426*) between hunted_fireflies and distance_bt看_participants (which was positively correlated with RMSSD).

6.1.3 Hierarchical multiple regression analyses (HMRA)

In the analyses done for LEGO, SIBs from the video-coding were significantly predicted by some internal state measurements (Table 2, LEGO). In particular, the RMSSD was included in all the significant models with the only exception of the Externalization. Instead in the prediction of Externalization, CDAmpsum and arousal_level were effective.

In the analyses done for MR, Initiation and Externalization were significantly predicted by some internal state measurements, game activity and SIBs from self-report measurements (Table 2, MR). In particular, desire_to_know_more and hunted_fireflies were included in all the significant prediction models.

6.2 Automatic classification

In order to better understand the deeper details of the link between SIBs with internal state activity and game activity, all the aforementioned features were brought together and nonlinear connections between these features were analyzed with ML techniques.

From the SIBs obtained from overt behaviors, as a first step, Initiation was analyzed in detail since it is considered as one of the most important elements of social interaction (Sigman et al. 1999). After all, social initiation is the primordial social act to improve social skills in general and lead a more autonomous life. The levels of social initiation were contextualized, based on the mean values of each interaction in each condition. Each participant's score was compared (lower or higher than) to the average number of initiations taking place in the experimental and control condition, and each participant's score was placed in one of the two classes generated, namely below average initiations or above average initiations.

In this regard, the problem is formulated as a binary classification task for both platforms. The below and above average initiations in the MR system were named as in_MR_below (n: 15) and in_MR_above (n: 14). The below and above average initiations in the LEGO experience were named as in_LEGO_below (n: 15) and in_LEGO_above (n: 14). In order to combine the information from all the uni-modal features, a single fused vector integrating the data from all the modalities was generated. The resulting feature sets for LEGO (64 features in total, 47 HRV, 12 EDA, 5 Questionnaires) and MR (74 features in total, 47 HRV, 12 EDA, 5 Questionnaires, 10 Game Activity) were quite large. Therefore, to reduce dimensionality of the data, to improve the classification performance, different feature selection techniques were used.

Using Weka Data Mining Software (Hall et al. 2009) the results of two different classifiers were compared. Neural Networks (NN), a state-of-the-art nonlinear classifier, and Decision Trees, an interpretable classifier to better understand the modeling process. Neural networks are seen as one of the most effective techniques to model complex nonlinear hypotheses even when the input feature space is large. They are not bounded by assumptions about the feature distributions. However, Neural Networks were recently viewed as 'black boxes' as they could not explain their predictions. On the other hand, decision trees are very intuitive data structures, and visualizing them can give valuable insights on their criteria for classifying new data.

6.2.1 Feature selection

Two types of feature selection strategies were used:

1. All Features (FS1): For LEGO (64 features, 47 HRV, 12 EDA, 5 Questionnaires) and for MR (74 features,

Table 2 Hierarchical Multiple Regression Analyses

Predicted Variable	Models	Predictor Variable	Significance
LEGO			
Initiation	Model 1	RMSSD	0.006
	Model 2	RMSSD, social_status	0.003
Response	Model 1	RMSSD	0.012
Externalization	Model 1	arousal_level, CDAmpsum	0.005
MR			
Initiation	Model 1	CDAmpsum, desire_to_know_more	0.032
	Model 2	CDAmpsum, desire_to_know_more, creature_changed	0.023
	Model 3	CDAmpsum, desire_to_know_more, creature_changed, hunted_insects	0.002
Response	Model 1	desire_to_know_more	0.064
Externalization	Model 1	valence_level, social_status, STAIC_state	0.054
	Model 2	valence_level, social_status, STAIC_state, desire_to_know_more, hunted_insects	0.023

47 HRV, 12 EDA, 5 Questionnaires, 10 Game Activity) were used as a result of feature extraction.

2. WrapperSubsetEval (FS2): A wrapper approach (with a best-first search method) to feature subset selection was used in order to find a minimum set of attributes that are tailored to the particular domain which maximize the performance of the classifiers. The results of the feature selection strategies can be found in Table 3.

6.2.2 Algorithms

The study includes an overall comparison between the models obtained with two feature selection strategies (FS1, FS2) for both platforms. The accuracy, precision, recall and F1-score metric were used to evaluate the models. The hyperparameters of each model are tuned based on two grids of parameters. The grid of parameters used for the NN model refers to a learning rate ranging from 0.0025 to

0.1 with the step size 0.0025 and momentum ranging from 0.1 to 0.5 with the step size 0.1. The grid of parameters used for the Decision Tree model refers to the confidence factor ranging from 0.1 to 0.9 with the step size 0.1 and minimum number of objects ranging from 1 to 4 with the step size 1. The best set of parameters leading to the best performance of the models were obtained across a 10-fold cross validation. Table 4 depicts the overall scores obtained for the feature sets FS1 and FS2, for each participant in a subject-independent manner for LEGO and the MR system, respectively. With automatic classification it was possible to show that social initiation level recognition scores based on internal state activity and game activity, as a combination of feature sets, for both LEGO and MR conditions are above chance level (ZeroR CV accuracy: 51.72%). Model scores were significantly different from the baseline performance. Moreover, the observed classification scores were not statistically different between LEGO and the MR system. Regarding the details of the decision tree generated using FS2 and discussed in Sect. 7.3 see Fig. 5.

Table 3 Feature Selection Strategies

Category	Type	Dataset
All (FS1)		
Time-domain, Frequency-Domain, Nonlinear Features	HRV	LEGO
Continuous Decomposition (CDA), Standard trough-to-peak (TTP) Analysis	EDA	
STAIc_state, arousal_level, valence_level, social_status, desire_to_know_more	Questionnaire	
Time-domain, Frequency-Domain, Nonlinear Features	HRV	MR
Continuous Decomposition (CDA), Standard trough-to-peak (TTP) Analysis	EDA	
STAIc_state, arousal_level, valence_level, social_status, desire_to_know_more	Questionnaire	
manipulate_prop, merge_char, point_at_action, hunted_fireflies, distance_bt看_participants, time_of_first_char, char_greeting, char_texture_change, char_being_idle, area_covered	Game Activity	
Wrapper Subset(FS2)		
pNN50(%), HF(ms2), HF(n.u.), HF(%), VLF(Hz), LF(ms2), Meanlinelength(beats)	HRV	LEGO
TTPnSCR, CDASCR	EDA	MR
DET, LF/HF	HRV	
manipulate_prop, time_of_first_char	Game Activity	
desire_to_know_more	Questionnaire	

Table 4 Classification Performance of Different Algorithms on LEGO and MR data

Method	FS1				FS2			
	Precision (%)	Recall	F-Measure	CV Acc.	Precision	Recall	F-Measure	CV Acc.
LEGO								
Decision Trees	24.1	24.1	24.1	24.1	69.6	69.0	68.5	68.9
Neural Networks	66.0	65.5	65.4	65.52	89.8	89.7	89.6	89.6
MR								
Decision Trees	62.3	62.1	61.5	62.0	84.3	82.8	82.6	82.7
Neural Networks	51.7	51.7	51.7	51.7	89.9	89.7	89.7	89.6

Fig. 5 Decision tree for the classification of social initiation levels in MR

7 Discussion

7.1 Hypothesis 1: The MR system generates a similar frequency of SIBs in children with ASC as LEGO

The video-coding of overt behaviors revealed no significant differences in the generated number of SIBs between the MR setting and the LEGO control condition. This finding was also consistent with the findings from automatic classification when the complex dynamics between all the features which contribute to the generation of different levels of SIBs (initiation) were analyzed. There were no significant differences between the evaluated automatic classification models in terms of performance for LEGO vs. MR environment. These results imply that the MR system is capable of fostering SIBs in children with ASC with similar success as the traditional strategies, such as playing with LEGO bricks. This is important since there is no human intervention in motivating the children in the MR system, whereas traditional strategies focus on the abilities of the therapist to foster SIBs in children. This does not mean that the MR system is an unsupervised system that does without experts. On the contrary, the MR system can, on the one hand, free the experts from having to mediate the session and allows them to focus on observing the evolution of the child within the intervention program. On the other hand, the MR system fosters SIBs in children with ASC in a homogeneous and unified manner for all children without incorporating human subjectivity. The presence of therapists has been found to interfere in traditional sessions since they represent an added human actor in the activity (Paul 2008; Kasari and Patterson 2012; Srinivasan et al. 2016). Because of this, they cannot be absolutely sure whether the social event generated by the

children with ASC is an initiation that comes from their own inner drive and will, or whether it is merely a response to previous actions started by the therapists themselves. In an MR experience, however, the system defines a context that mediates the session between the children without having the human figure get in the way. The experts can define the parameters of the MR system to adapt the session to each child and can intervene in the session only if they deem necessary.

Analyzing the questionnaires, it is seen that no significant difference existed between the two conditions regarding the reported level of knowledge of the play peer (social_status), and the desire to know more about the peer after playing (desire_to_know_more). In this regard, the promotion of collaboration in the MR environment is probably helping children with ASC to see their peers as valuable play partners, in a similar way to how they see peers in the LEGO condition. Given the challenges in social integration and maintaining social relationships in children with ASC (Strain 1983; Gupta et al. 2014), creating scenarios where users see each other as valuable partners could help in developing and maintaining friendship. The MR system offers the possibility of having a broad diversity of scenarios and can therefore provide even more possibilities to create collaboration situations, compared to a static unchanging approach such as the LEGO bricks are.

7.2 Hypothesis 2: The level of internal state activity during MR sessions will be related to the frequency of SIBs shown by the children with ASC in at least the same level as they are related for the LEGO sessions

Based on the Paired Sample T tests with the target variable MR/LEGO, the STAIC_state results were not significantly different between MR and LEGO. While it has been seen above that the MR condition is fostering SIBs, here it is seen that the children's perceived state anxiety, assessed through the STAIC_state questionnaire, was similar to that of the LEGO condition. This is an important finding, since the MR system is a new and unknown play context, while the LEGO condition is a very well known and familiar play context. The LEGO play context in general, tends to make the children either play solo building their own thing, or play together with the pieces from the start if the therapist mediates the experience and motivates children to build something together. On the other hand, the MR system provides different levels of relationship between users and the system. The MR system allows the users to start by doing solo activities and slowly and seamlessly brings the two together to allow them to discover joint play situations. Moreover, the MR system allows the children to step back from collaborative play and go back to solo playing at any

one time, and then back to collaboration again. We believe that this design approach (named encouraged collaboration, as opposed to enforced collaboration which is most often used in ICT for ASC), which gently encourages collaboration, created a comfortable setting for socialization and hence lowered potential anxiety coming from a novel and unknown situation.

On the other hand, the results obtained for the reported arousal/valence levels of children after playing show no significant difference between the two conditions. There is considerable evidence that factors specific to anxiety include negative valence as well as positive arousal, as mapped on the circumplex model of affect (Feldman 1995). In this regard, not seeing any significant difference in reported arousal/valence measurements between the MR and LEGO condition might also be related to the results found from the state anxiety scales. However, it should be noted that the reliability of self-reports from children with ASC might show fluctuations as they may have difficulties in interpreting their own experiences (Berthoz and Hill 2005).

Psychophysiology, as an objective measure, might provide more representative information about the children's internal state. Hence, this study monitored their physiology, and the results from both HRV and EDA did not exhibit statistically significant differences between MR and LEGO conditions. In our case, it is seen that results from physiological analysis seem to support the self-report analysis described above. Firstly, it should be noted that HRV indicates the parasympathetic nervous system (PNS) index and the PNS activity through vagal tone is associated with better social functioning (Laborde et al. 2017). Seeing both platforms induce a similar level of HRV might show that the choice of an MR full-body interactive system has an effect on socialization as much as a traditional setting. Besides, seeing positive correlation between the level of HRV and SIBs in LEGO might also validate the association between HRV and socialization mentioned before. However, the direction of this association were needed to be understood to be aligned with the previous findings. In this regard, Hierarchical Multiple Regression Analyses (HMRA) showed that with the level of HRV (RMSSD), it was possible to predict the frequency of SIBs (initiation and response) in LEGO. It can therefore be assumed that the higher HRV can be associated with higher frequency of SIBs in LEGO. These results were also consistent with the findings from automatic classification where the complex dynamics between all the features were analyzed (not just with HRV and few other features as in HMRA) which contribute to the generation of different levels of SIBs (initiation). Firstly, with automatic classification, it was possible to show that social initiation level recognition scores (based on internal state activity, game activity and questionnaires, as a combination of feature sets) for both LEGO and MR conditions were above

chance level. Moreover, model scores were significantly different from baseline performance. Secondly, the discovered features with the wrapper subset evaluation method (FS2) just included HRV-related features in the classification model for the LEGO condition. However, on the contrary, in the MR environment it was not possible to see a similar relationship (neither in correlation analysis, or in HMRA analysis, or in Automatic Classification) between HRV and SIBs. This might be because of the sensitivity of the HRV measurements in scenarios demanding physical activity. It is important to bear in mind that when Autonomic Nervous System (ANS) activity is assessed during exercise through HRV, the analysis of time and frequency domain measures of HRV may not yield adequate information (Boettger et al. 2010). In this regard, although the two conditions were designed as similar as possible, physical activity might be higher in the MR condition and hence, still getting in the way of understanding how HRV-related activation is associated with the SIBs in the MR environment (because of the allowed range of full-body activity). Perhaps, this is why it is not possible to see a correlation between HRV and SIBs, or any HRV-dominant features in predicting the levels of SIBs in the MR environment.

On the other hand, EDA is often considered one of the most useful indices of sympathetic nervous activity (SNS) (Boucsein 2012) and might provide different insights compared to HRV (Posada-Quintero et al. 2019). This finding was reflected in the results as well. It was seen that an EDA feature (CDAmpsum) was more effective than a HRV feature (RMSSD) in predicting the levels of SIBs (initiation) in the MR environment in HMRA based analyses. Moreover, EDA might be related to engagement in addition to state anxiety (Hernandez et al. 2014). Thus, seeing similar levels of EDA for both platforms can be associated with a similar response of children to externally presented stimuli in both conditions. Moreover, the results show that `desire_to_know_more` is coupled with an EDA measurement in the HMRA and in the Automatic Classification based prediction models of SIBs (initiation) in the MR environment and not in LEGO. This suggests that the MR environment is probably creating an adequate engaging environment for getting to know the peer and hence foster generation of SIBs.

7.3 Hypothesis 3: The count of Game Activity during the MR sessions will show a significant relation to the amount of SIBs shown by the children with ASC

In MR, there were no direct correlations between SIBs and internal state measurements of the children. However, there was a positive correlation between `social_status` and number of collected fireflies during the MR experience (`hunted_fireflies`). It could be argued that there might be an

indirect relationship between internal states and SIBs as it was also found that there was a positive correlation between `hunted_fireflies` and `distance_btw_participants` (which is a feature related to some of the internal state measurements). A possible explanation for this three-way relation can be as follows.

First of all, hunting fireflies was a simple introductory game mechanic that children could easily understand and perform alone, as a first “low floor” (Papert 1980) contact with the virtual environment. However, the MR system interaction design also envisioned this as opportunities to get the two children interested in one another through “search & discovery”. For example, if a child is catching many fireflies and the other cannot, then the child with less fireflies can have sufficient interest to ask the other how he has caught so many. Another example would be when catching fireflies actually leads to the climax with the caught fireflies magically fusing into a character. This can make the child that achieves this exclaim his surprise and attract the interest of the other, or want to share emotions and thoughts. Hence, despite this being designed as a solo activity, the use of this simple introductory mechanic seems to have been successful in making children engage with the peer, as results show an increase in children’s `social_status`.

Secondly, the positive correlation found between the `distance_btw_participants` and `hunted_fireflies` might also signal that a larger distance between participants might lead players to search the playing area more extensively, and hence collect more fireflies and discover different surprising elements. Thirdly, it should be noted that children without ASC might feel more comfortable in approaching others, compared to children with ASC (Gessaroli et al. 2013). Thus, the unnegotiated approach of the child without ASC toward the child with ASC might have created some level of discomfort in the child with ASC. This can be interpreted from the following findings: (i) positive correlation between `distance_btw_participants` and `RMSSD`; and (ii) positive correlation between `distance_btw_participants` and `valence_level`. Such a discomfort is not found in the LEGO environment because this setting provides a context in which the players were already within proximity of each other, sharing a focal point of joint attention on the table (despite them having to move around it). However, once the barrier in close proximity is achieved in the MR environment (when the approach is negotiated), there might be positive correlations between `RMSSD` and SIBs in MR as well (as long as the physical activity does not get in the way of the measures), as in LEGO. It is therefore likely that such connections exist between SIBs and internal state measurements of the children in MR.

Additionally, character merging (`merge_char`) might also show the presence of collaborative behaviors. A significant increase in the occurrence of these might indicate

the game’s potential to foster SIBs. However, the obtained negative correlations emerging between `valence_level` and the `merge_char` might be a sign of a weak implementation of this design aspect of the MR platform. One evidence of this is that it has been observed in several video recordings that the MR system was activating the character merge action when the participants’ creatures were close enough to each other, although the participants’ intentions seemed to be directed toward other activities. It is possible that the low contingency of these merge events might have led to a negative experience which is represented in the `valence_level`. It should be assured in future designs that no actions are triggered by the system unless it is clearly activated by the users.

Similarly, it was possible to observe negative correlations emerging between `valence_level` and the character “pointing at” gesture (`point_at_action`). Having the character pointing at something was designed to help the users discover the nearby elements which were interactive; hence, it is activated when the character of a user approaches the virtual elements. At that moment, the user’s creature points toward the element and makes an exclamatory remark. However, this action might be disruptive (probably reflected by low `valence_level`) for the child; e.g. imagine the child was trying to move his/her character toward a specific zone of the play area and suddenly the character stops following and does something unexpected. This could be an annoying situation for the child leading to the low valence that has been observed.

On the other hand, it was also possible to see how the collaborative act of manipulating virtual objects (props) through the characters of the children might have had a positive effect on them resulting in better social interaction. Aforementioned possible indirect relationships between SIBs (initiation) and internal state measurements could be observed from the prediction model of the initiation (model 3) derived from HMRA analysis. When the `hunted_fireflies` feature was added to the prediction model 2, which already included an EDA-based feature, it was possible to see an improvement in the significance of the prediction. Moreover, in automatic classification models, the possible indirect relationships between SIBs (initiation) and internal state measurements could also be discovered but in a more complex manner (in terms of predicting the level of initiation). Past research has indicated that different physiological measurements work together (D’Mello and Kory 2012) and may contribute more information when treated together. Besides, when the feature selection techniques were considered (e.g. FS2), the classification scores improved. Although it was not possible to interpret this complex relationship with the NN models, it was possible to see that Decision Trees provided interpretable results with the feature `manipulate_prop` and EDA feature (`TTPnSCR`). `TTPnSCR` (Number of significant skin conductance response (SCRs) within the response

window) is also another EDA feature which represents the phasic activity (<http://www.ledalab.de/documentation.htm>.) and can be related to the external stimuli from the MR system and engagement (Hernandez et al. 2014). Manipulate_prop is the measure of the collaboration act in the MR environment as the virtual elements may only be manipulated when both partners work together. The decision tree structure which has been obtained is as follows: The number of social initiations is above average in the child with ASC only when the TTP.nSCR is higher than a certain threshold (namely, 138) and the feature manipulate_prop is greater than three (i.e., 3 manipulated objects). In any other case, it is possible to observe below average social initiations. In this regard, the mechanism of manipulating props seems to be a significantly adequately designed game mechanics to bring the children together and get them to collaborate and become interested in the other. And that, the more props they manipulate, the better their relationship becomes. With an apparent threshold of 3 as a minimum to really get the initiations up above average. But also that to achieve this, a good previous engagement of the children in the game is needed. This could be related to many factors that need to be reassessed in the future. But for now it is possible to know that if the children are engaged (seen by $TTP.nSCR > 138$) then they have more possibilities of manipulating sufficient props to get involved with the peer and therefore get the initiations above average. However, one might argue that the explanation previously given around the negative correlation found between point_at_action and valence_level might be linked to the findings from the decision tree as having the character pointing at something could also lead to the discovery of potential objects to manipulate, since it is activated when both users approach the virtual elements together. Therefore, it could be argued that possible failure in the manipulation of a prop, when the child with ASC attempts this in a solo action, might have caused some frustration, which might explain also part of the low level of valence. However, the negative valence correlated with “the pointing of character” might be more likely associated with the fact that the character stops following the child and does something unexpected as explained before. This perhaps would tell us that in the future, the character should only point at props when the child (and the character) are in an idle situation rather than the character doing this action all the time. Then the character would be introducing a useful queue. The other feature which was effective in automatic classification models was time_of_first_char. This feature is related to the number of hunted fireflies in the first minutes of the game. After hunting a certain amount of fireflies, it was decided that fireflies would transform into a creature that would become a virtual partner for the user. As it was discussed before, hunting fireflies was a simple introductory game mechanic which children could do alone and easily understand as a first contact

with the system. Hence, despite the time it takes to collect the first necessary amount of fireflies, being designed as a solo activity, the use of this mechanic, when it is considered together with other features, seems to have been successful in making children engage with the peer in different social initiation levels, as results show above baseline prediction scores in the prediction of the level of social initiations.

8 Conclusion and future directions

This study provides new insights into the results from previous studies that indicated the potential of the designed MR system in fostering social interaction behaviors (SIBs) in children with Autism Spectrum Condition (ASC). Moreover, the present study has been one of the first attempts to thoroughly examine how a full-body interactive MR system can be used to foster the SIBs in at least a similar degree as can be achieved with a traditional therapy setting (e.g., with LEGO bricks). Within this big picture, the study also analyzed how SIBs taking place in these experiences are related to the internal states of the participants as well as the events generated and sensed by the system.

The results of this research support the idea that the designed MR system is capable of fostering SIBs in children with ASC with similar success as the LEGO setting, with an added advantage of being more flexible. This finding reported here shed new light on developing a tool that is mediating, guiding, and supporting the progress of the children in terms of practicing SIBs; and providing structure and assistance to therapists. Moreover, a further longitudinal study could assess the long-term effects of this tool, especially when it is used in place of a passive social skills training program that needs to be mediated by a therapist.

The results of this study also indicate that SIBs are related to some of the internal state measurements suggesting similar levels of anxiety levels in both platforms, better social functioning through HRV and engagement through EDA-based measurements. Nevertheless, physiological measures are susceptible to physical activity (especially in our case HRV). Indeed, special care must be taken although the two conditions have been designed as similar as possible. Perhaps, for future work, physiological measurements could be standardized with the quantity of body motion information from the participants, which could be tracked with special algorithms through camera systems. Additionally, as a next step, physiological measurements can be coupled with computer vision based internal state evaluation (e.g face and body gesture recognition) for more accurate evaluation. Implementation of such multi-modal objective evaluation tools can pave the way for carefully designing and assessing the novel tools that can be utilized by the ASC therapists and caregivers.

Finally, the results of this investigation also show that SIBs are related to some of the game activity measurements. The results have provided very useful insight on which game mechanics were useful and which were not. With this information, it is possible to improve the MR experience for the children, both by improving the current game mechanics, as well as designing new ones that can appeal to a broader range of children in the autism spectrum. It was also possible to see that basing the design on the encouraged collaboration approach is helpful in fostering SIBs. One of the important findings is that it is possible to keep up engagement in children by providing sufficient interaction potential, using game elements and mechanics such as hunting fireflies and manipulating props collaboratively. Also that, the children are allowed to step back from collaborative play and go back to solo playing at any one time.

Considerably more work will need to be done in the analysis of these full-body interaction MR systems to better help children with ASC in their social skills, as well as helping children without ASC better integrate children with ASC into society.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10055-021-00580-9>.

Acknowledgements This work has been funded by Spanish Ministry of Economy and Competitiveness under the Maria de Maeztu Units of Excellence Program (MDM-2015-0502).

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This information is available upon request

Data availability This information is available upon request.

Code availability This information is available upon request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Consent to participate This information is available upon request

Consent for publication This information is available upon request

Ethics approval This information is available upon request

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will

need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bauminger N (2002) The facilitation of social-emotional understanding and social interaction in high-functioning children with autism: intervention outcomes. *J Autism Dev Disord* 32(4):283–298
- Benedek M, Kaernbach C (2010) Decomposition of skin conductance data by means of nonnegative deconvolution. *Psychophysiology* 47(4):647–658
- Bernard-Opitz V, Sriram N, Nakhoda-Sapuan S (2001) Enhancing social problem solving in children with autism and normal children through computer-assisted instruction. *J Autism Dev Disord* 31(4):377–384
- Bernardini S, Porayska-Pomsta K, Smith TJ (2014) ECHOES: an intelligent serious game for fostering social communication in children with autism. *Inf Sci* 264:41–60
- Berthoz S, Hill EL (2005) The validity of using self-reports to assess emotion regulation abilities in adults with autism spectrum disorder. *Eur Psychiatry* 20(3):291–298
- Betella A, Verschure PFMJ (2016) The affective slider: a digital self-assessment scale for the measurement of human emotions. *PLOS ONE* 11(2):e0148037. <https://doi.org/10.1371/journal.pone.0148037>
- Bianchi-Berthouze N, Kim WW, Patel D (2007) Does body movement engage you more in digital game play? and why? In: International conference on affective computing and intelligent interaction, Springer, pp 102–113
- Boettger S, Puta C, Yeragani VK, Donath L, Mueller HJ, Gabriel HH, Baer KJ (2010) Heart rate variability, QT variability, and electrodermal activity during exercise. *Med Sci Sports Exerc* 42(3):443–448
- Borghi AM, Cimatti F (2010) Embodied cognition and beyond: acting and sensing the body. *Neuropsychologia* 48(3):763–773. <https://doi.org/10.1016/j.neuropsychologia.2009.10.029>
- Boucsein W (2012) *Electrodermal activity*. Springer, Berlin
- Braithwaite JJ, Watson DG, Jones R, Rowe M (2013) A guide for analysing electrodermal activity (EDA) and skin conductance responses (SCRs) for psychological experiments. *Psychophysiology* 49(1):1017–1034
- Brown J, Murray D (2001) Strategies for enhancing play skills for children with autism spectrum disorder. In: Education and training in mental retardation and developmental disabilities, pp 312–317
- Crowell C, Sayis B, Benitez JP, Pares N (2020) Mixed reality, full-body interactive experience to encourage social initiation for autism: comparison with a control nondigital intervention. *Cyberpsychol Behav Soc Netw* 23(1):5–9
- Di Paolo E, Rohde M, De Jaegher H (2010) Horizons for the enactive mind: Values, social interaction, and play. In: Enaction: towards a new paradigm for cognitive science. The MIT Press
- DMello S, Kory J (2012) Consistent but modest: a meta-analysis on unimodal and multimodal affect detection accuracies from 30 studies. In: Proceedings of the 14th ACM international conference on Multimodal interaction, pp 31–38
- van Dooren M, Janssen JH (2012) Emotional sweating across the body: comparing 16 different skin conductance measurement locations. *Physiol Behav* 106(2):298–304
- Feldman LA (1995) Valence focus and arousal focus: individual differences in the structure of affective experience. *J Pers Soc Psychol* 69(1):153

- Fowles DC, Christie MJ, Edelberg R, Grings WW, Lykken DT, Venables PH (1981) Publication recommendations for electrodermal measurements. *Psychophysiology* 18(3):232–239
- Gallo K, Annenberg J (2009) Eye power. *Science* 5:97
- Gessaroli E, Santelli E, di Pellegrino G, Frassinetti F (2013) Personal space regulation in childhood autism spectrum disorders. *PLoS ONE* 8(9):e74959
- Goffman E (1955) On face-work: an analysis of ritual elements in social interaction. *Psychiatry* 18(3):213–231
- Golan O, Baron-Cohen S (2006) Systemizing empathy: teaching adults with asperger syndrome or high-functioning autism to recognize complex emotions using interactive multimedia. *Dev Psychopathol* 18(2):591–617
- Goodwin MS, Groden J, Velicer WF, Lipsitt LP, Baron MG, Hofmann SG, Groden G (2006) Cardiovascular arousal in individuals with autism. *Focus Autism Other Dev Disabil* 21(2):100–123
- Gordon NJ, Fleisher WL (2010) Effective interviewing and interrogation techniques. Academic Press, London
- Gupta SS, Henninger WR, Vinh ME (2014) First steps to preschool inclusion. Brookes Publishing, Baltimore
- Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH (2009) The WEKA data mining software: an update. *ACM SIGKDD Explor Newsl* 11(1):10–18
- Hernandez J, Riobo I, Rozga A, Abowd GD, Picard RW (2014) Using electrodermal activity to recognize ease of engagement in children during social interactions. In: Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing, pp 307–317
- Hirstein W, Iversen P, Ramachandran VS (2001) Autonomic responses of autistic children to people and objects. *Proc R Soc Lond B* 268(1479):1883–1888
- Hourcade JP, Williams SR, Miller EA, Huebner KE, Liang LJ (2013) Evaluation of tablet apps to encourage social interaction in children with autism spectrum disorders. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 3197–3206
- Kasari C, Patterson S (2012) Interventions addressing social impairment in autism. *Curr Psychiatry Rep* 14(6):713–725
- Ke F, Im T (2013) Virtual-reality-based social interaction training for children with high-functioning Autism. *J Educ Res* 106(6):441–461. <https://doi.org/10.1080/00220671.2013.832999>
- Kirby AV, Dickie VA, Baranek GT (2015) Sensory experiences of children with autism spectrum disorder: in their own words. *Autism* 19(3):316–326
- Laborde S, Mosley E, Thayer JF (2017) Heart rate variability and cardiac vagal tone in psychophysiological research-recommendations for experiment planning, data analysis, and data reporting. *Front Psychol* 8:213
- LeGoff DB (2004) Use of LEGO® as a therapeutic medium for improving social competence. *J Autism Dev Disord* 34(5):557–571
- Liu C, Conn K, Sarkar N, Stone W (2008) Physiology-based affect recognition for computer-assisted intervention of children with Autism Spectrum Disorder. *Int J Hum Comput Stud* 66(9):662–677
- Lord C, Rutter M, Goode S, Heemsbergen J, Jordan H, Mawhood L, Schopler E (1989) Autism diagnostic observation schedule: a standardized observation of communicative and social behavior. *J Autism Dev Disord* 19(2):185–212
- Malinverni L, Mora-Guiard J, Padillo V, Valero L, Hervás A, Pares N (2017) An inclusive design approach for developing video games for children with Autism Spectrum Disorder. *Comput Hum Behav* 71:535–549
- McMahon CM, Vismara LA, Solomon M (2013) Measuring changes in social behavior during a social skills intervention for higher-functioning children and adolescents with autism spectrum disorder. *J Autism Dev Disord* 43(8):1843–1856
- Moore M, Calvert S (2000) Brief report: vocabulary acquisition for children with autism: teacher or computer instruction. *J Autism Dev Disord* 30(4):359–362. <https://doi.org/10.1023/A:1005535602064>
- Mora-Guiard J, Crowell C, Pares N, Heaton P (2016) Lands of fog: Helping children with autism in social interaction through a full-body interactive experience. In: Proceedings of the the 15th international conference on interaction design and children, ACM, pp 262–274
- Owens G, Granader Y, Humphrey A, Baron-Cohen S (2008) LEGO® therapy and the social use of language programme: an evaluation of two social skills interventions for children with high functioning autism and Asperger syndrome. *J Autism Dev Disord* 38(10):1944
- Papert S (1980) Mindstorms: children, computers, and powerful ideas basic books. Inc New York, NY
- Pares N, Masri P, Van Wolferen G, Creed C (2005) Achieving dialogue with children with severe autism in an adaptive multisensory interaction: the MEDATE project. *IEEE Trans Visual Comput Gr* 11(6):734–743
- Parsons S, Leonard A, Mitchell P (2006) Virtual environments for social skills training: comments from two adolescents with autistic spectrum disorder. *Comput Educ* 47(2):186–206
- Paul R (2008) Interventions to improve communication in autism. *Child Adolesc Psychiatr Clin N Am* 17(4):835–856
- Picard RW (2009) Future affective technology for autism and emotion communication. *Philos Trans R Soc London Ser B Biol Sci* 364(1535):3575–84. <https://doi.org/10.1098/rstb.2009.0143>
- Picard RW, Healey J (1997) Affective wearables. *Pers Technol* 1(4):231–240
- PLUX Wireless Biosignals SA (2020) biosignalsplux | Welcome. <https://biosignalsplux.com/>
- Posada-Quintero HF, Dimitrov T, Moutran A, Park S, Chon KH (2019) Analysis of reproducibility of noninvasive measures of sympathetic autonomic control based on electrodermal activity and heart rate variability. *IEEE Access* 7:22523–22531
- Prendinger H, Mori J, Ishizuka M (2005) Using human physiology to evaluate subtle expressivity of a virtual quizmaster in a mathematical game. *Int J Hum Comput Stud* 62(2):231–245
- Ramdom S, MacHalicek W, Rispoli M, Mulloy A, Lang R, O'Reilly M (2012) Computer-based interventions to improve social and emotional skills in individuals with autism spectrum disorders: a systematic review. *Dev Neurorehabil* 15(2):119–135. <https://doi.org/10.3109/17518423.2011.651655>
- Roussos M, Johnson A, Moher T, Leigh J, Vasilakis C, Barnes C (1999) Learning and building together in an immersive virtual world. *Presence Teleoperators Virtual Environ* 8(3):247–263
- Ruble L, Willis H, McLaughlin Crabtree V (2008) Social skills group therapy for autism spectrum disorders. *Clin Case Stud* 7(4):287–300
- Shaffer F, Ginsberg JP (2017) An overview of heart rate variability metrics and norms. *Front Public Health* 5:1–17. <https://doi.org/10.3389/fpubh.2017.00258>
- Sigman M, Ruskin E, Arbelle S, Corona R, Dissanayake C, Espinosa M, Kim N, López A, Zierhut C, Mervis CB (1999) Continuity and change in the social competence of children with autism, Down syndrome, and developmental delays. Monographs of the society for research in child development pp i–139
- Spielberger C, Edwards C, Lushene R, Montuori J, Platzek D (1973) State-trait anxiety inventory for children: preliminary manual. Consulting Psychologists, Palo Alto, CA
- Spss I (2015) IBM SPSS statistics for Windows, version 23.0. IBM Corp, New York
- Srinivasan SM, Eigsti IM, Gifford T, Bhat AN (2016) The effects of embodied rhythm and robotic interventions on the spontaneous and responsive verbal communication skills of children with

- autism spectrum disorder (asd): A further outcome of a pilot randomized controlled trial. *Res Autism Spectrum Disord* 27:73–87
- Strain PS (1983) Generalization of autistic childrens social behavior change: effects of developmentally integrated and segregated settings. *Anal Interven Dev Disabil* 3(1):23–34. [https://doi.org/10.1016/0270-4684\(83\)90024-1](https://doi.org/10.1016/0270-4684(83)90024-1)
- Strickland DC, McAllister D, Coles CD, Osborne S (2007) An evolution of virtual reality training designs for children with Autism and fetal alcohol spectrum disorders. *Topics Language Disord* 27(3):226–241
- Tarvainen MP, Niskanen JP, Lipponen JA, Ranta-Aho PO, Karjalainen PA (2014) Kubios HRV-heart rate variability analysis software. *Comput Methods Programs Biomed* 113(1):210–220
- Trepagnier C (1999) Virtual environments for the investigation and rehabilitation of cognitive and perceptual impairments. *NeuroRehabilitation* 12(1):63–72
- Wechsler D (1949) Wechsler intelligence scale for children. Psychological corporation, New York
- Welch KC, Lahiri U, Warren Z, Sarkar N (2010) An approach to the design of socially acceptable robots for children with autism spectrum disorders. *Int J Soc Robot* 2(4):391–403. <https://doi.org/10.1007/s12369-010-0063-x>
- Wilson M (2002) Six views of embodied cognition. *Psychon Bull Rev* 9(4):625–36. <https://doi.org/10.3758/BF03196322>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.