Evaluating Interactional Synchrony in Full-Body Interaction with Autistic Children

Grazia Ragone*
Department of Informatics, University of Sussex, Falmer BN1 9QJ, UK
g.ragone@sussex.ac.uk

Emeline Brulé
Department of Informatics, University of Sussex, Falmer BN1 9QJ, UK

Kate Howland
Department of Informatics, University of Sussex, Falmer BN1 9QJ, UK

ABSTRACT

Interactional synchrony, the spontaneous coordination of movements during interaction, is increasingly considered important in research on the development of non-verbal communication by autistic children. There is evidence that interventions using embodied-interaction technologies to support interactional synchrony are possible, but we do not have a shared framework in Human-Computer Interaction (HCI) for designing and evaluating such systems. We discuss existing measurement and evaluation tools used in experimental psychology and consider how the prevalent approach could be adapted to naturalistic HCI study contexts, with input from domain experts. We report on an exploratory case study evaluating a full-body interactive musical system with a group of ten autistic children. We provide methodological recommendations for the evaluation of future systems focusing on interactional synchrony, highlight limitations of current measurement tools and suggest mitigations.

CCS CONCEPTS

• **Human-centered computing** → Human computer interaction (HCI); HCI design and evaluation methods.

KEYWORDS

Autism, interactional synchrony, music, Motion Energy Analysis, evaluation methods, motion capture

ACM Reference Format:
Grazia Ragone, Emeline Brulé, and Kate Howland. 2022. Evaluating Interactional Synchrony in Full-Body Interaction with Autistic Children. In *Interaction Design and Children (IDC '22), June 27–30, 2022, Braga, Portugal.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3501712.3529729

1 INTRODUCTION

Existing Human-Computer Interaction (HCI) research on understanding and supporting non-verbal communication for autistic people has examined sequential synchrony (e.g., turn taking in discussion, imitation) and concurrent synchrony (e.g., joint action, mutual gaze) [1]. Recent autism research has suggested that interactional synchrony, i.e., the spontaneous coordination of movements during interaction, is a key form of concurrent synchrony supporting the development of motor, social and cognitive skills [2, 3]. Recent studies [2, 3] have indicated that autistic people display a lower degree of interactional synchrony and suggest that this finding may offer new pathways for interventions to support them, for instance using music therapy [4]. However, the construct of interactional synchrony is not yet guiding the design and evaluation of interactive systems that could be used to support such therapeutic interventions and their evaluation.

A significant obstacle is the difference in typical settings and aims between psychology research, which often studies interactional synchrony in tasks where both participants are seated in a controlled laboratory context; while interventions such as music therapy often use full-body interaction and may take place in a variety of settings, from homes to schools [5–7]. The prevalent tool for measuring interactional synchrony in autism research is Motion Energy Analysis (MEA), which allows for non-intrusive and inexpensive movement tracking using a simple camera. MEA is a video-image analysis software that adopts an automatic procedure to extract time-series movement data from the video recording, using the changes in the greyscale pixels between consecutive video frames [8].

Therefore, we investigate if and how MEA could be adapted for "in the wild" studies of realistic intervention using a wider range of movements. This exploratory study is two-fold: we first interviewed three experts, including the designer of MEA, on the challenges they would expect in using the tool to measure interactional synchrony in free-movement settings and the potential mitigations. We then used MEA in a case-study evaluation of a full-body interaction prototype to further identify implementation and analysis challenges. The system, called OSMoSIS, is based on music-therapy interventions supporting autistic people in developing their social and communication skills [9]. It generates musical patterns following movements of participants in real-time, providing means of expression other than traditional instruments which are sometimes a barrier for autistic children's participation [10]. It was designed to increase engagement in music therapy activities based on imitation and was used successfully for this purpose [10]. We do not know however if increased engagement translates in higher interactional synchrony, nor whether it is possible to use MEA to measure it in this context. Our contribution is primarily methodological, but through the application of the adapted method in the case study, we highlight opportunities for design and HCI considering interactional synchrony useful for the IDC community.

*Corresponding author.

IDC '22, June 27–30, 2022, Braga, Portugal

ACM ISBN 978-1-4503-9197-9/22/06…$15.00
https://doi.org/10.1145/3501712.3529729

2 BACKGROUND AND RELATED WORK

In this section, we give an overview of the concept of interactional synchrony and its implications for the design of interactive systems to support research and therapy related to non-verbal communication, with a focus on music therapy for autistic children.

2.1 Interactional Synchrony in Autism Research

Synchrony, where similar behaviours and movements are observed in persons interacting, is a ubiquitous mechanism in social behaviour and a form of non-verbal communication [11]. Within synchronising behaviours, interactional synchrony refers to the multimodal, embodied process, including movements, vocalisations and visual attention [12–14], which is believed to underlie and enable the development of mutual understanding, collaboration, empathy, smoother conversation and social cognition [15–17]. In the research literature, interactional synchrony is referred to using a wide range of terms, including interpersonal coordination, interpersonal adaptation, attunement, mimicry, stress contagion, emotional contagion, coupling, contingency, reciprocity, mutuality, and dyadic regulation [18]. Its key feature is the concurrent alignment of comparable processes, such as voluntary movements, when individuals interact with each other. Parent–child interactions and the interactional synchrony between them play a critical role in early social and language development [19, 20], with evidence that children and caregivers jointly coordinate movements and vocalisations in time [21, 22].

A recent study by Zampella [3] showed that autistic children displayed less interactional synchrony with familiar and unfamiliar partners, in comparison with non-autistic peers. Beyond group-level differences, interactional synchrony negatively correlated with autism symptom severity [3]. Other studies [23, 24] confirmed that autistic children are less likely than allistic peers to coordinate movements in time with known partners such as caregivers [25] and unknown partners such as researchers [23, 24]. This latter study [24] investigated interactional synchrony in structured (i.e., non-naturalistic) tasks, in which partners simultaneously rock in chairs or swing pendulums. Results suggested the association of reduced interactional synchrony to weaker social skills in autistic participants, which could explain subsequent difficulties such as joint attention [23, 24]. Research has also demonstrated that autistic adolescents display less interactional synchrony, using a periodic rhythmic task with coupled oscillators connected to the interacting agents [23]. The authors suggest weaker synchrony resulted in lower sensitivity and diminished attention to the movements of the other person, and that their findings concerning specific interactional synchronisation features of autistic people could be used as an objective, bio-behavioural marker [23].

Another study [26] investigated interpersonal synchronisation in a naturalistic situation, measuring synchrony between the participants and an avatar, and confirming a subtle impairment in autism compared to their non-autistic peers and children with developmental coordination disorder. The same study associated flawed interactional synchrony with differences in motor control and proprioception in the autistic group compared to the other groups [26]. On the same line, other studies [27] suggest differences in interactional synchrony are associated with differences in motor skills. In addition, these studies and others [27, 28] provide evidence for atypical interactional synchrony in autistic people and suggest it may lead to compromised social interaction by reducing opportunities for social-emotional reciprocity [3].

2.2 Tools and Technologies to Assess Interactional Synchrony with Autistic Participants

Measurements of interactional synchrony may be a valuable indicator of someone's attitude towards the interaction or their interaction partner. Previous research [22] has especially focused on assessing synchrony in the interaction between mother and infant, including manual annotation (e.g., coding events in on video recordings [29]) and computational analysis (e.g., time-series analysis from videos or sensors data [13]). In this paper we focus only on computational approaches because manual annotation, whilst accurate, is not scalable. Furthermore, it cannot be used in real-time because it cannot be built into systems to support live feedback.

Computational tools include tools for generating time-series of multimodal movement data, such as video, wearable devices or motion tracking sensors [30, 31]; and analysis tools measuring synchrony in and between groups of interacting agents [32]. But previous studies with autistic participants [28, 33] that evaluated interactional synchrony focused on stationary experiments in controlled environments. There appears to be a lack of work in experimental contexts where synchrony is measured that allows for more degrees of freedom of movement.

2.2.1 Generating data: motion tracking and depth sensors, video-tracking. Methods for generating time-series movement data include extracting parameters from sensors (motion tracking systems using depth-cameras or wearable devices [34, 35]), and employing video-tracking techniques [36–38]. Thus far, most studies of interactional synchrony measure movement of isolated body parts, using dedicated motion tracking devices measuring depth and 2D coordinates placed in the laboratory, measuring e.g., finger motion [39], eye movement [40], hand motion [41] or leg motion [42]. This includes the use of Microsoft Kinect or more advanced motion tracking systems using depth sensors [41]. Wearable sensors often are body-worn accelerometers [42]. While these methods can be more accurate than video-based techniques and can be used in combination with them [43], they have several limitations. Motion-tracking tools using depth sensors currently are used to generate data about movements and synchrony at specific body coordinates, but currently are not integrated with analysis tools to measure interactional synchrony [43].

Meanwhile, worn sensors are often inappropriate for dynamic naturalistic settings, especially with children who may not respond well to their presence. For example, one author [44] calculated the maximum value of the windowed cross-correlation between the participants' head movements, defined using tracking devices connected to the head. It would not be appropriate to consider placing such a system on autistic participants' heads as it would also undermine the child's spontaneity during the study, For the study of the interaction with autistic people, methods like video-based tracking techniques are preferred, having simplicity as their main advantage. Video footage taken with a single camera is sufficient for obtaining the time series of bodily movement [13, 45], with an accuracy similar to depth cameras [46]. There are two main approaches to generating data based on videos for measuring interactional synchrony: pose estimation, which similarly to motion tracking tracks a point representing a body coordinates, and frame differencing i.e., the capture human motion using pixel changes across a plane or in a region [47].

OpenPose [46], which combines computer vision and deep learning and automatically detects the joints' coordinates in the human body, is a pose estimation tool. It calculates the movement score after summing the distances between 17 body coordinates (i.e., eyes, nose, neck, shoulders, elbows, wrists, mid-hip, left hip, right hip, knees, and ankles) across frames, enabling to measure whole-body movements and evaluate synchrony. OpenPose software is pre-trained, so it is not necessary to train the algorithm to estimate the coordinates of body parts. Although it lacks a graphical user interface and requires coding skills, its accessibility is significantly high because it is available for Mac, Windows, and Linux users.

By contrast, MEA, the tool used in this study, uses frame differencing i.e., it captures human motion using pixel changes across a plane or in a region. Both MEA and OpenPose allow for the movement of multiple persons to be tracked simultaneously and OpenPose has the advantage of reduced sensitivity to background noise and lighting conditions compared to MEA. However, previous studies show they have a very similar accuracy [48], with MEA being easier to use due to its graphical interface and its straightforward integration with analysis tools.

2.2.2 Time series analyses and tools. Time-series analysis tools are based mainly on statistical approaches. Interpersonal coordination appears when two or more individuals coordinate their behaviour in a time series, and it can be analyzed in either time- or frequency domains [49]. The amount of movement is mapped on the y-axis while the x-axis corresponds to the timeline in the time domain. In the frequency domain, instead, the movement is plotted on the y-axis while the x-axis displays the frequency components. Research on synchrony has focused on the similarity of rhythm and timing, which can be interpreted as a frequency-domain event [47]. Hence, a simple correlation applied to assess time-domain coordination cannot be employed to evaluate synchrony in the frequency domain. Automated tools enable objective measurements of the synchrony between interacting agents throughout the time series.

In past research, once time-series were obtained, particularly in structured controlled settings, methods like the Fourier transform were used. But they have severe practical limitations. For example, Fourier transforms require assuming a stable frequency or repetitive pattern during the entire interaction [50]. Other methods used in less structured environments include the cross-wavelet transform [50], the simultaneous analysis of two signals in the frequency and time domain, which require determining a dominant rhythm before the analysis. However, in unstructured settings such as conversational interaction, it was not possible to predict the rhythm of the conversation. Therefore, using the cross-wavelet transform, authors employed the pseudo-synchrony experimental paradigm [51, 52], in which randomly matched time series data are used as a baseline against which to compare synchronous ones.

Therefore, with this paper, we tested MEA's usability in dynamic full-body interactions between a facilitator and an autistic child where we investigate MEA's advantages and limitations in semi-experimental settings. MEA has been used in the last decade to measure interactional synchrony in research related to autism [53, 54]. It is built on two distinct modules: 1) MEA which allows the computation of movement from video pixel changes and 2) the rMEA R package [8], which makes available the computation of synchrony from that pixel quantification data using windowed cross-correlation. The MEA process consists in recording two video streams resulting from the interaction between two persons. In both videos, the movement by the persons detected is measured as a function of time by an image pixels difference computation (Figure 1A). Then, the time-lagged cross-correlation between the two movement functions is computed to establish if the agents move synchronously [45, 55] (Figure 1B). Comparing these two computations shows whether the measured synchrony level is higher than expected by chance.

Since people interacting might not move precisely simultaneously, it may result in a brief lag in reaction to each other; this computation is done many times, comparing each window of each of the two agents that are within a lag of -5+5 secs. Also, authors from previous studies [56] introduced 'randomly shuffled pseudo-synchrony'(Figure 1D) where the data of one of the two persons is shuffled around in blocks encompassing more considerable periods. Calculating the cross-correlations on this shuffled pseudo-interaction still relates the movements of the same two persons, but now at two different moments, therefore no longer representing a real interaction. Interactional synchrony is thus illustrated as a heatmap (Figure 1C), where in the legend on the right side of the map, the colour code represents the strength of correlation in the given time-lag matches, starting from the white meaning no synchrony, and the red at the top meaning high synchrony.

2.3 Technologies in Autism Interventions and Research

Previous HCI research has sought to support playful experiences, within [57] or outside of interventions. Within interventions, technology has been used for increasing effectiveness, as well as diagnosis and evaluation of intervention effectiveness. Technologies to support behavioural interventions often engage the whole body, through e.g., the design of smart-home environments [58–60], interactive multimodal systems [61–63], virtual environments [10, 64, 65,] and robotics [66].

2.3.1 Designing full-body interaction. Technologies have the potential to support and evaluate therapeutic interventions. Full-body interaction systems may be especially well suited for use by autistic participants because they allow engagement, free-movement, exploration opportunities, and social activities, by promoting social interaction and collaboration in multi-user situations. This section presents an overview of the full-body interaction systems developed in recent years in HCI for therapeutic and research purposes. Theoretical foundations of such systems derive from theories that

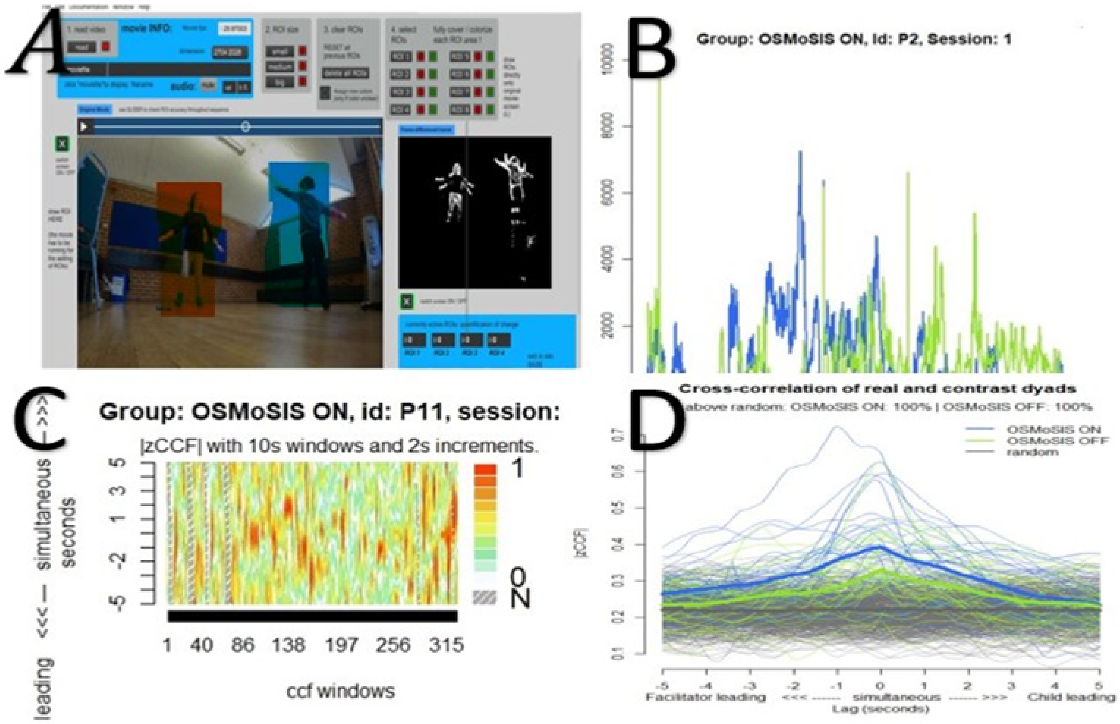

Figure 1: MEA workflow (A: MEA output display with various information and settings areas displayed, included the drawn ROI's which is Red for the facilitator and Blue for the child, while interacting; B: plot function of raw data in rMEA to represent different settings of the synchronisation algorithm; C: Heatmap of synchronisation in the dyad #Participant11, #session 1. The horizontal axis represents time, one interval for each correlation window. The vertical axis represents the various lags, from −5 (facilitator leading) to +5 (child leading), the colour code represents the strength of correlation in the given time-lag matches; D: The lag-plot of normalized correlations between lags. Social presence (area of superiority of real vs random correlations) can be observed to be higher in Sounds On than in Sounds Off dyads. The highest correlation in Sounds On dyads is expressed at about 0 s lag with the facilitator leading the interaction. Similarly, we observe in Sounds Off dyads, high correlation favours facilitator leading phases.

cognition and communication skills are developed through embodied interaction [67]. HCI researchers have developed virtual agents [65, 68] or robots [66] that emulate directing attention, conveying emotion and communicative behaviours.

One of the most used technologies in full-body interaction is Kinect (for example [69]), which enables to accurately track and process body gestures for a low cost**.** Full-body interaction systems let multiple users interact in the same scenario, permitting face-to-face communication and movement in exploratory virtual environments. Examples include Mediate [70], a Multisensory interactive environment for an interface between autistic and allistic users to allow creative exploration through tactile, visual, and auditory stimuli. Similar research projects are Lightpools or Ball del Fanalet [71]. Another project is Lands of Fog [72], a collaborative virtual, full-body interactive environment fostering social interaction in high functioning autistic children. GenPlay [73] as the successor of Land of Fog, was initially developed for two Kinect cameras and an interactive projection on the floor. The goal was to create a system that would be portable to classrooms and hospitals, which would inspire the children's creativity and collaboration. Another project is the Magic Room [74] which encourages interaction with "smart" objects and the whole space through body movement and object manipulation, offering various combinations of stimuli. Full-body interaction can offer multisensory stimuli to keep autistic children focused through motor-focused therapeutic intervention [75, 76].

2.3.2 Technologies for music therapy. However, while interactional synchrony might be an underlying mechanism supporting the development of the skills targeted by these systems, and contribute to explain their success, it is not currently designed for, nor is it a construct used for their evaluation. Furthermore, little research focuses on enhancing the interaction with a human facilitator in motor-focused therapy. One such intervention is music therapy, which has been shown to be helpful for autistic children to develop a higher degree of interactional synchrony [77], as well as their communicative, motor and social skills [9, 73]. Music-therapy can adapt to children's rhythms and emotional needs in a flexible way [78, 79]. Hence, music-therapy can provide a pleasant and non-frightening means to stimulate social interactive skills [78]. Some autistic children face barriers to actively participate in music-therapy activities, due to the complexity and fine-motor skills requirements of traditional musical instruments [80, 81]. Technology can help overcome motor constraints and enable anyone in creating complex musical sounds [78]. This includes virtual environments, interactive multimodal systems, and robots [80].

Interactive multimodal systems adopted in music therapeutic include a variety of input devices (e.g., fabric-based) [82] or gestures [10] to play and generate musical sounds. Moreover, music therapy often focuses on collaboration, either with others or with the therapist, which makes it an especially good area to research interactional synchrony. But music therapists rarely use quantitative data, especially multimodal data such as information about movements, in their work. This is in part due to lacking the tools to do so [80, 83], and further motivates the use of MEA over other approaches to evaluate interactional synchrony: it could allow close collaboration with practitioners. The case-study chosen for our exploratory study of therefore uses a full-body musical interactive system to support a collaborative activity in music therapy, in a series of realistic settings. This system appears to increase children's engagement [84, 85], which in turn could impact interactional synchrony.

3 EVALUATING INTERACTIONAL SYNCHRONY IN FULL-BODY INTERACTION

Given the interest in full-body interaction for both supporting and assessing the development of motor and social skills in autistic children, we focus on a realistic task conducted in a non-laboratory setting. To explore how MEA can be used in this type of context and HCI researchers can contribute to this knowledge area, we first gathered the opinion of three experts in autism and MEA about the challenges and benefits they would expect when using MEA in this context. We then explored challenges with implementation and analysis of results through a case-study with a full-body technology prototype for music therapy with autistic children. We chose an exploratory case-study approach [86] to identify the empirical feasibility of using MEA in this context, both for collecting, analysing, and presenting data. It allows to explore in detail how the tool can be implemented, and key difficulties of which researchers should be aware.

3.1 Expert survey

We collected the opinion of three experts, two autism researchers who used MEA for their own studies, as well as the designer and developer of the MEA tool. This allowed us to identify known limitations for the use of the tool in a different context and to guide the design of the case study presented in the following section.

3.1.1 Data collection and participants. The experts were sent the following questions: 1) In which experimental setups do you think MEA/ rMEA tools could be used? 2) What limitations/issues with the data would you expect from using these tools in naturalistic study contexts with unconstrained participant movement?

They were given the choice of answering them by writing or orally. When answering orally, they were sent the transcript for reviewing. The participants were Devyn Glass (DG), a researcher in the field of autism who has been using MEA and rMEA to examine social motor synchrony during table-top games played on iPads [87]; Dr Alexandra Georgescu (AG) who has used MEA to investigate interpersonal coordination differences between clinical groups [28], and for digitally assisted autism diagnostics based on intra- and interpersonal coordination [28, 53]; and Dr Fabian Ramseyer (FR), a psychotherapist and researcher whose focus is the temporal dynamics of human synchronisation. He is also the programmer and developer of the MEA software.

3.1.2 Data analysis. Given that the responses were authored and reviewed by three domain experts, a summary approach to analysis was most appropriate. We used a basic open-coding approach to identify the key themes in the responses, which were 1) potential benefits of using MEA in the context of full-body interaction in realistic settings; 2) expected differences between the design of experiments using MEA in autism research traditionally and studies focusing on full-body movements; and 3) expected limitations or difficulties the case-study should focus on. We summarised the key points made by experts according to these three headings, and then asked them to review our summary for inaccuracies. In the results, respondents are identified by their initials.

3.1.3 Findings. MEA was designed to capture and analyse stationary interactions, in fact one of the first steps in the process is drawing the Region of Interest (ROIs, i.e., specific areas, such as a single participant or a participant body part) with different colours to distinguish the agents. In the exploratory study for example the child's ROI was also of the same blue colour, and the facilitator's ROI was always red (Figure 1A). The ROI was designed to be static, imagining a typical stationary setting with, for example, the client and the therapist sitting on a chair. The limitations of these studies, however, are that they do not necessarily capture well the specificities of the movements of autistic people.

Potential benefits of using this tool in the context of full body interaction. DG notes: "autistic children may have more movement variability than non-autistic children". This difference may be captured by allowing freedom of movement through more expansive spaces where autistic people can move spontaneously: "free-flow movement will allow us to detect synchrony more accurately for the autistic group as it provides them the opportunity to move dynamically and naturally". DG stated MEA is useful in controlled circumstances. DG affirmed "MEA can capture spontaneous synchrony in naturalistic settings; however, it might be best for static activities". According to DG measuring interactional synchrony in dynamic contexts: "would also allow more accurate detection of multimodal synchrony. This could allow us to examine whether autistic and non-autistic people synchronise in similar or different ways". DG described an instance where one partner's arm movements appeared: "closely coupled with their partner's hand movements". Employing MEA in open full-body interactions or classroom studies: "may be difficult due to participants crossing over into each other's ROI, typical of such dynamic settings".

Hence, for the study of interactional synchrony and autism to progress, researchers could investigate full-body interaction with systems that enables functions of moving ROI where agents move freely around the environment. This would allow the distinction between the agents and, with an advanced system, even the opportunity to make distinction of the different body parts which is something that the last motion tracking devices (e.g., Kinect) would allow.

Expected differences between the design of experiments using MEA in autism research traditionally and studies focusing on full body movements. AG recognizes the great value of the rMEA package: "to investigate interpersonal coordination with continuous data (pixel change time series) in a wide range of setups for the experimental study of nonverbal behaviour". In particular, a main advantage in autism research is that it enables the "quantification of movement without the need of wearable technology". This is helpful, given the sensory sensitivity in autism". However, AG recognizes limitations in using rMEA in a dynamical setting. For example, she notes: "one of the most potential issues stems from the fact that the ROI definition with MEA is done manually and the ROIs are not dynamic. They need to be carefully drawn to include only the area of interest but the entire range of movement for the entire length of the analyzed clip". This is not always easy, will result in ROIs of varying sizes and there can be many instances of region crossing (e.g., one person's leg enters the ROI of the other). She continues: "Differences in the size of the drawing of regions of interest, however, can be addressed through standardisation or rescaling of the data in rMEA".

We support the idea from Glass and Yuill [88] that tailored studies, where children are engaged in motivating activities (e.g., with music or related to preferred interests), autistic participants can be observed in a more natural, semi-experimental setting. Such setups could potentially counter previous findings that highlighted poor interactional synchrony in autistic children. Hence, traditional stationary studies on interactional synchrony may be limited in the exploration of such a complex phenomena.

Expected limitations or difficulties the case-study should focus on. FR confirms the recognition of a range of potential difficulties in different domains using MEA. He adds: "the inability of the system to differentiate between interaction partners, if they move in front of each other". This means that in processing the raw data, certain limitations of MEA must be considered and kept in control. Also, FR adds that there is a "qualitative limitation such that MEA usually will not account for the type of body-part involved in movement dynamics". In the setting of the full-body interaction, there will be no differentiation into body parts or moving parts. "This is a limit, but on the other hand, MEA will capture the dynamics of movement, and the quantification of synchrony in these dynamic aspects may also be viewed as a potential plus". According to FR, MEA allows capturing the "co-regulation and coordination at an abstract level which is highly relevant in itself, looking at dynamics and not merely movement quality, capturing a much more deeply based type of attunement between child and interaction partner". In addition, capturing the coordination at an abstract level of movement is difficult to detect for an observer, simply relying on the visual input provided in real-time. Thus, the time-series generated with MEA will open up an avenue into a quality (movement dynamics) that is difficult to "see" without the aid of information technologies. FR argues: "Using MEA in dynamic settings, there are certain limitations about the type of data you will get with MEA, it will only be possible to extract information on movement dynamics, not on the direction, body-part, expression, etc., of how these movements are performed". According to FR, MEA shows limitations of movement's representation from a technical perspective, deriving motion in the 2-dimensional plane. It results in a "potential bias in the direction concerning the camera". It will underestimate movements, fundamental in full-body interaction, "in the sagittal plane (towards or away from the camera) compared to activities in the other planes". Furthermore, low amounts of bias could be introduced if participants wore highly different types of clothing. Given that MEA operates on frame-differences, higher-contrast clothing will over-estimate movement while low-contrast will under-estimate movement. Such a potential bias may be taken care of by analyzing and controlling for subjects' grey-scale values and background. Ideally, such aspects should be stable across recordings.

In summary, while the three experts were interested in the use of MEA in realistic full-body interaction such as therapeutic contexts, expected difficulties include the static function of ROI, and limitations such as the inability to differentiate between interacting partners when moving in different ROI and between body's parts when the agents are in different areas.

3.2 Exploratory Case-Study

In this section we present and discuss in more detail how we implemented MEA and examined its viability as a method to measure interactional synchrony in a full-body interaction context. We conducted an empirical study of the use of a system called OSMoSIS with 10 participants who were engaged in tasks involving their observation and reproduction of four selected movements with and without the association to selected sounds. The case-study was designed to explore how MEA could be used for evaluating interactional synchrony in a naturalistic, free-movement setting. We present the findings from using MEA in this setting and highlight the challenges and resulting limitations for each step of our study.

3.2.1 System. OSMoSIS (Figure 2) is an interactive musical system that tracks body movements and transforms them into sounds using the Microsoft Kinect v2 [69] motion capture system. The system uses bespoke software that processes and transforms into sounds the data stream generated by the Kinect system as it captures the body movements of one or more individuals and detects interacting movements. It is designed for use in a music therapy setting, with the sounds intended to encourage synchronous movement between an autistic child and a facilitator. The system's design was inspired by previous findings showing that music is a relative strength of autistic people [89], and the movement sonification [90, 91] was intended to promote the multisensory integration to perception and self-motion.

3.2.2 Participants. A total of 10 children (8 boys, 2 girls) aged 5-11 (M age = 8.27; SD =1.42) were recruited through a Special

Table 1: Session design

Warm-up	First Condition	Second Condition	Conclusion
5 minutes	10 minutes with sounds OFF	10 minutes with sounds ON	Farewell

Educational Needs (SEN) school, a mainstream school, and a family-support group. They all had a previous diagnosis of an Autism Spectrum Condition. Participants were voluntarily enrolled in the study, and their legal caregivers gave informed consent to the study (Details of research ethics approval reference hidden for review). The participants did not receive any compensation for their attendance. All participants were invited to take part in four sessions, but seven participants were only able to take part in three, because incoming COVID-19 restrictions preventing them from attending the last session, resulting in 33 sessions. Four sessions were determined as appropriate based on short term music therapy intervention practices and represented a good compromise for data validation and to meet the needs of the participants, their families and teachers.

3.2.3 Study Design and Procedure. The study was designed to investigate whether having the OSMoSIS system sounds turned on led to greater interactional synchrony between the participant and the facilitator than when the sounds were off, with MEA used to measure interactional synchrony. We conducted the study in two different locations, according to how the participants were recruited. The first location was a typical teaching room of a mainstream school, with all the tables and chairs put apart, allowing the child and the facilitator to move freely without obstacles. The second location was a playroom of a parent-support group located in a local library. Most of the participants (9 out of 10) were not familiar with OSMoSIS prior to the study or to similar music therapies. The conditions of Sounds On and Sounds Off were balanced. In each condition four sounds associated to four movements were proposed. Each sound/movement lasted 3 minutes. The design of each session for this study observed the structure of Sounds On always following Sounds Off. After the warm-up where the facilitator welcomed each child with a welcome song including the name of each child, the Sounds Off were designed with the facilitator proposing the same four movements proposed in Sounds On., but without sounds (OSMoSIS deactivated). At the end of each session, the facilitator waved to the child, singing a farewell song, again including the name of the child. All the four sessions had the same structure. The sounds produced in real time by the OSMoSIS system in response to the child and facilitator's movements were expected to motivate children to interact with the facilitator and to result in greater interactional synchrony. Participants were asked to either move freely or imitate the facilitator who proposed four simple activities, inspired by music therapy exercises, like (1) hopping, (2) rolling onto floor, (3) marching and (4) moving arms with OSMoSIS software active (when responding in real-time with sounds to participants and facilitator's movements) or inactive (with no sounds). When the software was active, each sound elicited a specific movement. The study presented two different conditions 1) with sounds (Sounds On) elicited from the software and 2) without sound (Sounds Off). Thus, we analysed a total of 66 sessions (with two sub-sessions for each of the 33 sessions). The procedure for each session is shown in Table 1.

3.2.4 Data collection. **Description**: First, we video recorded all the sessions positioning a Go PRO camera in one of the corners of the room out of children's view. The camera was placed out of view to avoid distracting the children. The workflow of video editing (Figure 2) took a significant amount of time to watch, examine and crop videos when issues like artefacts or crossover were present.

Cross-correlations were transformed (Fisher's Z), and their absolute values were combined, generating one global value of synchrony for each session's interaction. The function of the fundamental values allowed both positive and negative cross-correlations to be added to the synchrony measure, including the anti-phase correlations, which is possible to visualize on a heatmap (Figure 1C). These values resulted as dependent variables in our analysis. The rMEA package provides a function to control coincidental synchrony (Figure 1B), computing pseudo interactions for each real one and then to compare both. To compute the pseudo synchrony, we shuffled the real-time series with pseudo one (Figure 1D). This procedure kept the time structure of the real data permuting only the temporal location of the segments 30s, yielding a distribution of pseudo synchrony values for each real dyad.

Challenges and limitations: We estimate that around 10 percent of data was lost in each session considering the sum of crossover, noise in time series and instances of one of the agents (child or the facilitator) away from the camera. Indeed, the perspective of the camera position can cause self-occlusion (Figure 3) or cause the loss of movement data due to the crossover between therapist and child. Looking at our experience with MEA, we found a few limitations, which stemmed from spending a great amount of time decreasing the issues and assuring data validity.

3.2.5 Data processing. **Description**: The researcher followed a systematic procedure to process the video recorded before using MEA (Figure 2), in order to minimize possible causes of error that can bring in artefacts in the data. The resulted 75 videoclips were imported into MEA software and for each one, we selected a ROI for each participant (including head and body. See Figure 1A). Absolute changes in grayscale values detected from the ROI of each individual were recorded separately as a stream of data, generating two (the facilitator and the child) continuous time series measuring the amount of movement in each region interacting. Subsequently, with an open-source R package (rMEA), we imported, filtered, and analyzed dyadic time-series of nonverbal behaviour generated by the MEA software.

Challenges and limitations: The biggest challenge in data processing was the frame difference in videos where children or the facilitator were performing quick and often overlapping movements. The rMEA could not differentiate time-series in those cases, showing then loads of zeros that needed to be taken out. We had

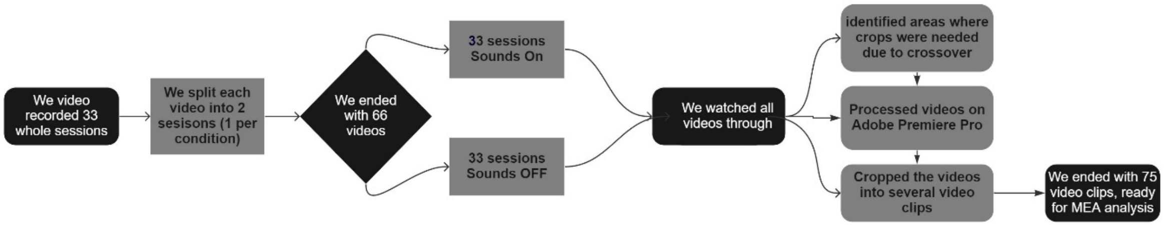

Figure 2: Videos preparation workflow

Figure 3: Examples of crossover and self-occlusion.

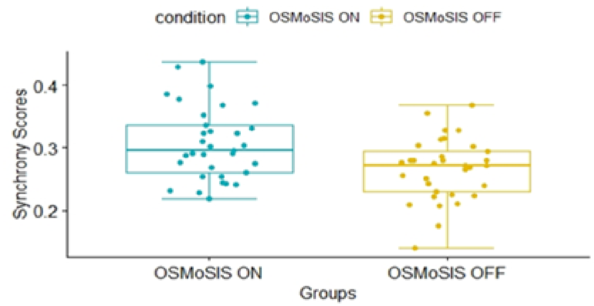

Figure 4: Boxplot of difference between conditions.

to remove those data to the detriment of the richness of data interaction. As raised in the expert interviews regarding the use of MEA in dynamic contexts, occurrences of the 'crossover' (Figure 3), where the two participants intersected in the camera view were a significant issue in the free-movement setting.

We had to split each video into multiple clips to make our analysis as accurate as possible in targeting interactional synchrony and remove unfiltered keyframes, video noise or rapid rhythmic patterns in behaviour quite common in full dynamic body-interaction. One of the issues we encountered was of scale with the interaction of the facilitator and the child, both not in a steady position. If one person appears closer to the camera, slight movements will change the significant pixel value [92]. On the other hand, if a person farther from the camera recreates the same action, MEA will quantify this as less movement because fewer pixels change. In most cases, we overcame these issues by dividing each video session into several video clips (Figure 2) and using different sized Regions of Interest (Figure 1A) and z-transforming (standardising) the data. Some data, as results of the agents' interaction, were lost, including important details of the interaction.

3.2.6 Data Analysis. **Description**: The rMEA package (Figure 1B) provides an interface for assessing interactional synchrony. The synchrony was measured through a windowed cross-lagged correlation of the motion energy time series of the child and the facilitator in every 5-minutes interaction. The resulting correlation measures the similarity of the two-resulting time-series as a function of the displacement (lag), putting each one with the other. Time lags of up to +/- 5s are applied in steps of 10 seconds. Thus, the time series are shifted by 10 seconds and then correlated (Figure 1B). Also, an important function is the graphical representation of the various lags in the vertical axis of the heatmap (Figure 1C) representing from −5 the facilitator leading and to +5, the child leading.

Challenges and limitations: One limitation of using MEA is that it may be challenging to perform motion tracking for multiple individuals at once [36, 93] because the ROIs are likely to overlap, particularly when the target agents are located near each other. It can be addressed using either multiple cameras or methods of simultaneous tracking segmentation of different body parts in a bottom-up fashion [94].

Findings on interactional synchrony: To examine within participants difference in interactional synchrony score between the two conditions of Sounds On and Sounds Off, we first tested the distribution of each condition, finding that the data in each condition were normally distributed. Then we tested the distribution homogeneity and found an equal variance [F (1,64) p = 0.30]. Paired-samples Welch Two sample T-Test were used to comparing means between Sounds On and Off conditions. These tests showed that there was a significant main effect on the interactional synchrony score (t = -3.096, df = 62.354, p = < .05) between the two conditions and Sounds On shows a larger distribution. We ran a repeated-measures analysis of variance (ANOVA) and found significant differences, as within-participant factors by conditions when OSMoSIS On (M=0.307, SD=0.0582) compared to OSMoSIS Off (M=0.266, SD=0.0494) (4).

4 DISCUSSION

In this section we discuss the use of MEA to evaluate interactional synchrony between dyads interacting while using a full-body interaction system. Although the approach shows some limitations and inaccuracies when used for full-body interaction in a naturalistic system, we argue that gains in ecological validity and a potential gain for autism research supports the use of MEA for future research on interactional synchrony, as well as hopefully encouraging the design of new systems to support interactional synchrony.

4.1 Addressing issues with data collection and processing

The crossover between participants, which required the cutting of some videoclips, implied the inability to process and visualize some of the key moments during the full-body interaction. While the chosen method is automatic, the constraints on the physical setup (camera position and movement of the child and facilitator) of the recordings needed to be strict to result in reliable measurements. One could argue that this reduces the ecological validity of the measure. It would be helpful to expand the reasons in HCI for developing tools enabling quantifying synchrony in a larger perspective, with dynamic ROIs for example which would include other qualities of the dyad's interaction. MEA's approach would be a good starting point for a new (maybe multi-modal) method where interactional synchrony can be measured regardless of the stationary or dynamic agents' movement. Video observation still results as the best option in a dynamic context to support a great degree of action of autistic agents. However, we need to consider the other problems that technology would need to predict, such as the occlusion or the sudden change of position in the space (Figure 3). Furthermore, these findings confirmed that interactional synchrony is a complex construct, difficult to detect when it manifests at micro-level. Because of its social significance, accurate measures result as a potentially important pathway for understanding the social difficulties, in autism research and beyond.

4.2 Implications for the design of systems to support interactional synchrony

A major difficulty for the design of technologies to support autistic people is the large variations in the difficulties they face [95]. If, as suggested by recent research [27, 94], difficulties with motor skills contribute to explain difficulties in other domains and could impact their severity, then systems supporting synchronous motor activities with a facilitator could improve the support offered to many in this group. Large scale systems are ideally suited for this purpose, and systems like MEA could be used as a standardised tool to compare the impact of different designs on interactional synchrony and ultimately relate it to gains in the area of motor skills.

Discussing the findings in detail is beyond the scope of this paper; however, being able to identify differences in interactional synchrony between conditions confirms MEA can be used for identifying the impact of systems on the activity. Moreover, our study suggests that autistic children responded well to musical stimuli, confirming previous research findings [9, 61, 80] on music therapy for autistic children. We hope that this will encourage further research into supporting music therapists, as well as research into the design of new instruments using technologies. These have the potential to be especially customisable, thereby addressing the needs or musical sensitivity of individuals who may not be able to use traditional musical instruments [80, 81].

4.3 Future research

The latest advancements in depth-sensing, which consists of a wide range of technologies such as active or passive triangulation, time-of-flight, and ultrasound, offer researchers in the field of no verbal communication a wide range of opportunities. For example, issues like crossover might be solved using more sensors located in different corners in the environment. The vast selection of machines generated a dynamic field of research, with ongoing innovation and motivating the development of many applications across several technical areas. Systems we used in our exploratory study like OSMoSIS can benefit from these technologies and implement in their software the function of measuring interactional synchrony through data extracted from the Kinect SDK. Future research should investigate autistic children's experiences and perspectives about the musical stimuli, from a qualitative and quantitative perspective. Given the potential importance of interactional synchrony for autistic children, this paper suggests that other means for measuring autistic children's interactional synchrony should be explored. Our suggestions for future work will provide us with a deeper understanding of how-to best support, through measurement's tools, the autistic children's verbal and non-verbal interaction, in turn, how this support can be implemented in an environment such as schools or homes. This study represents an essential guideline for HCI since nonverbal communication is central to social interaction [96, 97] in general and in particular for autistic users. Hence, the result of this study supports the idea that computational methods to study interactional synchrony could provide automatic and objective tools to investigate interactive skills in several conditions other than autism.

5 CONCLUSION

We demonstrate that assessment of interactional synchrony is challenging in free-movement study settings. Interactional synchrony is an important aspect to investigate in autism research, which is at the crossroad of several disciplines. In fact, the interdisciplinary expertise resulted from the consultation of the three expert users in this study highlighted new insights useful for researchers and designers in the HCI community to develop new tools or implementing existing ones for assessment and interventions. The HCI and IDC community can benefit from this paper to evaluate other systems and contribute for example on the development and validation of new tools or techniques to automatically assess incidence of synchrony and shape interactive patterns. With autistic participants the use of non-obtrusive sensors is likely to be the best approach in allowing freedom in movements and expression with the emergence of interactional synchrony as a multimodal phenomenon in nonverbal communication.

ACKNOWLEDGMENTS

The technical assistance provided by Dr Johann R. Kleinbub and NSPE Gregorio Baldari during data processing with MEA and rMEA was greatly appreciated for the finalisation of the data analysis.

SELECTION AND PARTICIPATION OF CHILDREN

Ethical approval for this study was obtained from the University of Sussex Sciences & Technology Cross-Schools Research Ethics Committee. A total of 10 children (8 boys, 2 girls) aged 5-11 were recruited through a local Special Educational Needs (SEN) school, a mainstream school, and a family-support group. They all had a previous diagnosis of Autism Spectrum Condition. Participants were voluntarily enrolled in the study, and their legal caregivers gave informed consent to the study. All participants were invited to take part in four sessions, where always another adult was present (e.g., Teaching assistant, parent, or caregiver).

REFERENCES

[1] Feldman R. (2007). Infant biological foundations synchrony and developmental outcomes. Curr. Dir. Psychol. Sci. 16, 340–345. 10.1111/j.1467-8721.2007.00532.x

[2] Fitzpatrick, P., Romero, V., Amaral, J. L., Duncan, A., Barnard, H., Richardson, M. J., & Schmidt, R. C. (2017). Social motor synchronization: Insights for understanding social behavior in autism. Journal of Autism and Developmental Disorders, 47(7), 2092–2107. https://doi.org/10.1007/s10803-017-3124-2

[3] Zampella, C. J., Csumitta, K. D., Simon, E., & Bennetto, L. (2020). Interactional Synchrony and Its Association with Social and Communication Ability in Children With and Without Autism Spectrum Disorder. Journal of Autism and Developmental Disorders, 50(9), 3195–3206. https://doi.org/10.1007/s10803-020-04412-8

[4] Nielsen, J. B., & Holck, U., (2020) Synchronicity in improvisational music therapy – Developing an intersubjective field with a child with autism spectrum disorder, Nordic Journal of Music Therapy, 29:2, 112-131, DOI: 10.1080/08098131.2019.1680571

[5] Boso, M., Emanuele, E., Minazzi, V., Abbamonte, M., & Politi, P. (2007). Effect of long-term interactive music therapy on behavior profile and musical skills in young adults with severe autism. The journal of alternative and complementary medicine, 13(7), 709-712.

[6] Beer, J. M., Boren, M., & Liles, K. R. (2016, March). Robot assisted music therapy a case study with children diagnosed with autism. In 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI) (pp. 419-420). IEEE.

[7] Kim, J., Wigram, T., & Gold, C. (2009). Emotional, motivational and interpersonal responsiveness of children with autism in improvisational music therapy. Autism, 13(4), 389-409.

[8] Kleinbub, J. R., & Ramseyer, F. T. (2021). rMEA: An R package to assess nonverbal synchronization in motion energy analysis time-series. Psychotherapy research : journal of the Society for Psychotherapy Research, 31(6), 817–830. https://doi.org/10.1080/10503307.2020.1844334

[9] Bharathi, G., Venugopal, A., & Vellingiri, B. (2019). Music therapy as a therapeutic tool in improving the social skills of autistic children. The Egyptian Journal of Neurology, Psychiatry and Neurosurgery, 55(1), 44. https://doi.org/10.1186/s41983-019-0091-x.

[10] Ragone, Grazia,. 2020. Designing Embodied Musical Interaction for Children with Autism. In The 22nd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '20). Association for Computing Machinery, New York, NY, USA, Article 104, 1–4. DOI:https://doi-org.ezproxy.sussex.ac.uk/10.1145/3373625.3417077

[11] Koole, S. L., & Tschacher, W. (2016). Synchrony in psychotherapy: A review and an integrative framework for the therapeutic alliance. Frontiers in Psychology, 7.

[12] Rolf, M., Hanheide, M., & Rohlfing, K. (2009). Attention via Synchrony: Making Use of Multimodal Cues in Social Learning. Autonomous Mental Development, IEEE Transactions On, 1, 55–67. https://doi.org/10.1109/TAMD.2009.2021091

[13] Paxton, A., & Dale, R. (2013). Argument disrupts interpersonal synchrony. Quarterly Journal of Experimental Psychology, 66(11), 2092–2102. https://doi.org/10.1080/17470218.2013.853089

[14] Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. Journal of Personality and Social Psychology, 76(6), 893–910. https://doi.org/10.1037/0022-3514.76.6.893

[15] Wiltermuth, S. S., & Heath, C. (2009). Synchrony and Cooperation. Psychological Science, 20(1), 1–5. https://doi.org/10.1111/j.1467-9280.2008.02253.x

[16] Stel, M., & Vonk, R. (2010). Mimicry in social interaction: Benefits for mimickers, mimickees, and their interaction. British Journal of Psychology (London, England: 1953), 101(Pt 2), 311–323. https://doi.org/10.1348/000712609X465424

[17] Miles, L. K., Griffiths, J. L., Richardson, M. J., and Macrae, C. N. (2010a). Too late to coordinate: contextual influences in behavioral synchrony. Eur. J. Soc. Psychol. 40, 52–60. doi: 10.1002/ejsp.721

[18] Delaherche, E., Chetouani, M., Mahdhaoui, A., Saint-georges, C., Viaux, S., & Cohen, D. (2012). Interpersonal Synchrony: A Survey of Evaluation Methods across Disciplines. IEEE Transactions on Affective Computing, 3, 349–365. https://doi.org/10.1109/T-AFFC.2012.12

[19] Bakeman, R., & Adamson, L. B. (1984). Coordinating Attention to People and Objects in Mother-Infant and Peer-Infant Interaction. Child Development, 55(4), 1278–1289. https://doi.org/10.2307/1129997

[20] Goldin-Meadow, S., & Butcher, C. (2000). Pointing toward two-word speech in young children. To appear in S. Kita (Ed.), Pointing: Where language, culture, and cognition meet. Manuscript under review.

[21] Jaffe J, Beebe B, Feldstein S, Crown CL, Jasnow MD. Rhythms of dialogue in infancy: coordinated timing in development. Monogr Soc Res Child Dev. 2001;66(2):i-viii, 1-132. PMID: 11428150.

[22] Leclère, C., Viaux, S., Avril, M., Achard, C., Chetouani, M., Missonnier, S., & Cohen, D. (2014). Why Synchrony Matters during Mother-Child Interactions: A Systematic Review. PLoS ONE, 9(12), e113571. https://doi.org/10.1371/journal.pone.0113571

[23] Fitzpatrick, P., Diorio, R., Richardson, M. J., & Schmidt, R. C. (2013). Dynamical methods for evaluating the time-dependent unfolding of social coordination in children with autism. Frontiers in Integrative Neuroscience, 7, Article 21.

[24] Fitzpatrick, P., Romero, V., Amaral, J. L., Duncan, A., Barnard, H., RichardsonM. J., *et al.* (2017b). Social motor synchronization: Insights for understandingsocial behavior in autism. J. Autism. Dev. Disord. 47, 2092–2107.

[25] Marsh, K. L., Isenhower, R. W., Richardson, M. J., Helt, M., Verbalis, A. D., Schmidt, R. C., & Fein, D. (2013). Autism and social disconnection in interpersonal rocking. Frontiers in Integrative Neuroscience, 7, Article 4. https://doi.org/10.3389/fnint.2013.00004

[26] Xavier, J., Gauthier, S., Cohen, D., Zahoui, M., Chetouani, M., Villa, F., Berthoz, A., & Anzalone, S. (2018). Interpersonal Synchronization, Motor Coordination, and Control Are Impaired During a Dynamic Imitation Task in Children With Autism Spectrum Disorder. Frontiers in Psychology, 9, 1467. https://doi.org/10.3389/fpsyg.2018.01467

[27] Granner-Shuman, M.; Dahan, A.; Yozevitch, R.; Gvirts Problovski, H.Z. (2021) The Association among Autistic Traits, Interactional Synchrony and Typical Pattern of Motor Planning and Execution in Neurotypical Individuals. Symmetry 2021, 13, 1034. https://doi.org/10.3390/sym13061034

[28] Georgescu, A. L., Koeroglu, S., Hamilton, A. F. de C., Vogeley, K., Falter-Wagner, C. M., & Tschacher, W. (2020). Reduced nonverbal interpersonal synchrony in autism spectrum disorder independent of partner diagnosis: A motion energy study. Molecular Autism, 11(1), 11. https://doi.org/10.1186/s13229-019-0305-1

[29] Pouw, W., Trujillo, J. P., & Dixon, J. A. (2020). The quantification of gesture–speech synchrony: A tutorial and validation of multimodal data acquisition using device-based and video-based motion tracking. Behavior Research Methods, 52(2), 723–740. https://doi.org/10.3758/s13428-019-01271-9

[30] Romero, V., Amaral, J., Fitzpatrick, P., Schmidt, R. C., Duncan, A. W., & Richardson, M. J. (2017). Can Low-cost Motion Tracking Systems Substitute a Polhemus System when Researching Social Motor Coordination in Children? Behavior Research Methods, 49(2), 588–601. https://doi.org/10.3758/s13428-016-0733-1

[31] Gashi, S., Di Lascio, E., Santini, S. 2019. Using Unobtrusive Wearable Sensors to Measure the Physiological Synchrony Between Presenters and Audience Members. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 3, 1, Article 13 (March 2019), 19 pages. DOI:https://doi-org.ezproxy.sussex.ac.uk/10.1145/3314400

[32] Ayache, J., Connor, A., Marks, S., Kuss, D. J., Rhodes, D., Sumich, A., & Heym, N. (2021). Exploring the "Dark Matter" of Social Interaction: Systematic Review of a Decade of Research in Spontaneous Interpersonal Coordination. Frontiers in Psychology, 12. https://www.frontiersin.org/article/10.3389/fpsyg.2021.718237

[33] Bloch, C., Vogeley, K., Georgescu, A. L., & Falter-Wagner, C. M. (2019). INTRApersonal Synchrony as Constituent of INTERpersonal Synchrony and Its Relevance for Autism Spectrum Disorder. Frontiers in Robotics and AI, 6, 73. https://doi.org/10.3389/frobt.2019.00073

[34] Ward, J. A., Richardson, D., Orgs, G., Hunter, K., & Hamilton, A. (2018, October). Sensing interpersonal synchrony between actors and autistic children in theatre using wrist-worn accelerometers. In Proceedings of the 2018 ACM International Symposium on Wearable Computers (pp. 148-155).

[35] Schmid Mast, M., Gatica-Perez, D., Frauendorfer, D., Nguyen, L., & Choudhury, T. (2015). Social Sensing for Psychology: Automated Interpersonal Behavior Assessment. Current Directions in Psychological Science, 24(2), 154–160. https://doi.org/10.1177/0963721414560811

[36] Choi, W., Pantofaru, C., & Savarese, S. (2013). A General Framework for Tracking Multiple People from a Moving Camera. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(7), 1577–1591. https://doi.org/10.1109/TPAMI.2012.248

[37] Fujiwara, K., & Yokomitsu, K. (2021). Video-based tracking approach for nonverbal synchrony: A comparison of Motion Energy Analysis and OpenPose. Behavior Research Methods, 53(6), 2700–2711. https://doi.org/10.3758/s13428-021-01612-7

[38] Tschacher, W., Rees, G. M., & Ramseyer, F. (2014). Nonverbal synchrony and affect in dyadic interactions. Frontiers in Psychology, 5, 1323. https://doi.org/10.3389/fpsyg.2014.01323

[39] Oullier, O., de Guzman, G. C., Jantzen, K. J., Lagarde, J., & Kelso, J. A. S. (2008). Social coordination dynamics: Measuring human bonding. Social Neuroscience, 3(2), 178–192. https://doi.org/10.1080/17470910701563392

[40] Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. Cognitive Science, 29(6), 1045–1060. https://doi.org/10.1207/s15516709cog0000_29

[41] Weichert, F., Bachmann, D., Rudak, B., & Fisseler, D. (2013). Analysis of the Accuracy and Robustness of the Leap Motion Controller. Sensors (Basel, Switzerland), 13(5), 6380–6393. https://doi.org/10.3390/s130506380

[42] Schmidt, R. C., Carello, C., & Turvey, M. T. (1990). Phase transitions and critical fluctuations in the visual coordination of rhythmic movements between people. Journal of Experimental Psychology: Human Perception and Performance, 16, 227-247.

[43] Avril, M., Leclère, C., Viaux, S., Michelet, S., Achard, C., Missonnier, S., Keren, M., Cohen, D., & Chetouani, M. (2014). Social signal processing for studying parent–infant interaction. Frontiers in Psychology, 5. https://www.frontiersin.org/article/10.3389/fpsyg.2014.01437

[44] Watanabe T. A study of motion-voice synchronization. Bulletin of JSME. 1983;26(222):2244–50. 10.1299/jsme1958.26.2244

[45] Ramseyer, F., & Tschacher, W. (2011). Nonverbal synchrony in psychotherapy: Coordinated body movement reflects relationship quality and outcome. Journal of Consulting and Clinical Psychology, 79(3), 284–295.

[46] Nakano, N., Sakura, T., Ueda, K., Omura, L., Kimura, A., Iino, Y., Fukashiro, S., & Yoshioka, S. (2020). Evaluation of 3D Markerless Motion Capture Accuracy Using OpenPose With Multiple Video Cameras. Frontiers in Sports and Active Living, 2. https://www.frontiersin.org/article/10.3389/fspor.2020.00050

[47] Fujiwara K., Daibo I. (2014). The extraction of nonverbal behaviors: using video images and speech-signal analysis in dyadic conversation. J. Nonverbal Behav. 38 377–388. 10.1007/s10919-014-0183-3

[48] Schmid Mast, M., Gatica-Perez, D., Frauendorfer, D., Nguyen, L., & Choudhury, T. (2015). Social Sensing for Psychology: Automated Interpersonal Behavior Assessment. Current Directions in Psychological Science, 24(2), 154–160. https://doi.org/10.1177/0963721414560811

[49] Issartel J., Marin L., Gaillotb P., Bardainnec T., Cadopia M. (2006). A practical guide to time—frequency analysis in the study of human motor behavior: the contribution of wavelet transform. J. Motor Behav. 38 139–159. 10.3200/JMBR.38.2.139-159

[50] Issartel, J., Bardainne, T., Gaillot, P., & Marin, L. (2015). The relevance of the cross-wavelet transform in the analysis of human interaction—A tutorial. Frontiers in Psychology, 5, Article 1566. https://doi.org/10.3389/fpsyg.2014.01566

[51] Bernieri, F. J., Reznick, J. S., & Rosenthal, R. (1988). Synchrony, pseudosynchrony, and dissynchrony: Measuring the entrainment process in mother-infant interactions. Journal of Personality and Social Psychology, 54(2), 243–253. https://doi.org/10.1037/0022-3514.54.2.243

[52] Bernieri, F. J., & Rosenthal, R. (1991). Interpersonal coordination: Behavior matching and interactional synchrony. In R. S. Feldman & B. Rimé (Eds.), Fundamentals of nonverbal behavior (pp. 401–432). Cambridge University Press; Editions de la Maison des Sciences de l'Homme.

[53] Georgescu, A., Koehler, J., Weiske, J., Vogeley, K., Koutsouleris, N., & Falter-Wagner, C. (2019). Machine Learning to Study Social Interaction Difficulties in ASD. Frontiers In Robotics And AI, 6. doi: 10.3389/frobt.2019.00132

[54] Koehler, J. C., Georgescu, A. L., Weiske, J., Spangemacher, M., Burghof, L., Falkai, P., Koutsouleris, N., Tschacher, W., Vogeley, K., & Falter-Wagner, C. M. (2021). Brief Report: Specificity of Interpersonal Synchrony Deficits to Autism Spectrum Disorder and Its Potential for Digitally Assisted Diagnostics. Journal of autism and developmental disorders, 10.1007/s10803-021-05194-3. Advance online publication. https://doi.org/10.1007/s10803-021-05194-3

[55] Boker, S. M., Xu, M., Rotondo, J. L., & King, K. (2002). Windowed cross–correlation and peak picking for the analysis of variability in the association between behavioral timeseries. Psychological Methods,7(3), 338–355. doi: 10.1037/1082-989X.7.3.338

[56] Ramseyer, F., & Tschacher, W. (2006). Synchrony: A Core Concept for A Constructivist Approach to Psychotherapy. Constructivism in the Human Sciences, 11(1–2), 150–171.

[57] Valencia, K., Rusu, C., Quiñones, D., & Jamet, E. (2019). The Impact of Technology on People with Autism Spectrum Disorder: A Systematic Literature Review. Sensors (Basel, Switzerland), 19(20), 4485. https://doi.org/10.3390/s19204485

[58] Burleson, W., Newman, N., & Brotman, R. (2012). Empowering Independent Living for People with Autism: Designing Supportive, Low-Cost, Interactive E-Health Environments. In M. Bang & E. L. Ragnemalm (Eds.), Persuasive Technology. Design for Health and Safety (pp. 13–30). Springer. https://doi.org/10.1007/978-3-642-31037-9_2

[59] Nagib, W., & Williams, A. (2017). Oward an autism-friendly home environment. Taylor & Francis, 32,(2). https://doi-org.ezproxy.sussex.ac.uk/10.1080/02673037.2016.1181719

[60] Tentori, M., Escobedo, L., & Balderas, G. (2015). A Smart Environment for Children with Autism. IEEE Pervasive Computing, 14, 42–50. https://doi.org/10.1109/MPRV.2015.22

[61] Forti, S., Colombo, B., Clark, J., Bonfanti, A., Molteni, S., Crippa, A., Antonietti, A., & Molteni, M. (2020). Soundbeam imitation intervention: Training children with autism to imitate meaningless body gestures through music. Advances in Autism, ahead-of-print(ahead-of-print). https://doi.org/10.1108/AIA-07-2019-0023

[62] McGowan, J., Leplâtre, G., & McGregor, I. (2017). CymaSense: A Novel Audio-Visual Therapeutic Tool for People on the Autism Spectrum. 62–71. https://doi-org.ezproxy.sussex.ac.uk/10.1145/3132525.3132539

[63] Parsons, S., Mitchell, P., & Leonard, A. (2004). The use and understanding of virtual environments by adolescents with autistic spectrum disorders. Journal of Autism and Developmental Disorders, 34(4), 449–466. https://doi.org/10.1023/b:jadd.0000037421.98517.8d

[64] Valori, I., McKenna-Plumley, P. E., Bayramova, R., & Farroni, T. (2021). Perception and Motion in Real and Virtual Environments: A Narrative Review of Autism Spectrum Disorders. Frontiers in Psychology, 12, 2869. https://doi.org/10.3389/fpsyg.2021.708229

[65] Rajendran, G. (2013). Virtual environments and autism: A developmental psychopathological approach: Virtual environments and autism. Journal of Computer Assisted Learning, 29(4), 334–347. https://doi.org/10.1111/jcal.12006

[66] Alcorn, A. M., Ainger, E., Charisi, V., Mantinioti, S., Petrović, S., Schadenberg, B. R., Tavassoli, T., & Pellicano, E. (2019). Educators' Views on Using Humanoid Robots With Autistic Learners in Special Education Settings in England. Frontiers in Robotics and AI, 6, 107. https://doi.org/10.3389/frobt.2019.00107

[67] Lux, V., Non, A. L., Pexman, P. M., Stadler, W., Weber, L. A. E., & Krüger, M. (2021). A Developmental Framework for Embodiment Research: The Next Step Toward Integrating Concepts and Methods. Frontiers in Systems Neuroscience, 15. https://www.frontiersin.org/article/10.3389/fnsys.2021.672740

[68] Wang, I., & Ruiz, J. (2021). Examining the Use of Nonverbal Communication in Virtual Agents. International Journal of Human–Computer Interaction, 37(17), 1648–1673. https://doi.org/10.1080/10447318.2021.1898851

[69] Kinect V2, Microsoft inc. USA Kinect for Windows features [Internet]. Cited 23 March 2015. Available from: http://www.microsoft.com/en-us/kinectforwindows/meetkinect/features.aspx.

[70] Gumtau, S., Newland, P., Creed, C., & Kunath, S. (2005). MEDIATE – a responsive environment designed for children with autism. https://doi.org/10.14236/ewic/AD2005.14

[71] Hoberman, P., Pares, N., and Pares, R. (1999). Interaction-Driven VirtualReality Application Design: El ball del fanalet or lightpools. In Proceedings of International Conference on Virtual Systems and Multimedia. 40, 55, 87

[72] 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. Proceedings of the The 15th International Conference on Interaction Design and Children, 262–274. https://doi-org.ezproxy.sussex.ac.uk/10.1145/2930674.2930695

[73] Crowell, C., Sayis, B., & Bravo, A., & Paramithiotti, A., (2018). GenPlay: Generative Playscape. 1-6. 10.1145/3170427.3180653.

[74] Garzotto, F., & Gelsomini, M. (2018). Magic Room: A Smart Space for Children with Neurodevelopmental Disorder. IEEE Pervasive Computing, 17, 1. https://doi.org/10.1109/MPRV.2018.011591060

[75] Lima, J. L., Axt, G., Teixeira, D. S., Monteiro, D., Cid, L., Yamamoto, T., Murillo-Rodriguez, E., & Machado, S. (2020). Exergames for Children and Adolescents with Autism Spectrum Disorder: An Overview. Clinical Practice and Epidemiology in Mental Health: CP & EMH, 16, 1–6. https://doi.org/10.2174/1745017902016010001

[76] Rafiei Milajerdi, H., Sheikh, M., Najafabadi, M. G., Saghaei, B., Naghdi, N., & Dewey, D. (2021). The Effects of Physical Activity and Exergaming on Motor Skills and Executive Functions in Children with Autism Spectrum Disorder. Games for Health Journal, 10(1), 33–42. https://doi.org/10.1089/g4h.2019.0180

[77] Wigram, T.; Gold, C. Music therapy in the assessment and treatment of autistic spectrum disorder: Clinical application and research evidence. Child Care Health Dev. 2006, 32, 535–542.

[78] Srinivasan, S.M.; Bhat, A.N. A review of "music and movement" therapies for children with autism: Embodied interventions for multisystem development. Front. Integr. Neurosci. 2013, 7, 22.

[79] Geretsegger, M.; Elefant, C.; A Mössler, K.; Gold, C. Music therapy for people with autism spectrum disorder. Cochrane Database Syst. Rev. 2014, 2014, CD004381.

[80] Ragone, G., Good, J., & Howland, K. (2021). How Technology Applied to Music-Therapy and Sound-Based Activities Addresses Motor and Social Skills in Autistic Children. Multimodal Technologies and Interaction, 5(3), 11. doi:10.3390/mti5030011

[81] Magee, W.L. & Burland, K. (2008). Using electronic music technologies in clinical practice: opportunities, limitations and clinical indicators. British Journal of Music Therapy, 22(1), 3-15.

[82] Cibrian, F.L. *et al.*, 2017 BendableSound: An elastic multisensory surface using touch-based interactions to assist children with severe autism during music therapy. Int J Hum-CompStudies. 107, May (2017), 22–37.

[83] Ragone, G., Good, J., & Howland, K. (2020). OSMoSIS: Interactive sound generation system for children with autism. Proceedings of the 2020 ACM Interaction Design and Children Conference: Extended Abstracts, 151–156. https://doi.org/10.1145/3397617.3397838

[84] Hoehl, S., Fairhurst, M., & Schirmer, A. (2020). Interactional synchrony: Signals, mechanisms and benefits | Social Cognitive and Affective Neuroscience Oxford Academic. Oxford Academic, 16(1–2), 5–18. https://doi-org.ezproxy.sussex.ac.uk/10.1093/scan/nsaa024

[85] Ragone, Grazia,. 2020. Designing Embodied Musical Interaction for Children with Autism. In The 22nd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '20). Association for Computing Machinery, New York, NY, USA, Article 104, 1–4. DOI:https://doi-org.ezproxy.sussex.ac.uk/10.1145/3373625.3417077

[86] Campbell, D. T., & Yin, R. K. (2018). Case study research and applications: Design and methods | Campbell, Donald Thomas; Yin, Robert K. | download (Sixth). SAGE Publications. https://b-ok.cc/book/3705537/e8d32d

[87] Glass, D., & Yuill, N. (2020). Synchronised shared spaces: Design considerations for tablet technology. IDC. https://doi.org/10.1145/3397617.3397846

[88] Glass & Yuill (2021). Social Motor Synchrony in Autism Spectrum Conditions: A Systematic Review [Manuscript submitted for publication]. School of Psychology, University of Sussex.

[89] Quintin, E.-M. (2019). Music-Evoked Reward and Emotion: Relative Strengths and Response to Intervention of People With ASD. Frontiers in Neural Circuits, 13, 49.

[90] Stanton, T. R., & Spence, C. (2020). The Influence of Auditory Cues on Bodily and Movement Perception. Frontiers in Psychology, 10, 3001. https://doi.org/10.3389/fpsyg.2019.03001

[91] Effenberg, A. (2005). Movement Sonification: Effects on Perception and Action. Multimedia, IEEE, 12, 53–59. https://doi.org/10.1109/MMUL.2005.31

[92] Cao, Z., Simon, T., Wei, S.-E., & Sheikh, Y. (2017). Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. ArXiv:1611.08050 [Cs]. http://arxiv.org/abs/1611.08050

[93] Park, S., & Aggarwal, J. K. (2006). Simultaneous tracking of multiple body parts of interacting persons. Computer Vision and Image Understanding, 102(1), 1–21. https://doi.org/10.1016/j.cviu.2005.07.011

[94] Mohd Nordin, A., Ismail, J., & Kamal Nor, N. (2021). Motor Development in Children With Autism Spectrum Disorder. Frontiers in Pediatrics, 9. https://www.frontiersin.org/article/10.3389/fped.2021.598276

[95] American Psychiatric Association, 2013. (2013). Diagnostic and statistical manual of mental Disorders (5th ed.). Arlington, VA: Author.

[96] Argyle, M., Bodily Communication. Methuen, London, (1975).

[97] Pentland, A., Honest Signals - How They Shape Our World. MIT Press, Cambridge, MA, (2008).