Machine Learning to Study Social Interaction Difficulties in ASD

Alexandra Livia Georgescu [1,2*†], Jana Christina Koehler [3*†], Johanna Weiske [3],
Kai Vogeley [2,4], Nikolaos Koutsouleris [3] and Christine Falter-Wagner [3,5]

[1] Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom, [2] Department of Psychiatry and Psychotherapy, University Hospital of Cologne, Cologne, Germany, [3] Department of Psychiatry and Psychotherapy, Medical Faculty, LMU Munich, Munich, Germany, [4] Institute of Neuroscience and Medicine, Cognitive Neuroscience (INM-3), Research Center Juelich, Jülich, Germany, [5] Institute of Medical Psychology, Medical Faculty, LMU Munich, Munich, Germany

Autism Spectrum Disorder (ASD) is a spectrum of neurodevelopmental conditions characterized by difficulties in social communication and social interaction as well as repetitive behaviors and restricted interests. Prevalence rates have been rising, and existing diagnostic methods are both extremely time and labor consuming. There is an urgent need for more economic and objective automatized diagnostic tools that are independent of language and experience of the diagnostician and that can help deal with the complexity of the autistic phenotype. Technological advancements in machine learning are offering a potential solution, and several studies have employed computational approaches to classify ASD based on phenomenological, behavioral or neuroimaging data. Despite of being at the core of ASD diagnosis and having the potential to be used as a behavioral marker for machine learning algorithms, only recently have movement parameters been used as features in machine learning classification approaches. In a proof-of-principle analysis of data from a social interaction study we trained a classification algorithm on intrapersonal synchrony as an automatically and objectively measured phenotypic feature from 29 autistic and 29 typically developed individuals to differentiate those individuals with ASD from those without ASD. Parameters included nonverbal motion energy values from 116 videos of social interactions. As opposed to previous studies to date, our classification approach has been applied to non-verbal behavior objectively captured during naturalistic and complex interactions with a real human interaction partner assuring high external validity. A machine learning approach lends itself particularly for capturing heterogeneous and complex behavior in real social interactions and will be essential in developing automatized and objective classification methods in ASD.

Keywords: autism spectrum disorder, machine learning, nonverbal synchrony, support vector machine, motion energy analysis, classification, intrapersonal synchrony, nested cross-validation

INTRODUCTION

Autism spectrum disorder (ASD) is an umbrella term for neurodevelopmental conditions characterized by severe difficulties in social interaction and communication, as well as by repetitive behaviors and restricted interests (American Psychiatric Association, 2013). The prevalence rates of ASD are on the rise (Elsabbagh et al., 2012) and diagnostic services are experiencing an increased demand, in particular in adults seeking diagnostic advice (Murphy et al., 2011). Diagnostics according to medical guidelines are time-consuming, the clinical assessment is complicated by the phenotypic heterogeneity and the language-dependency of assessment with verbal skills being affected by the ASD.

Recently, computational methods of classification have been employed to increase diagnostic reliability and efficiency (Thabtah, 2018). In particular, machine learning (ML) employs algorithms to uncover patterns in complex datasets, which are utilized to improve decision making. ASD diagnostics come down to a decision-making problem that can be supported by automated models (classifiers) using ML to decide whether a newly assessed patient has ASD or not. This works by splitting available data into a training set, on which an algorithm is trained, which is then applied to a test set, resulting in a measure of accuracy of the resulting model. Without making assumptions ML finds classification solutions in a data-driven, bottom-up approach that can be applied to individual prediction making (Dwyer et al., 2018). The primary purposes of using ML are (1) to reduce assessment time to reach a diagnostic decision in order to provide quicker access to health care services, (2) to improve diagnostic reliability, and (3) diagnostic validity by reducing dimensionality of input data so as to identify those features that have the most diagnostic value in ASD (Thabtah, 2018). However, first applications of ML in studies on autism diagnostics have been inconsistent in terms of methodology and outcome, with inconsistent classification accuracy and specificity.

The aim of the present paper is twofold: First, we aim to give an overview of previous research that has attempted to apply ML methods to the classification of ASD, while suggesting guidelines for future research in terms of setup and algorithm design. Second, in a proof-of-principle analysis of data from a social interaction study we aim to establish the potential of using full-body non-verbal behavior data extracted from video recordings of naturalistic social interactions to classify autistic adults.

MACHINE LEARNING APPLICATIONS IN THE CLASSIFICATION OF ASD

First ML attempts in ASD have been used with the aim of shortening ADOS [Autism Diagnostic Observation Schedule, (Lord et al., 2000)] and ADI-R [Autism Diagnostic Interview, (Lord et al., 1994)] administration time by item-reduction yielding a classification accuracy of autism vs. typically-developing (TD) individuals of up to 99.9% (Wall et al., 2012a,b; Bone et al., 2016). In a similar attempt to predict case status words and expressions contained in 8 year old children's developmental

evaluations across a network of multiple clinical sites were used for algorithm development (Maenner et al., 2016) with 86.5% prediction accuracy and high concordance with the respective clinician. Home videos of children have been rated by naïve and/or expert raters in terms of ASD-typical behavior and ratings fed into a predictive model along with other features of the diagnostic process (Glover et al., 2018; Tariq et al., 2018). However, while all these first studies using ML in ASD yield fairly high accuracies, the features utilized for classification are still highly subjective and not independent of the respective clinician who bases the diagnostic decision on just those features (circularity). Importantly, when using subjectively influenced data, resulting classification algorithms must be validated in an independent sample in order to prevent circularity.

An increasing number of studies are also using ML to separate individuals with ASD from TD individuals based on neuroimaging data. For example, Ecker et al. (2010) used regional gray and white matter volume measures from whole-brain structural MRI scans of individuals with ASD to investigate their diagnostic value. They used a common variant of ML, the support vector machine (SVM). This is an algorithm aiming at finding a boundary (the so-called "hyperplane") that can be used to optimally classify groups while being able to generalize to new cases (Dwyer et al., 2018). In their sample, the SVM correctly classified individuals with ASD and controls on the basis of their neuroanatomy with about 80% accuracy (Ecker et al., 2010). These original observations are supported by findings from several other neuroimaging studies with similar levels of classification accuracy in younger age groups (Wee et al., 2014), females with ASD (Calderoni et al., 2012) and with various anatomical and functional measurements (Coutanche et al., 2011). These results based on objective data are very promising, although not widely applicable due to high costs.

WHOLE-BODY MOVEMENTS AS A FEATURE IN ML ALGORITHMS IN ASD

Another source of objective data with high potential for diagnostics can be found in the motor domain. Approximately 80% of children with ASD are suspected to exhibit pronounced motor difficulties (Green et al., 2009). Difficulties with balance, gait, movement speed and timed movements have demonstrated to hold a high level of discrimination between children with ASD and TD children (Jansiewicz et al., 2006) and correlate strongly with measures of social and communicative functioning (Parma and de Marchena, 2016). Hence, movement parameters of social interactions in ASD should be investigated for their potential as a diagnostic marker.

Particularly relevant for ASD motor symptomology are gestures and non-verbal communicative behaviors (Georgescu et al., 2014). Accordingly, atypical non-verbal behavior has been included in the DSM-5 criteria for ASD. Yet, the assessment is not straightforward or standardized so far and is hampered by the fact that non-verbal behavior is not necessarily reduced in ASD, but abnormal in the *quality* of its temporal coordination with

own verbal output (de Marchena and Eigsti, 2010) and that of an interaction partner. Literature provides evidence for aberrations in temporal processing (Allman and Falter, 2015) and time experience in ASD (Vogel et al., 2019), potentially affecting non-verbal communication. In fact, findings have shown that ASD can be characterized by increased temporal resolution associated with the severity of (non-verbal) communication impairments in ASD (Falter et al., 2012, 2013; Menassa et al., 2018; but see Isaksson et al., 2018).

Recently, movement in ASD has taken up increasing interest (for a review see Bo et al., 2016). In a proof-of-concept study to explore whether low-functioning children with ASD could be identified by means of a kinematic analysis of a simple motor task, 15 children with ASD and 15 TD children (2–4 years) were asked to pick up a ball and drop it into a hole while their movements were recorded using a motion tracker (Crippa et al., 2015). Seventeen kinematic parameters were extracted from the upper-limb movement and seven of these were found significant for discrimination. The classifier distinguished ASD from non-ASD with a classification accuracy of 96.7%, suggesting the validity of assuming a motor signature of ASD. Reach and throw movements of 10 ASD and 10 TD children were analyzed for "peculiar features" using ML and fed into a classification algorithm yielding an accuracy of 92.5% (Perego et al., 2009). Furthermore, Li et al. (2017) extracted 40 kinematic parameters of imitative movements and identified 9 of them that best describe variance of participant groups, resulting in a classification accuracy of 93%.

These studies demonstrate the potential of using kinematic biomarkers in diagnostics of ASD. However, the movements under investigation were staged, thus, highly unnatural. Yet, it has been established that individuals with ASD have particular difficulties with spontaneous "on-line" social interaction requiring intuitive decisions and behavior (Redcay et al., 2013) constituting an urgent need to move this type of research to more external validity and investigate movement in a more naturalistic context.

CLASSIFICATION USING INTRAPERSONAL SYNCHRONY: A PROOF-OF-CONCEPT STUDY

Whole-body movements in more naturalistic conversations were tested for their classification potential in 29 high functioning adults with ASD and 29 TD individuals. The data for this investigation came from a study on interpersonal coordination in dyadic interactions (Georgescu et al., under revision). The autistic participants were diagnosed and recruited at the Autism Outpatient Clinic of the Department of Psychiatry, University Hospital Cologne, Germany. The sample included only patients with the diagnoses high-functioning autism (ICD-10: F84.0) or Asperger syndrome (ICD-10: F84.5). Two medical specialists confirmed the diagnosis independently in clinical interviews, according to the criteria of the International Classification of Diseases (ICD-10) and supplemented by

extensive neuropsychological examination. The TD sample was recruited online from the student and staff population at the University of Cologne and the University Hospital of Cologne, Germany. The study was conducted with the approval of the local ethics committee of the Medical Faculty of the University of Cologne. Participants were paired to conduct five 5 min social interaction tasks. Conversational dyads consisted of either two TD individuals, two individuals with ASD or a TD individual with an individual with ASD. An ice-breaker task, two debating tasks, a meal-planning task and a roleplay were included resulting in a total of 145 videos of social interactions (for more information, see Georgescu et al., under revision). All conversations were recorded in a room with standardized artificial lighting and using a high-definition video camera (Panasonic DV C Pro HD P2), mounted on a tripod 320 cm away from the chairs which were 60 cm apart from each other. Since one of the MIXED dyads did not understand instructions on the ice-breaker task, for the purpose of this analysis the whole task was abandoned, resulting in a total of 116 videos submitted for final analysis. Intrapersonal Synchrony between the head and upper body was quantified using Motion Energy Analysis, a widely used semi-automated frame-differencing method that continuously monitors the amount of movement occurring in manually pre-defined regions of interest and the method of lagged cross-correlations (Nagaoka and Komori, 2008; MEA; Altmann, 2011; Ramseyer and Tschacher, 2011). MEA offers the advantage of a constraint-free, objective analysis tool for non-verbal behavior (e.g., Ramseyer and Tschacher, 2011; Schmidt et al., 2012; Paxton and Dale, 2013). This method has been used to capture body movement in different contexts (e.g., Grammer et al., 1999; Ramseyer and Tschacher, 2011, 2014; Schmidt et al., 2012, 2014; Paxton and Dale, 2013). MEA and other frame-differencing methods have been successfully used in clinical research before (e.g., Kupper et al., 2015) and in particular in autism (Noel et al., 2017; Romero et al., 2017, 2018). We followed the MEA pipeline described in Ramseyer and Tschacher (2014). We manually selected two regions of interest (ROI) for each participant, covering (1) the head and (2) the rest of the body including the legs. Changes in grayscale values in these ROIs were detected and separately recorded as two continuous time series measuring the amount of movement in the head and the body region of each person. Data were submitted for quantification of Intrapersonal Synchrony (for more information on the MEA procedure in general, please see Ramseyer and Tschacher, 2014 and on this sample, Georgescu et al., under revision). Input time series were smoothed and scaled to account for different-sized ROIs using custom software in R (package rMEA, Kleinbub and Ramseyer, 2019) and cross-correlated in windows of 60 s with a time lag of ± 5 s (step size 0.04 s). Windows were not allowed to overlap. The resulting 1,004 lagged cross-correlations were then z-standardized and aggregated over the four conditions for every participant, yielding 4,016 features per participant which were implemented in the open-source machine learning tool NeuroMiner (https://www.pronia.eu/neurominer/). A support vector machine with linear kernel was chosen as a classification algorithm, a multivariate supervised learning technique widely

TABLE 1 | Performance metrics of the ASD vs. TD SVM classifier.

True positives/true negatives	28/16
False positives/false negatives	13/1
Accuracy [%]	75.9
Sensitivity [%]	96.6
Specificity [%]	55.2
Area under the curve	0.71

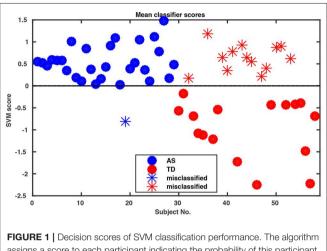*For detailed explanation of performance metrics please refer to Dwyer et al. (2018).*

FIGURE 1 | Decision scores of SVM classification performance. The algorithm assigns a score to each participant indicating the probability of this participant as belonging to Group 1 or 2 (in our case ASD vs. TD) where the decision boundary between the two groups is zero. Notably, our algorithm misclassified only one of the ASD participants.

used in psychiatric research (Bone et al., 2016; Duda et al., 2016). Our repeated nested k-fold cross-validation (CV) structure consisted of 10-folds and five permutations for the outer cross-validation cycle (CV_2) and repeated 5-by-5-fold inner cross-validation cycle (CV_1), with participants being shuffled prior to each definition of folds. This way, the data available for training was maximized while ensuring enough heterogeneity within the inner test sample to avoid overfitting and create stable models. Parameter optimization was performed in CV_1, while model performance was evaluated in CV_2. Prior to analysis, data was preprocessed using principal component analysis (PCA) for dimensionality reduction, retaining the principal components that cumulatively explained 80% of the variance in each CV_1 fold, and subsequently, scaled feature-wise from 0 to 1. The slack parameter C was estimated in the inner CV cycle using eight parameters ranging from 0.015625 to 16. Overall classification performance resulted in 75.9% accuracy (**Table 1**). Remarkably, sensitivity was 96.6%, correctly classifying all but one individual with ASD (**Figure 1**).

Thus, with a portable and inexpensive video-setup in a naturalistic setting and a semi-automated analysis pipeline, we reached a good diagnostic classification of ASD within four 5 min interaction excerpts on the mere basis of objective motion data. Feeding further clinical and interaction variables into the

BOX 1 | Minimum requirements for reliable clinical application of ML in ASD research (adapted from Dwyer et al., 2018)

- Combination of objective variables and standard diagnostic measures as input features to classify ASD.
- Use of nested CV as a standard procedure.
- Prevent unstable model outcomes through k-fold CV.

algorithm promises a high potential for classification (see Future Perspectives section).

METHODOLOGICAL ISSUES IN MACHINE LEARNING APPROACHES TO CLASSIFYING ASD

Unlike e.g., Bone et al. (2016) or Li et al. (2017), most ML studies in ASD research have relied on simple cross-validation (CV) methods. This increases the likelihood of choosing an overly optimistic model (Cawley and Talbot, 2010). We therefore suggest the application of a second layer of CV to allow for parameter selection and model performance evaluation to not be performed on the same data and to prevent overfitting. The test fold is completely held out until parameter optimization within the inner CV cycle is achieved by splitting the training data once more into an (inner) test and (inner) training set. The optimized models can then be tested for generalizability on the outer test fold. This so-called nested CV maximizes generalizability and has now been established as a gold standard procedure in psychiatric research (Dwyer et al., 2018). In order to account for the small sample sizes in ASD research, often predictions are made in a leave-one-out approach whereby only one individual's data is held out in the test set while parameters are optimized on the others (Crippa et al., 2015; Li et al., 2017). Especially, for ASD with its highly heterogeneous phenotype, leave-one-out creates overly variable test sets, rendering model outcomes unstable (Varoquaux et al., 2017). This can be prevented through k-fold nested CV and simultaneous permutation of individual data sets within the inner cross-validation cycle (Dwyer et al., 2018). An overview of best-practice standards is outlined below.

FUTURE PERSPECTIVES

Impairments of non-verbal communication are seen across the entire spectrum of ASD warranting the use as a behavioral biomarker. Yet, its intricacy requires multivariate analysis methods to capture complex interdependencies across domains. Machine learning offers the potential to incorporate high-dimensional data for the detection of underlying mechanisms and classification if certain minimum practice requirements are fulfilled (see **Box 1**).

In our proof-of-principle study, we were able to classify high-functioning adults with ASD from TD adults on the mere basis of non-verbal intrapersonal motion synchrony in

social interactions with an accuracy of 75.9%, which can be regarded a conservative estimate on the basis of a state-of-the art ML approach. Due to relatively small sample sizes available with high phenomenological heterogeneity in ASD, it is of utmost importance to choose adequate methods of cross-validation in order to maximize generalizability. The use of repeated nested cross-validation prevents overfitting and should be incorporated as a standard procedure in ML applications. However, given our rather limited sample size, the next steps for future research will be to apply the resulting algorithm to a completely new and larger data set and to investigate its transdiagnostic specificity across different psychiatric disturbances.

Future research should furthermore consider combining multiple non-verbal communication parameters and clinical data (e.g., questionnaires) in order to improve prediction and classification accuracy further and to possibly detect potential associations across domains. For instance, peculiarities in eye-gaze (Merin et al., 2007; Georgescu et al., 2013) and facial expression (McIntosh et al., 2006) in ASD demonstrate feasible approaches.

One future avenue would be to explore methods to quantify non-verbal behavior in a fully-automated fashion. In the present proof-of-principle study, a dataset was used that was analyzed using MEA, a classic frame-differencing approach. It has been shown that MEA is able to capture movements and even complex coordinative patterns to a similar extent as more expensive motion capture equipment such as the Polhemus (Romero et al., 2017). A main advantage for autism research of this method of extracting whole-body motor movement is that it does not involve any wearable technology. Given the hypersensitivity exhibited by many individuals with ASD, not having to add any attachable piece of equipment or body suit to their bodies is helpful. However, while MEA automatically detects pixel changes, corresponding regions of interest are drawn in manually. Although resulting values are standardized, there remains a subjective component. Computer vision tools that can estimate the coordinates of limb positions and even extract gaze location and body poses would offer similar benefits while balancing out subjective biases in the motion extraction process (Marín-Jiménez et al., 2014; Mehta et al., 2017; Tome et al., 2017; Cao et al., 2018). In addition, they offer even more flexibility, given it could be possible to include less strict and standardized experimental setups (no requirement for standardized camera or lighting conditions). However, the validity for movement extraction compared to other standard motion capture methods has not been demonstrated yet. Moreover, such tools vary greatly with respect to their susceptibility to tracking failures, or the type of videos they can support (single vs. multiple agent, indoor vs. outdoor etc.). Overall, with the current methodology that is available for motion extraction, the present semi-automated method offers a realistically applicable diagnostic value. Nevertheless, incredible advances are being

made (Li et al., 2018; Tran et al., 2018) such that they are very promising tools for future non-verbal behavior in autism research and beyond.

Taken together, given the recent advances in predictive psychiatry, adequately applied ML offers the potential to fully capture the autistic phenotype in all its complexity with sufficient specificity across psychiatric disorders with a special focus on the spontaneous non-verbal behavior during social encounters with others and irrespective of clinician or site.

DATA AVAILABILITY STATEMENT

The video datasets generated and analysed during the current study are not publicly available due this being identifiable patient data from a sample that did not consent to their data being shared in any form.

ETHICS STATEMENT

Written informed consent was obtained from all participants in accordance with the Declaration of Helsinki (1964). All participants received a monetary compensation for their participation of 50 Euros and were debriefed at the end. The study was conducted with the approval of the local ethics committee of the Medical Faculty of the University of Cologne.

AUTHOR CONTRIBUTIONS

AG, JK, and CF-W contributed equally to the drafting of this manuscript. AG provided the data. JW and JK performed the statistical analysis. All authors contributed to the manuscript revision, read, and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Allman, M., and Falter, C. (2015). "Abnormal timing and time perception in autism spectrum disorder?: a review of the evidence," in *Time Distortions in Mind: Temporal Processing in Clinical Populations,* eds A. Vatakis and M. Allman (Leiden; Boston, MA: Brill), 37–56.

Altmann, U. (2011). "Investigation of movement synchrony using windowed cross-lagged regression," in *Analysis of Verbal and Nonverbal Communication and Enactment. The Processing Issues.* Lecture Notes in Computer Science, Vol. 6800, eds A. Esposito, A. Vinciarelli, K. Vicsi, C. Pelachaud, and A. Nijholt (Berlin; Heidelberg: Springer).

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®).* American Psychiatric Publishing. doi: 10.1176/appi.books.9780890425596

Bo, J., Lee, C.-M., Colbert, A., and Shen, B. (2016). Do children with autism spectrum disorders have motor learning difficulties? *Res. Autism Spectr. Disord.* 23, 50–62. doi: 10.1016/j.rasd.2015.12.001

Bone, D., Bishop, S. L., Black, M. P., Goodwin, M. S., Lord, C., and Narayanan, S. S. (2016). Use of machine learning to improve autism screening and diagnostic instruments: effectiveness, efficiency, and multi-instrument fusion. *J. Child Psychol. Psychiatr.* 57, 927–937. doi: 10.1111/jcpp.12559

Calderoni, S., Retico, A., Biagi, L., Tancredi, R., Muratori, F., and Tosetti, M. (2012). Female children with autism spectrum disorder: an insight from mass-univariate and pattern classification analyses. *NeuroImage* 59, 1013–1022. doi: 10.1016/j.neuroimage.2011.08.070

Cao, Z., Hidalgo, G., Simon, T., Wei, S.-E., and Sheikh, Y. (2018). OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields. *ArXiv:1812.08008.* doi: 10.1109/CVPR.2017.143

Cawley, G. C., and Talbot, N. L. C. (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. *J. Machine Learn. Res.* 11, 2079–2107.

Coutanche, M. N., Thompson-Schill, S. L., and Schultz, R. T. (2011). Multi-voxel pattern analysis of fMRI data predicts clinical symptom severity. *NeuroImage* 57, 113–123. doi: 10.1016/j.neuroimage.2011.04.016

Crippa, A., Salvatore, C., Perego, P., Forti, S., Nobile, M., Molteni, M., et al. (2015). Use of machine learning to identify children with autism and their motor abnormalities. *J. Autism Dev. Disord.* 45, 2146–2156. doi: 10.1007/s10803-015-2379-8

de Marchena, A., and Eigsti, I. M. (2010). Conversational gestures in autism spectrum disorders: asynchrony but not decreased frequency. *Autism Res.* 3, 311–322. doi: 10.1002/aur.159

Duda, M., Ma, R., Haber, N., and Wall, D. P. (2016). Use of machine learning for behavioral distinction of autism and ADHD. *Transl. Psychiatr.* 6:e732. doi: 10.1038/tp.2015.221

Dwyer, D. B., Falkai, P., and Koutsouleris, N. (2018). Machine Learning Approaches for Clinical Psychology and Psychiatry. *Annu. Rev. Clin. Psychol.* 14, 91–118. doi: 10.1146/annurev-clinpsy-032816-045037

Ecker, C., Rocha-Rego, V., Johnston, P., Mourao-Miranda, J., Marquand, A., Daly, E. M., et al. (2010). Investigating the predictive value of whole-brain structural MR scans in autism: a pattern classification approach. *NeuroImage* 49, 44–56. doi: 10.1016/j.neuroimage.2009.08.024

Elsabbagh, M., Divan, G., Koh, Y.-J., Kim, Y. S., Kauchali, S., Marcín, C., et al. (2012). Global prevalence of autism and other pervasive developmental disorders. *Autism Res.* 5, 160–179. doi: 10.1002/aur.239

Falter, C. M., Braeutigam, S., Nathan, R., Carrington, S., and Bailey, A. J. (2013). Enhanced access to early visual processing of perceptual simultaneity in autism spectrum disorders. *J. Autism Dev. Disord.* 43, 1857–1866. doi: 10.1007/s10803-012-1735-1

Falter, C. M., Elliott, M. A., and Bailey, A. J. (2012). Enhanced visual temporal resolution in autism spectrum disorders. *PLoS ONE* 7:e32774. doi: 10.1371/journal.pone.0032774

Georgescu, A. L., Kuzmanovic, B., Roth, D., Bente, G., and Vogeley, K. (2014). The use of virtual characters to assess and train non-verbal communication in high-functioning autism. *Front. Human Neurosci.* 8:807. doi: 10.3389/fnhum.2014.00807

Georgescu, A. L., Kuzmanovic, B., Schilbach, L., Tepest, R., Kulbida, R., Bente, G., et al. (2013). Neural correlates of "social gaze" processing in high-functioning autism under systematic variation of gaze duration. *NeuroImage* 3, 340–351. doi: 10.1016/j.nicl.2013.08.014

Glover, E., Garberson, F., Abbas, H., and Wall, D. P. (2018). Machine learning approach for early detection of autism by combining questionnaire and home video screening. *J. Am. Med. Inform. Assoc.* 25, 1000–1007. doi: 10.1093/jamia/ocy039

Grammer, K., Honda, M., Juette, A., and Schmitt, A. (1999). Fuzziness of nonverbal courtship communication unblurred by motion energy detection. *J. Pers. Soc. Psychol.* 77:487. doi: 10.1037//0022-3514.77.3.487

Green, D., Charman, T., Pickles, A., Chandler, S., Loucas, T., Simonoff, E., et al. (2009). Impairment in movement skills of children with autistic spectrum disorders. *Dev. Med. Child Neurol.* 51, 311–316. doi: 10.1111/j.1469-8749.2008.03242.x

Isaksson, S., Salomäki, S., Tuominen, J., Arstila, V., Falter-Wagner, C. M., and Noreika, V. (2018). Is there a generalized timing impairment in Autism Spectrum Disorders across time scales and paradigms? *J. Psychiatr. Res.* 99, 111–121.

Jansiewicz, E. M., Goldberg, M. C., Newschaffer, C. J., Denckla, M. B., Landa, R., and Mostofsky, S. H. (2006). Motor signs distinguish children with high functioning autism and Asperger's syndrome from controls. *J. Autism Dev. Disord.* 36, 613–621. doi: 10.1007/s10803-006-0109-y

Kleinbub, J. R., and Ramseyer, F. (2019). *rMEA Synchrony in Motion Energy Analysis (MEA) Time-Series.* R package version 1.1.0. Available online at: https://CRAN.R-project.org/package=rMEA

Kupper, Z., Ramseyer, F., Hoffmann, H., and Tschacher, W. (2015). Nonverbal synchrony in social interactions of patients with schizophrenia indicates socio-communicative deficits. *PLoS ONE* 10:e0145882. doi: 10.1371/journal.pone.0145882

Li, B., Sharma, A., Meng, J., Purushwalkam, S., and Gowen, E. (2017). Applying machine learning to identify autistic adults using imitation: an exploratory study. *PLoS ONE* 12:e0182652. doi: 10.1371/journal.pone.0182652

Li, Z., Gavrilyuk, K., Gavves, E., Jain, M., and Snoek, C. G. M. (2018). Videolstm convolves, attends and flows for action recognition. *Comp. Vis. Image Underst.* 166, 41–50. doi: 10.1016/j.cviu.2017.10.011

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., et al. (2000). The autism diagnostic observation schedule—generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J. Autism Dev. Disord.* 30, 205–223. doi: 10.1023/A:1005592401947

Lord, C., Rutter, M., and Le Couteur, A. (1994). Autism diagnostic interview-revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J. Autism Develop. Disord.* 24, 659–685. doi: 10.1007/BF02172145

Maenner, M. J., Yeargin-Allsopp, M., Van Naarden Braun, K., Christensen, D. L., and Schieve, L. A. (2016). Development of a machine learning algorithm for the surveillance of autism spectrum disorder. *PLoS ONE* 11:e0168224. doi: 10.1371/journal.pone.0168224

Marín-Jiménez, M. J., Zisserman, A., Eichner, M., and Ferrari, V. (2014). Detecting people looking at each other in videos. *Int. J. Comput. Vis.* 106, 282–296. doi: 10.1007/s11263-013-0655-7

McIntosh, D. N., Reichmann-Decker, A., Winkielman, P., and Wilbarger, J. L. (2006). When the social mirror breaks: deficits in automatic, but not voluntary, mimicry of emotional facial expressions in autism. *Dev. Sci.* 9, 295–302. doi: 10.1111/j.1467-7687.2006.00492.x

Mehta, D., Sridhar, S., Sotnychenko, O., Rhodin, H., Shafiei, M., Seidel, H.-P., et al. (2017). Vnect: real-time 3d human pose estimation with a single rgb camera. *ACM Trans. Graph.* 36:44. doi: 10.1145/3072959.3073596

Menassa, D. A., Braeutigam, S., Bailey, A., and Falter-Wagner, C. M. (2018). Frontal evoked γ activity modulates behavioural performance in Autism Spectrum Disorders in a perceptual simultaneity task. *Neurosci. Lett.* 665, 86–91. doi: 10.1016/j.neulet.2017.11.045

Merin, N., Young, G. S., Ozonoff, S., and Rogers, S. J. (2007). Visual fixation patterns during reciprocal social interaction distinguish a subgroup of 6-month-old infants at-risk for autism from comparison infants. *J. Autism Dev. Disord.* 37, 108–121. doi: 10.1007/s10803-006-0342-4

Murphy, D. G. M., Beecham, J., Craig, M., and Ecker, C. (2011). Autism in adults. New biological findings and their translational implications to the cost of clinical services. *Brain Res.* 1380, 22–33. doi: 10.1016/j.brainres.2010.10.042

Nagaoka, C., and Komori, M. (2008). Body movement synchrony in psychotherapeutic counseling: A study using the video-based quantification method. *IEICE Trans. Inf. Syst.* 91, 1634–1640. doi: 10.1093/ietisy/e91-d.6.1634

Noel, J.-P., De Niear, M. A., Lazzara, N. S., and Wallace, M. T. (2017). Uncoupling between multisensory temporal function and nonverbal turn-taking in autism spectrum disorder. *IEEE Trans. Cognit. Develop. Syst.* 10, 973–982. doi: 10.1109/TCDS.2017.2778141

Parma, V., and de Marchena, A. B. (2016). Motor signatures in autism spectrum disorder: the importance of variability. *J. Neurophysiol.* 115, 1081–1084. doi: 10.1152/jn.00647.2015

Paxton, A., and Dale, R. (2013). Frame-differencing methods for measuring bodily synchrony in conversation. *Behav. Res. Methods* 45, 329–343. doi: 10.3758/s13428-012-0249-2

Perego, P., Forti, S., Crippa, A., Valli, A., and Reni, G. (2009). "Reach and throw movement analysis with support vector machines in early diagnosis of autism," in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (Minneapolis, MN), 2555–2558. doi: 10.1109/IEMBS.2009.5335096

Ramseyer, F., and Tschacher, W. (2011). Nonverbal synchrony in psychotherapy: coordinated body movement reflects relationship quality and outcome. *J. Consult. Clin. Psychol.* 79, 284–295. doi: 10.1037/a0023419

Ramseyer, F., and Tschacher, W. (2014). Nonverbal synchrony of head- and body-movement in psychotherapy: different signals have different associations with outcome. *Front. Psychol.* 5:979. doi: 10.3389/fpsyg.2014.00979

Redcay, E., Dodell-Feder, D., Mavros, P. L., Kleiner, M., Pearrow, M. J., Triantafyllou, C., et al. (2013). Atypical brain activation patterns during a face-to-face joint attention game in adults with autism spectrum disorder. *Hum. Brain Mapp.* 34, 2511–2523. doi: 10.1002/hbm.22086

Romero, V., Amaral, J., Fitzpatrick, P., Schmidt, R. C., Duncan, A. W., and Richardson, M. J. (2017). Can low-cost motion-tracking systems substitute a Polhemus system when researching social motor coordination in children? *Behav. Res. Methods* 49, 588–601. doi: 10.3758/s13428-016-0733-1

Romero, V., Fitzpatrick, P., Roulier, S., Duncan, A., Richardson, M. J., and Schmidt, R. C. (2018). Evidence of embodied social competence during conversation in high functioning children with autism spectrum disorder. *PLoS ONE* 13:e0193906. doi: 10.1371/journal.pone.0193906

Schmidt, R. C., Morr, S., Fitzpatrick, P., and Richardson, M. J. (2012). Measuring the dynamics of interactional synchrony. *J. Nonverbal. Behav.* 36, 263–279. doi: 10.1007/s10919-012-0138-5

Schmidt, R. C., Nie, L., Franco, A., and Richardson, M. J. (2014). Bodily synchronization underlying joke telling. *Front. Hum. Neurosci.* 8:633. doi: 10.3389/fnhum.2014.00633

Tariq, Q., Daniels, J., Schwartz, J. N., Washington, P., Kalantarian, H., and Wall, D. P. (2018). Mobile detection of autism through machine learning on home video: a development and prospective validation study. *PLoS Med.* 15:e1002705. doi: 10.1371/journal.pmed.1002705

Thabtah, F. (2018). Machine learning in autistic spectrum disorder behavioral research: a review and ways forward. *Inform. Health Soc. Care* 44, 278–297. doi: 10.1080/17538157.2017.1399132

Tome, D., Russell, C., and Agapito, L. (2017). "Lifting from the deep: convolutional 3d pose estimation from a single image," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Honolulu, HI), 2500–2509. doi: 10.1109/CVPR.2017.603

Tran, D., Wang, H., Torresani, L., Ray, J., LeCun, Y., and Paluri, M. (2018). "A closer look at spatiotemporal convolutions for action recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Salt Lake City, UT), 6450–6459. doi: 10.1109/CVPR.2018.00675

Varoquaux, G., Raamana, P. R., Engemann, D. A., Hoyos-Idrobo, A., Schwartz, Y., and Thirion, B. (2017). Assessing and tuning brain decoders: cross-validation, caveats, and guidelines. *NeuroImage* 145, 166–179. doi: 10.1016/j.neuroimage.2016.10.038

Vogel, D., Falter-Wagner, C. M., Schoofs, T., Krämer, K., Kupke, C., and Vogeley, K. (2019). Interrupted time experience in autism spectrum disorder: empirical evidence from content analysis. *J. Autism Dev. Disord.* 49, 22–33. doi: 10.1007/s10803-018-3771-y

Wall, D. P., Dally, R., Luyster, R., Jung, J.-Y., and DeLuca, T. F. (2012a). Use of artificial intelligence to shorten the behavioral diagnosis of autism. *PLoS ONE* 7:e43855. doi: 10.1371/journal.pone.0043855

Wall, D. P., Kosmicki, J., DeLuca, T. F., Harstad, E., and Fusaro, V. A. (2012b). Use of machine learning to shorten observation-based screening and diagnosis of autism. *Transl. Psychiatr.* 2:e100. doi: 10.1038/tp.2012.10

Wee, C.-Y., Wang, L., Shi, F., Yap, P.-T., and Shen, D. (2014). Diagnosis of autism spectrum disorders using regional and interregional morphological features. *Hum. Brain Mapp.* 35, 3414–3430. doi: 10.1002/hbm.22411