

Digital Phenotyping Autism: Investigating Objective Vocal and Movement Quantification for Characterizing Autism Severity

By: Emilin Mathew

[1] Background Research (06/09/2020-07/19/2020):

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder with deficits in socialization, communication, and repetitive/ritualistic behaviors (Faras, 2010). These social deficiencies may greatly impair daily functioning and persist throughout adulthood. Epidemiologically, the prevalence of diagnoses has been rapidly increasing. The CDC has reported that 1 in 54 American children were diagnosed in 2020 (Autism Speaks, ND). Rates have increased 119.4 % over the course of 10 years and 10% within the past 2 years. Unfortunately, such data does not even account for the large number of undiagnosed/misdiagnosed cases of patients.

A major cause of misdiagnoses is the lack of an objective detection metric for ASD. The current gold standard is the Autism Diagnostic Observation Schedule (ADOS) which is a 40-60-minute exam. There are four modules to account for different levels of development. For example, the module administered to toddlers is in play format whereas kids over 30 months are tested in conversational format. Results are divided into two behavioral domains: Social Affect (SA) and Restricted and Repetitive Behaviors (RRB). SA measures social activity such as eye contact, initiation of communication, response to name, joint attention, etc. RRB includes excessive movement such as hand flapping (Akshoomoff et al, 2006).

There are many flaws with this current method: 1) The ADOS may inaccurately assert ASD as the sole cause for patients' symptoms. Social communication and interaction related diseases get misdiagnosed with ASD such as learning disorders, depression, or Attention-deficit/hyperactivity disorder. To clarify, a symptom such as irregular facial expressions have a lot of underlying causes so it is difficult to pinpoint as ASD. This issue is augmented by the fact that several psychiatric disorders (autism, schizophrenia, bipolar disorder, and major depression) have genetic and phenotypic similarities (Cardno et al, 2014). 2) The result of the diagnosis is a total score that cannot represent the nuances and subtleties amongst ASD patients. All symptoms have varying severities that a simple number cannot encompass (Jabr, 2012). 3) Psychological evaluations are inevitably subjective because psychologists perceive patient activities differently. One might believe that the patient is regularly pointing while the other believes it's abnormal (Ehmke, 2020). 4) The ADOS was not created to account

for/detect temporal changes in symptoms. This is problematic due to the dynamic nature of the disease, especially in linguistic development (Frye, 2018).

Therefore, it is clear new criteria must be investigated for evaluating the severity of ASD symptoms. One potential predictor is vocalization patterns; ASD patients are known to have difficulty in verbal communication and social practices such as hand gestures and eye contact. Specifically, it is difficult to initiate and maintain conversations. Conversational turn counts (CTC), a measure for defining the number of conversations, is when either an adult or child initiates a discussion and the other responds with vocal responses within 5 seconds. These child-adult alterations can be measured by Language Environment Analysis (LENA). LENA provides adult and child word count, classifies child vocalizations, and indicates CTCs. It is groundbreaking in that it provides automated and objective data to understand linguistic development.

Another metric that can be used in ASD diagnosis not covered in the ADOS is distances patients maintain between themselves and others during social interaction. ASD patients typically engage in behavioral avoidance and are unresponsive to social interaction, which can be observed through expressive control of space (Gessaroli et al, 2013). Self generated distance is vital in sensorimotor functions such as action regulation and maintenance of gait (Honma et al, 2019). These abilities may be impaired due to cerebellar deficits which are associated with postural control, gait, sensory integration, and reduced connectivity in motor nodes (Kindregan et al, 2015). A novel automated coding technology to objectively record these data is motion-capture systems. Kinect sensors can be installed to the ceiling and note spatiotemporal data of ASD patients and adults' location during the examination (Cohen et al, 2014).

[2] Problem Statement (07/20/2020-08/1/2020):

This investigation analyzes the potential in using distances children with ASD maintain from adults during the examination and their conversational abilities as new criteria for evaluating the severity of ASD symptoms.

[3] Hypotheses (08/2/2020-08/22/2020):

If the average distance of children with ASD from the parent is measured during ADOS, then there will be a statistically significant correlation with Social Affect score.

If the average distance of children with ASD from the examiner is measured during ADOS, then there will be a statistically significant correlation with Social Affect score.

If conversational exchanges initiated by children with ASD are calculated, then a decrease of conversation initiations will have a statistically significant correlation with an increase in the Social Affect score and Total ADOS score.

[4] Experimental Setup (08/23/2020-09/5/2020):

[A] Materials

- Vocalization recordings of children diagnosed with ASD
- Movement Data of positions of child, examiners, and parents over time during ADOS
- ADOS-2 Raw Scores Obtained from licensed clinical psychologist
- R Studio
- Language Environment Analysis (LENA)
- Excel (Data Tool Pack)

[B] Procedures

Subject Sample Description: I only worked with de-identified data obtained from the University of Miami's psychology lab under the PI Dr. Daniel Messinger. No interactions with human subjects were made. Sample had 18 males and 7 females averaging 38 months in age who were diagnosed with ASD. 17 children were Latino, 7 were Black, and 1 was white.

Analyze Vocal Data:

1. Export LENA data to an Excel file.
2. Delete all columns except that of vocal type, segment duration of CTC, number of CTC, and the child ID number.
3. Move the entire vocal type column one row lower than the segment duration and CTC columns. This matches up the initiator of the turn count to the number of turn counts and duration.
4. Select all data values and click F5, empty data sets.
5. Delete the selected rows to eliminate extraneous data. These do not contain CTC information.
6. Count both CHN_CHF and CXN_CXF vocal IDs.
7. Note the duration of each turn count and number of CTC in each series.

Analyze Movement Data:

1. Save motion capture data as text files with child, examiner, and parent/guardian's X and Y coordinates.
2. Import data into R to calculate average distances between child and parent and child and examiner. Here's sample code:

```
d <- read.delim("/System/Volumes/Data/Users/emilinmathew/Desktop/Data for
2021/ADOS_Ceiling Tracking for HeatMaps/ Output_1024_exp.txt")

names(d)<- c("ParentX", "ParentY", "ChildX", "ChildY", "Ex1X", "Ex1Y")

PaCh<- sqrt((d$ParentX-d$ChildX)^2+(d$ParentY+d$ ChildY)^2)

mean(PaCh)
```

Design Heatmaps on R:

1. Save motion capture data as text files with child, examiner, and parent/guardian's X and Y coordinates.
2. Import data into R using the following sample code:

```
d <- read.table("/System/Volumes/Data/Users/emilinmathew/Desktop/Data for
2021/ADOS_Ceiling Tracking for HeatMaps/New ones to
do/ADOS1057_C_EXP_103019_output.txt", header=FALSE)

names(d)<- c("ParentX", "ParentY", "ChildX", "ChildY", "Ex1X", "Ex1Y")

require(MASS)

dens_child <- kde2d(d$ChildX, d$ChildY, h=75, n=50)

filled.contour(dens_child, xlim =range(0,1000), ylim = range(0,1000))
```

Conduct Regressions:

1. On the Data tab in excel, click Data Analysis.
2. Select regression, and the ranges for the sample. Check labels and include the heading of each column (If doing multiple regressions and there are multiple X variables, place the columns adjacent).
3. Conduct Stepwise regression: Delete variable with the highest p- value and rerun.

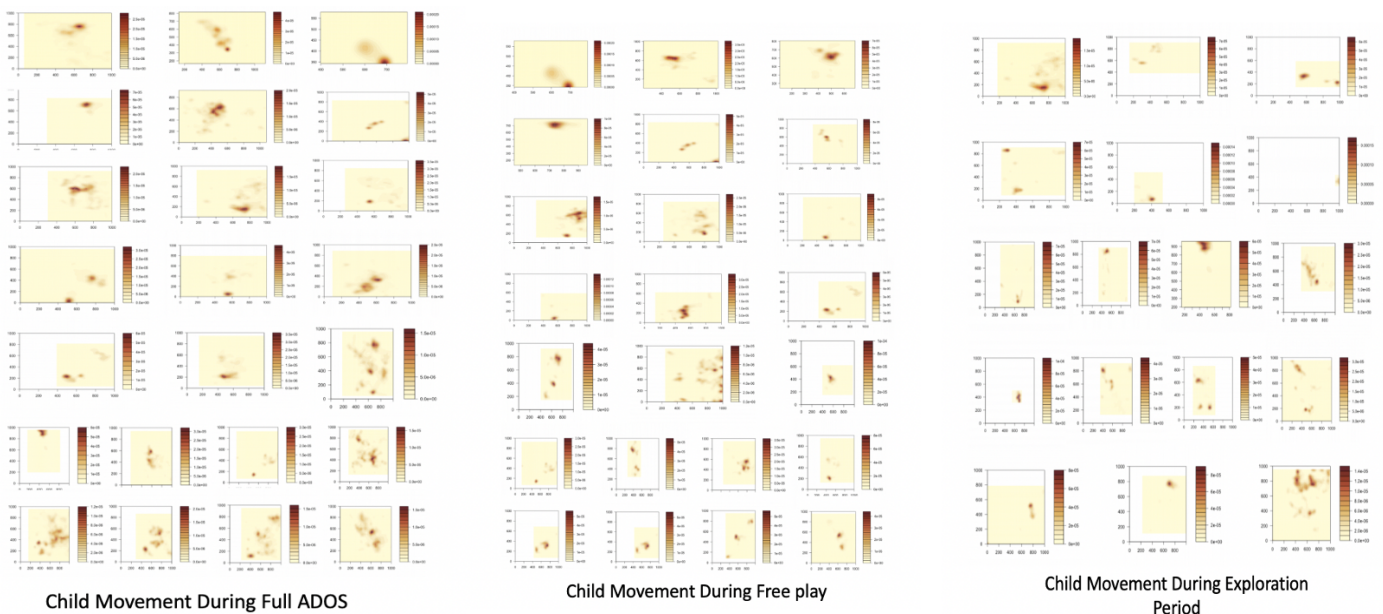
4. Repeat until overall significance F is <0.05.
5. With each model, check the adjusted R square to see if the model's becoming more fit.
6. Calculate the regression line: Dependent Variable = intercept + coefficient variable * respective variable.

[5] Variables (09/6/2020-09/19/2020):

- Independent Variables: Total number of Conversational Turn Counts, number of Conversational Exchanges initiated by the child, average turns in a Conversational Exchange, Duration of each Conversational Exchange, Distance from child to parent during ADOS, distance from child to examiner during ADOS,
- Dependent Variables: Total ADOS score, Social Affect score, Restricted and Repetitive Behavior Score
- Constants: Metric for evaluating ADOS score, audio extraction technique, machine learning method for movement
- Control is not required for this correlation study

[6] Data (09/20/2020-10/10/2020):

Heatmaps: Although there is some movement during the structured ADOS heatmap, the children largely remained located in one region for the majority of the exam. However, during the free play and exploration periods, the children had little to no movement and most were located in the periphery of the room. This is indicative of ASD behavioral abnormalities because the children are expected to move around and play.



[7] Results & Analysis (10/11/2020-12/6/2020):

Vocalization was quantified through number of child CE initiations, duration of CE, total number of CTCs, and average number of CTCs in one CE. Movement data was obtained during the ADOS clinical examination with computer vision analysis from ceiling mounted cameras.

Table 1: Univariate Regression Results

		Total ADOS Score	Social Affect Score	Restricted and Repetitive Behavior Score
Child Initiated Conversational Exchanges	Pearson Correlation	-0.356	* -0.442	0.024
	p value	0.081	* 0.027	0.910
Conversational Turn Counts	Pearson Correlation	-0.363	* -0.414	-0.073
	p value	0.075	*0.040	0.727
Distance from Child to Examiner	Pearson Correlation	0.091	0.129	-0.051
	p value	0.667	0.537	0.809
Distance from Child to Parent	Pearson Correlation	0.234	0.237	0.129
	p value	0.260	0.255	0.540

This is the Pearson's correlation table of the variables independently predicting the scores from the ADOS. The two significant models were the child initiated conversational exchanges and total conversational turn counts predicting the social affect score. Pearson's correlation of ± 1 means perfect relationship and p value is significant at 0.05.

Table 2: Predicting the Social Affect Score

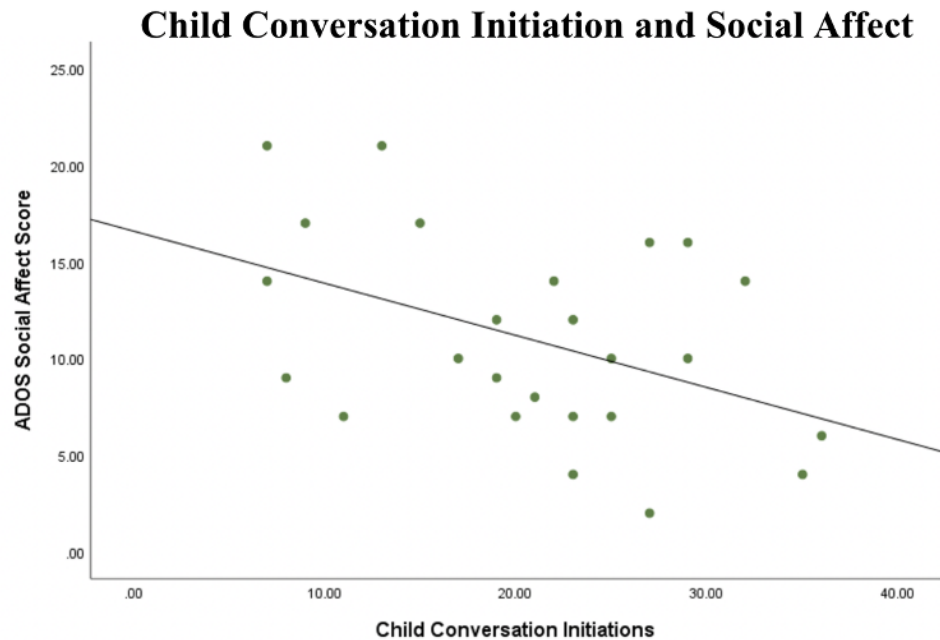
Multiple Regression Predicting the Social Affect Score			
Independent Variables	F- Stat	Adjusted R ²	P-value
Distance from child to parent, Distance from child to examiner, Number of CE initiated by Child, Number of CTC in series, Segment Duration of CE	1.763	0.137	0.169
Distance from child to parent, Number of CE initiated by Child, Number of CTC in series, Segment Duration of CE	2.261	0.173	0.099
Distance from child to parent, Number of CE initiated by Child, Number of CTC in series	3.087	0.261	*0.049
Distance from child to parent, Number of CE initiated by Child	3.524	0.174	*0.047
Number of CE initiated by Child	5.591	0.161	*0.027

This is the stepwise multiple regression where row 1 has all independent variables and the least correlated variable is removed one at a time. The only significant independent variable in all models with these variables was the number of CTCs initiated by Child. This signifies that the child to parent and child to examiner distance during the ADOS was not able to predict the social affect score neither independently nor in conjunction with the vocalization variables.

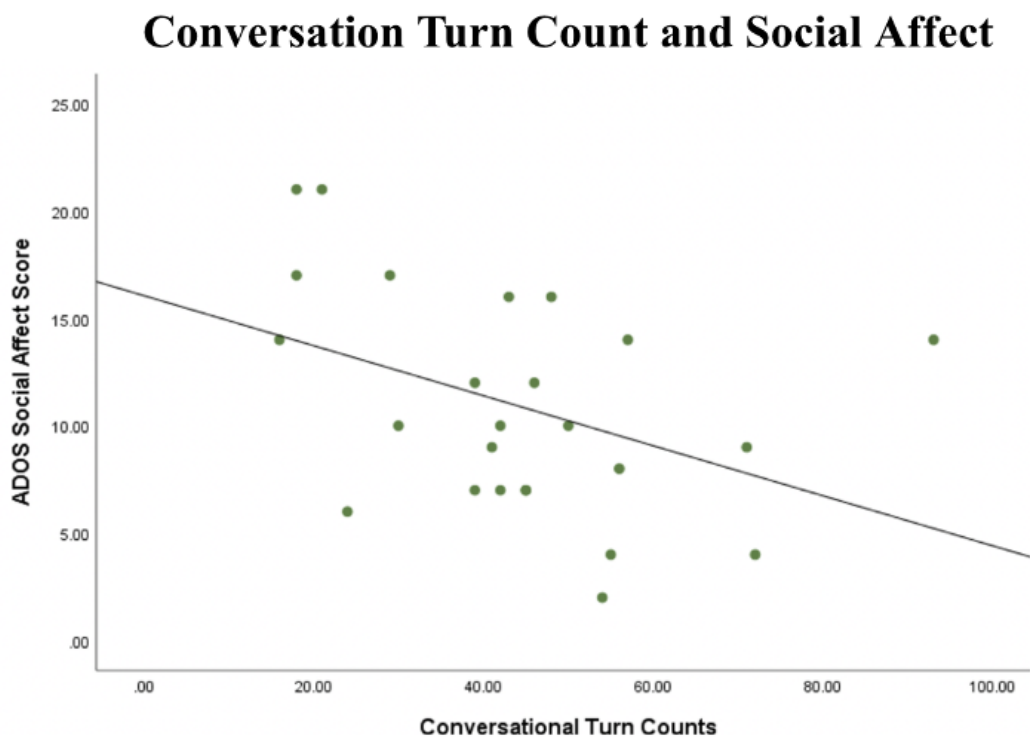
Table 3: Best Fit Model to Predict Social Affect Score

Best Fit Model to predict Social Affect Score: Adjusted R ² : 0.161				
Model	Unstandardized Coefficients	Standard Error	t	p value
Intercept	16.593	2.563	6.474	0.000
Number of Conversational Exchanges initiated by the Child	-0.2698	0.1141	-2.364	0.0269

This table displays the best fit model with the predictor variable being the number of conversational exchanges initiated by the child predicting the Social Affect score. The linear regression, $y = -0.2698x + 16.593$, shows that a decrease in CE initiated by the child correlates to an increase in Social Affect score.



This scatterplot displays the correlation between conversation initiations by the child and the social affect score. There is a strong negative correlation ($p=0.0269$).



This scatterplot displays the correlation between total turn counts in a conversation and the social affect score. There is a strong negative correlation ($p=0.040$).

[8] Analysis, Conclusions, and Applications (12/7/2020-12/19/2020):

Autism Spectrum Disorder (ASD) affects one in 58 US children (Brian et al, 2019). A major issue in this field is the lack of objectivity in diagnostic tests which results in incomplete diagnoses and interferes with prescribing relevant medications. The current leading diagnostic evaluation or the Autism Diagnostic Observation Schedule (ADOS) is a semi-structured exam that continues to have numerous downfalls. It cannot distinguish nuances in symptoms shared with other psychiatric conditions and is subjective by nature of requiring a clinician to analyze symptoms.

This study digitally phenotyped two novel criteria for diagnosing autism: 1) distances the child maintained between parent and examiner during the full ADOS and free play and 2) linguistic development: number of conversational exchanges initiated by the child, duration of each conversation, average turns in a conversation exchange, and total turns in a conversation. I hypothesized that the two distance variables would have correlations with the social affect domain of the ADOS score and that the number of children initiated conversational exchanges would be correlated with the social affect and total ADOS score.

Distances were obtained from ceiling motion tracking technology which measured the movements of child, parent, and examiner during the ADOS examination. This collects automated and precise data down to the hundredth of a second. The mean distances of individuals in the room were calculated via coding in RStudio to hold the immense volume of data. For the linguistic data, adult and child vocalization was collected during the full ADOS diagnostic assessment. It was analyzed by a software called Language Environment Analysis (LENA) which records and classifies vocalizations, initiations of conversation, duration of speaking, etc. I mined this data to isolate the conversations initiated by the child and associated information.

I conducted Pearson's correlation and multiple regression analyses to test the correlation of movement and vocal variables to four dependent variables. Overall, no models were able to predict the total ADOS score and RRB score. The strongest correlations and statistical significance was when testing with the SA score. The most fit model was able to account for 16.1% of total variance of the SA score. Therefore, the variable in this model, the number of conversations initiated by Child, serves as useful tools in predicting ASD. Other significant findings include that the child is the main agent that determines the ADOS score: The CTC initiation variable is entirely dependent on children initiating conversations and child-adult movement is also determined by whether the child wishes to move or not.

This investigation provides necessary insight into using movement and linguistic measurements as predictors for autism. These criteria are significant because they provide a new framework for conducting autism diagnoses. Unlike previous subjective diagnostics which require costly and time-intensive training for clinicians, this computational method creates new avenues for objectively assessing autism symptoms using automated technology.

The benefits in using LENA and motion capture technology is three-fold. 1) They can be utilized in natural settings such as a classroom which allows patients to act freely in supportive, non-critical environments as both language and distance can be measured in daily settings. This can assist in early diagnosis because it can be integrated seamlessly into spaces inhabited by young children (such as daycare) and be analyzed for at risk populations. 2) They're objective and automated, deriving precise measurements with extreme temporal accuracy. This eliminates the subjectivity introduced by the ADOS. 3) These results can be compared across time and also reflect gradations in symptoms much freely. Conversely, the ADOS assigns a number as a diagnostic score without accounting for such fluctuation.

[9] Reflections (12/20/2020-12/26/2020):

The study can be expanded by separating the vocalization data collected during the full ADOS into subsections of the exam i.e., have vocal data recorded from only during the free play and exploration period. This would allow researchers to compare the distance and vocal variables more precisely and avoid covariance in data. Additionally, I used ADOS-2 raw scores instead of the Calibrated Severity Scores (CSS). Evaluating the CSS could provide more insight as literature suggests it is more standardized amongst samples. Finally, the sample size of 25 should be expanded to include more patients. This size prevented clear trends in individual item level scores because of variation which is why conclusive results were not obtained about individual symptoms. When I attempted to run logistic regressions, the predictor variable gave low p-values for many items. However, scatterplots showed that this correlation was not following a trend and where due to chance from the small sample size. Having more patients can give a clearer understanding of how movement and vocal standards can predict individual symptom severity.

Future research should expand on this novel method of digital phenotyping where movement tracking and vocal analysis techniques predict behavioral symptom severity. Although I obtained recordings/movement data from a 60-minute test, the programs used can collect data from a longer duration of time to have a more holistic understanding of patient movement and language. For example, these systems can be integrated at daycare centers for at risk populations or purchased by families with familial history of autism. Research can also be conducted on

using machine learning technology to identify and track the movement of multiple children in a room to have results faster rather than conducting individual tests.

This study used children approximating 38 months in age with speech development at nonverbal- single word phrase. Future research can analyze children at older ages with phrase speech to analyze if the linguistic variables can still predict the social affect score. Furthermore, other variables can also be incorporated to produce more fit multiple regression models so researchers can jointly test variables.

[10] Bibliography:

Akshoomoff, N., Corsello, C., & Schmidt, H. (2006). The Role of the Autism Diagnostic

Observation Schedule in the Assessment of Autism Spectrum Disorders in School and Community Settings. Retrieved January 01, 2021, from

<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1868476/>

Autism Statistics and Facts. (n.d.). Retrieved January 01, 2021, from

<http://www.autismspeaks.org/autism-statistics>

Cardno, A., & Owen, M. (2014, May). Genetic relationships between schizophrenia,

bipolar disorder, and schizoaffective disorder. Retrieved from

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3984527/>

Ehmke, R. (2020, February 28). Autism Evaluation and Diagnosis. Retrieved January 01,

2021, from <https://childmind.org/article/what-should-evaluation-autism-look-like/>

Faras, H., Al Ateeqi, N., & Tidmarsh, L. (2010). Autism spectrum disorders. Retrieved

January 01, 2021, from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2931781/>

Frye, R. (2018, August). Social Skills Deficits in Autism Spectrum Disorder: Potential Biological Origins and Progress in Developing Therapeutic Agents. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6105175/>

Gessaroli, E., Santelli, E., Di Pellegrino, G., & Frassinetti, F. (2013, September 23). Personal space regulation in childhood autism spectrum disorders. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3781155/>

Jabr, F. (2012, January 30). By the Numbers: Autism Is Not a Math Problem. Retrieved from <https://www.scientificamerican.com/article/autism-math-problem/>

S. (2020). International Rules For Pre-College Science Research Guidelines For Science And Engineering Fairs 2020–2021. Retrieved from <https://sspcdn.blob.core.windows.net/files/Documents/SEP/ISEF/2021/Rules/Book.pdf>