Using Neural Networks to Show the Lack of Cultural Relativity in Autism Screeners

COLUMBIA UNIVERSITY | QMSS
QUANTITATIVE METHODS IN THE SOCIAL SCIENCES
MASTERS PROGRAM IN THE INSTITUTE FOR SOCIAL AND ECONOMIC RESEARCH AND POLICY

By Zachary Zivalich

Phone: 661-400-7890 | Email: zlz2001@columbia.edu

COLUMBIA UNIVERSITY | QMSS
QUANTITATIVE METHODS IN THE SOCIAL SCIENCES
MASTERS PROGRAM IN THE INSTITUTE FOR SOCIAL AND ECONOMIC RESEARCH AND POLICY

Aim

The purpose of this study is to address the potential bias within autism screeners as a result of lack of cultural relativity. I hypothesize that common autistic traits manifest differently across gender and sociocultural groups, which is a reality not accounted for by screeners at this moment.

Abstract

Autism Spectrum Disorder is a developmental disability that has a prevalence rate between .9 to 1.7% of the U.S. population with similar rates in developed countries. Epidemiological surveys show that the disorder is not associated with racial or genetic factors, yet this is in direct contrast with observed prevalence rates - where white children are significantly more likely to be diagnosed than Black or Hispanic children. Despite this, little has been done to explain how this imbalance came to be. I address the skewed prevalence rates by looking at methodological bias in Autism Screeners. Autism screeners are questionnaires acting as a first step in diagnosis – truncating individuals seeking a diagnosis by scoring autistic traits. I use an open source dataset, filled out by app takers from all over the world (n=2496), to address the issue of bias within the questions of autism screeners. These screeners are designed to gauge autistic traits. Yet, trait scores varied as a product of race and gender, with white men scoring higher overall. I use a logistic regression to show the likelihood of diagnosis, which is higher for white app users. Lastly, I build a neural network to predict whether a test taker is a white male or not, providing evidence that the screener's questions have been designed around more available white data. Autistic traits are the de facto symptom of diagnosis. It is possible that the dearth of nonwhite children in early autism studies lead to said traits being based on predominantly white male characteristics.

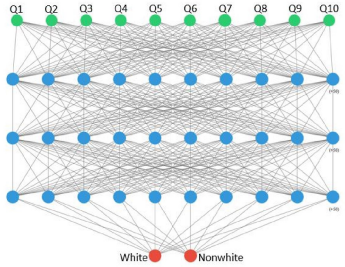


Fig. 1: Neural Network

Data and Methods

Two datasets were used for this study- one filled out by the guardians of toddlers and another by adults. These datasets are provided by Dr. Fadi Thabtah who collected it via smartphone app (ASD Tests) which is easily downloaded from the Android and Apple app store. The participants fill out varied autism screeners, which are based on the Q-CHAT and Autism Quotient surveys. Participants receive a score which determines whether they are recommended to pursue professional diagnosis. The dataset is prone to selection bias, as anyone can download the app and fill out the survey, meaning statistically it cannot be applied to a wider population. However, the prevalence rates for the samples (~70% and ~30%) are well above the national average of 1.5%. Therefore, it is likely that this dataset is comprised of participants who are much more likely to be autistic. I run a logistic regression to see which ethnicities and genders are more or less likely to reach diagnosis. Then, I compare the average scores of white and nonwhite test takers through a two-sample t-test. Finally, I run a neural network to see if a deep learning algorithm can pick up on a score pattern only white male participants display.

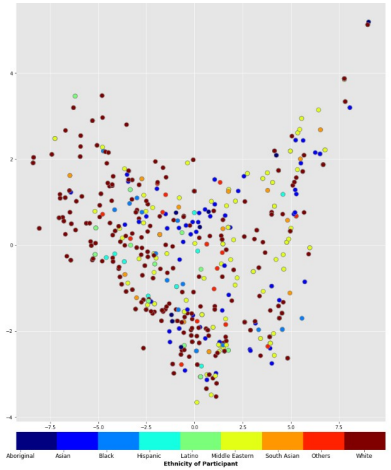


Fig. 2: Isometric Map of Adult Score Patterns

Results

Across both age groups, white toddlers and adults scored higher on each question then nonwhites. Scores were compared through a two-sample two-way T test. Every T-value was statistically significant but one question (A8), meaning there was below a 5% probability the difference in means were found by chance. However , this did not mean every ethnicity groups scored less than white test takers.

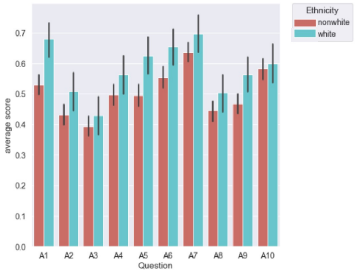


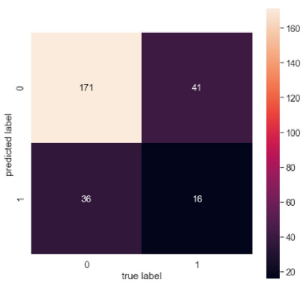
Fig. 3: Average scores for Adult White vs Nonwhite

Black and Latino Male test takers were more likely to receive a recommendation on both questionnaires then White males. However, these three groups were very likely to be diagnosed when compared with other groups. For example, Black and White Male toddlers were 2.05 and .85 log-odds more likely to be diagnosed then Middle Eastern Male toddlers.

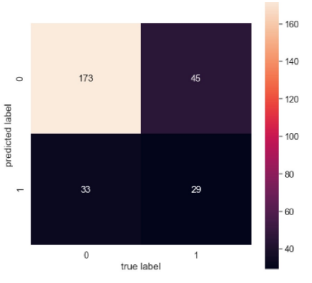
	coef	std err	z	P>|z|
Intercept	0.4082	0.117	3.480	0.001
male:C(Ethnicity)[Hispanic]	0.8239	0.445	1.850	0.064
male:C(Ethnicity)[Latino]	1.4636	0.769	1.904	0.057
male:C(Ethnicity)[white European]	0.8048	0.193	4.165	0.000
male:C(Ethnicity)[asian]	0.4185	0.185	2.260	0.024
male:C(Ethnicity)[black]	2.0195	0.614	3.291	0.001
male:C(Ethnicity)[middle eastern]	-0.0335	0.228	-0.147	0.883
male:C(Ethnicity)[mixed]	-0.4082	1.007	-0.405	0.685
male:C(Ethnicity)[south asian]	0.1026	0.347	0.296	0.767

Table 1: Multivariate Logit Regression of Male Toddlers

Fig. 4 and 5: Decision Matrix of Toddler and Adult Neural Network (respectively)



	accuracy	f1_score	precision	recall
0	0.708333	0.554904	0.553394	0.557148



	accuracy	f1_score	precision	recall
0	0.721429	0.621254	0.615849	0.63066

A Neural network was then designed to classify test takers into nonwhite male and white male categories across both questionnaires based only on the participants score patterns. After some tuning, It was able to predict which class each participant belonged to with an accuracy of ~70% and an f1_score of ~60%. Just by training on how a participants answers the 10 questions, the model could guess the ethnicity to a moderate degree of accuracy. However, it more so excelled at classifying who was not a white male then who was.

Conclusion

•White scored higher than nonwhite test takers on average across every question on both screeners.

•Latino and Black male test takers were more likely to receive a diagnosis then white males. However, all three were much more likely to receive a diagnosis then Nonwhite, Black, or Hispanic males.

•The likelihood of diagnostic recommendation varying by ethnicity and gender in differing degree.

• The neural network was able to predict white male test takers to ~70% accuracy with ~60% f1_score for both datasets, signaling at least one measurable pattern that is dependent on race.

Discussion

As the awareness of autism continues to rise, psychiatric practitioners need to better understand how the behavioral symptomology of autism manifests itself as a product of socio-cultural factors. It is important that autistic children get access to the resources necessary to excel at an early age so that they do not develop the depression and suicidal tendencies shown to appear in higher rates in autistics diagnosed after childhood. Despite the selection bias in the sample, this sample's specific qualities do well in simulating the barriers of diagnosis for those who believe they are their children may be autistic. When the dataset is understood to be composed of participants seeking a diagnosis because of self realized symptomology, it is easy to see that certain groups of people struggle to receive the diagnosis they need. Going forward, follow up studies can select a true random sample to see the sociocultural differences in the wider population of autistic adults and toddlers to verify these findings.

Acknowledgements

Data comes from Dr. Fadi Thabhtah of creator of ASD tests and co-creator of Autism AI. His website can be found here https://fadifayez.com/