Spring 5-21-2022

Using Machine Learning to Analyze Children’s Drawings as Indicators of Mental Well-Being

Nicholas R. Lawrie
University of San Francisco, nlawrie12@gmail.com

Using Machine Learning to Analyze Children's Drawings as Indicators of Mental Well-Being

JEL Classifications: I15

Nick Lawrie

Department of Economics
University of San Francisco
2130 Fulton St.
San Francisco, CA 94117

Thesis Submission for the Masters of Science Degree
in Applied Economics

email: *nrlawrie@usfca.edu*

May 20, 2022

Abstract: Human figure drawings are a well-studied diagnostic tool for emotional distress in children. Cleft lip/palate is one of the most common birth defects in the world, and has been shown to negatively impact emotional well-being in childhood which can have negative economic consequences in adulthood. Utilizing a dataset of human figure drawings from children in India and survey data on mental health, this paper will assess the impact of corrective surgery on mental health outcomes, as well as assess the validity of the drawing emotional indicators themselves. The results indicate that while the emotional indicators may not be valid in predicting emotional distress in this sample, there is a positive relationship between corrective surgery and mental health outcomes.

The author wishes to thank Bruce Wydick for his help and guidance in completing this research, and Shreeja Guha for her assistance in translating the drawings which were used for this analysis.

1. Introduction

Human figure drawings (HFDs) have long been studied as a screening tool for emotional distress in young children. Researchers have been interested in drawings because of their ability to reveal emotional distress in children who may not have the vocabulary or self-awareness to adequately communicate their emotional condition. Machine learning, while having been studied for decades in the computer sciences, is not widely used in the social sciences due to its inherent focus on accurately predicting out-of-sample data rather than explaining causal relationships between endogenous and exogenous variables. For researchers in fields such as psychology and economics, the latter focus on causal relationships has more practical research applications. Using HFDs by Syrian refugee children, Baird et al. (2020) conducted a study which found that LASSO regression, a common machine learning method, was successful in selecting statistically significant emotional indicators when compared to the results of traditional subset regression analysis. Wydick et al. (2022) have collected a dataset of HFDs by children in India which includes corresponding survey data on mental well-being and family demographic information for the purpose of conducting an impact evaluation on cleft palate/lip (CLP) corrective surgery. This paper will proceed as a cross-validation study on the validity of the emotional indicators described in the psychology literature for this sample, using indices based on the drawing indicators, indices based on self-reported survey data, and cross-validated LASSO regressions. The paper will be organized as follows. Section 2 will provide an overview of related literature in the fields of economics, psychology, and machine learning. Section 3 will provide detail on the dataset and research methods that will be used. Section 4 will report results, which will immediately be followed by discussion in Section 5. Concluding remarks will be provided in Section 6.

2. Literature Review

Mental well-being is a growing topic of interest in economic research and falls into two broad categories: examining if income has a causal effect on an individual's happiness and determining the economic impact of mental illness. Regarding the former, research has shown that the relationship of income on happiness is ambiguous and that there may be no causal relationship between the two (Graham, 2005; Jaunky et al., 2020; Pitukhina et al., 2021). However, the latter poses this relationship in the reverse: can mental illness have a negative impact on income or economic development? Major depressive disorder alone is one of the most common illnesses in the world and is estimated to reduce productivity by at least $31 billion in the United States (de Quidt & Haushofer, 2016). De Quidt and Haushofer (2016) conducted a theoretical study on poverty traps which showed that depressed individuals develop negative outlook on returns to labor and become more withdrawn and less productive as a result, which further exacerbates the negative effects of return on labor. Patel (2013) has similarly argued that depression in adolescents is a global health priority due to its prevalence and its long-term effects which carry over into adulthood. With respect to development economics, research has shown that health problems in children can have long-term negative consequences on schooling and human capital accumulation (Currie & Stabile, 2006). Currie & Vogl (2013) noted in their research that these effects may be exacerbated in developing countries where children do not have access to proper medical care. Individuals born with CLP, one of the most common birth defects in the world, have widely reported lower scores on quality of life and more frequently experience depressive symptoms later in life (Ali et al., 2021; Gowda et al., 2013; Nagappan et al., 2019; Pinquart & Shen, 2011; Wehby & Cassell, 2010).

Finding an accurate and cost-effective method to screen for emotional trauma in children is of great importance because children may be reluctant to expressing trauma or lack the self-awareness to verbally communicate complex or abstract ideas. HFDs have been well-studied in the field of psychology, and there have been numerous studies which provide evidence that a child's drawing can provide meaningful insight into their current emotional state (Baird et al., 2020; Farokhi & Hashemi, 2011; Glewwe et al., 2018). Koppitz (1968) in her research showed that the presence of two or more emotional indicators on a child's human figure drawing is associated with higher levels of emotional distress. An emotional indicator is defined as an indicator which is "not primarily related to a child's age and maturation but reflects his anxieties, concerns, and attitudes" (Koppitz, 1968, p. 35). These emotional indicators include, but are not limited to, the relative size of the human figure in the drawing, asymmetry in the figure, and the position of the figure's hands and limbs. As a screening tool, human figure drawing are helpful in clinical settings because children, regardless of how reticent they may be in expressing emotional distress, will often readily produce a drawing upon request. Koppitz (1968) noted in her research the importance of using a specific and homogeneous prompt for the child to create their human figure drawing, as having the child produce a drawing of any subject of their choosing invalidates their drawing of clinical significance.

As mentioned, this paper will build upon a study conducted by Baird (2020), which found that drawings by Syrian refugee children could serve as an accurate screening method for emotional trauma, and that the LASSO regression method could accurately select statistically significant indicators of emotional distress. Machine learning currently is not widely used in the social sciences, though its use has expanded in recent years. Athey (2019), who works in the field of economics, argues that machine learning in the computer sciences has expanded under the goal of creating predictive algorithms rather than explaining relational causality, which is the primary topic of interest in the social sciences. Nevertheless, Yarkoni (2017), who works in the field of psychology, expresses caution in the use of regression methods because the mechanical purpose of a regression is to explain the variation within a given sample and is therefore prone to over-fitting. This leads to the development of models which do not retain their accuracy when applied to data outside the sample. However, this does not imply that regression analysis is irrelevant. As previously mentioned, OLS regression is a superior analytical tool for explaining causality, and machine learning itself faces the same issues of over-fitting. Similar to Baird's 2020 study, this paper proposes a specific case study for the application of machine learning in analyzing children's drawings and determine if machine learning is a relevant method in analyzing these drawings as indicators of mental-well-being.

Machine learning as a field has been in development since the 1960's. The end goal of machine learning is to develop a model which automatically improves itself when introduced to new data. However, machine learning as an area of research is still relatively new, and an autonomous self-learning algorithm has yet to be developed (Jordan & Mitchell, 2015). The LASSO regression method used by Baird et al. in their study was first introduced by Tibshirani (1996). The LASSO regression creates an algorithm which selects coefficients which have a significant impact on the dependent variable and reduces insignificant variables gradually to zero by introducing a penalty for the number of parameters (Rajaratnam, 2016; Wang, 2007). In this way, the LASSO regression optimizes the number of variables by selecting the fewest variables that have the greatest explanatory power on the endogenous variable. The LASSO regression is an ideal tool for economists as it utilizes the predictive ability of machine learning while still creating an explanatory model. In terms of prediction, random forests have been introduced as an accurate and efficient method for accurately categorizing class variables (Breiman, 2001). Other machine learning methods such as ridge regression, neural networks, and k-nearest neighbors have also been used in a variety of cross-validation studies to test for accuracy, though the results of these studies have been mixed. For example, it has been shown that machine learning can out-perform linear regression analysis in predicting bid prices of public works projects and return-on-investment in the film industry (Kim & Jung, 2019; Kim et al., 2020). However, Chen (2021) testing 13 different machine learning and linear regression models on predicting mortgage delinquency and found that none of the models were successful in predicting out-of-sample data, with all accuracy scores falling below 40%.

In the intersection of machine learning and HFDs, a study conducted by Eitz (2012) tested the accuracy of machine learning algorithms in categorizing human sketches against the accuracy of manual human identification. Using the same set of 20,000 human-drawn sketches spread across 250 categories, humans were able to accurately identify the images in these sketches 73% of the time, compared to a 56% accuracy rate for machine learning algorithms. However, Monica et al. (2019) tested the k-nearest neighbor classifier method for categorizing drawings and found that this method achieved a 76.8% accuracy rate. Refaat & Atiya (2009) found that a Support Vector Machined Kernel method could correctly categorize drawings with a 92.8% accuracy rate. These results have two critical implications. First, thorough cross-validation is necessary to determine if machine learning is a reliable tool for analyzing HFDs, especially considering the end goal of using these algorithms to identify emotional indicators and screen for emotional distress. Second, this implies that analyzing drawings with machine learning cannot be a substitute for clinical expertise and should not be the sole screening method used. Be that as it may, the obvious benefit to studying machine learning with respect to this paper is that an algorithm in theory could process a set of drawings and select those which show signs of emotional distress in a fraction of the time that it would take manually. One day, this algorithm may even accomplish this task with greater accuracy than humans ourselves. The question of whether it is ethical to solely employ machine learning to identify signs of emotional distress should not be ignored, though this debate falls outside the scope of this paper. However, the two critical implications of machine learning identification mentioned earlier should prompt the reader to exercise caution when implementing machine learning in psychological diagnostic settings in its current stage of development.

3. Research Method

3.1. Dataset

The dataset which will be used in this study is a set of 453 drawings from children in India. Summary statistics are reported in Table 1. The survey sample consisted of 219 households where at least one child with CLP is present. 223 households surveyed did not have a child with CLP which served as a counterfactual sample. The siblings of the children with CLP were also surveyed and asked to complete drawings. Each child was provided with the same prompt for their drawing: "Draw a picture of your immediate family (your parents and your siblings together). In your drawing, please indicate where your parents are, and where you are." It is important that each child was provided with the same prompt so that we can conclude that every emotional indicator that is included in the drawing is a result of the child's own volition and not because of differences in instructions. Each child was provided paper and colored pencils to complete their drawing. The colors provided include, but are not limited to, blue, black, grey, red, orange, green, and purple. Due to limitations resources in the field during project implementation, there may have been differences in the colors that were available to each child.

Each child was also asked to complete a survey to assess their emotional well-being, and demographic information for each child and their family also collected. For this study, we will focus on similar parameters to those utilized in Baird et al.'s (2020) study for our model, which will include age, gender, birth order (1=first born child, 2=second born child, etc.), required surgeries (defined as the number of surgeries required to restore life outcomes to near normalcy), and number of surgeries received. Our dependent variables will be indices for depression and anxiety, which will be constructed using either drawing indicators or survey data. The survey data for anxiety and depression are coded with discrete values from 1 to 5, with 5 representing severe depression and anxiety, and 1 representing no depression or anxiety.

3.2. Emotional Indicators

Because the main goal of this paper is to assess the validity of the emotional indicators themselves as well as the accuracy of LASSO regressions, we will use well-established emotional indicators as our dependent variables in our model. These emotional indicators will mimic those used in Baird et al.'s (2020) study and as originally developed by Koppitz (1968). These variables will include shading of face or body, missing nose or mouth, frowning or crying, use of dark colors, use of only a single color, poor figural integration (disconnected head or limbs from the body), smiling, use of light cheery colors, sketchy broken lines, tiny figure (smaller than 2.5 cm), faint lines, tiny head, and lack of detail (Baird et al., 2020). Each of these indicators has been shown to be significant in predicting higher levels of emotional distress in children. Each indicator will be represented by a dummy variable, which will be coded as 1 if the indicator is present in the drawing. We will also introduce two new indicators: smaller in relative size and separated from parents. Smaller in relative size will be defines as a child drawing themselves smaller in size compared to their siblings. A child who draws themselves as the same size of their siblings but drew themselves and all siblings as smaller than their parents will not be considered as having drawn themselves relatively smaller. Separated from parents will be defined as a child who did not draw themselves physically next to either parent in their drawing. Our prediction is that a score of 1 for both of these indicators will be associated with higher levels of anxiety and depression.

3.3. Regression Model Construction

The regression analysis will be conducted with respect to two dependent variable types: indices based on drawing indicators and indices based on survey data. The regression analyses will further be conducted in pairs with anxiety and depression as the dependent variables of interest respectively. For indices based on drawing indicators, the main question of interest is whether having cleft or receiving corrective surgery has a significant effect on emotional well-being using the drawings themselves as the method of measurement. We will analyze three index types: principal components, Kling indices, and Anderson indices. Each index will be inversely signed so that a negative coefficient indicates a poorer score on depression or anxiety, and each will be constructed using the indicators associated with each condition of interest. For anxiety, these include shading of the face, missing nose or mouth, frowning or crying, use of dark colors, use of only a single color, poor figural integration, use of sketchy or broken lines, smiling inversely, and use of light colors inversely. For depression, these include shading of the face, missing nose or mouth, frowning or crying, use of dark colors, use of a single color, poor figural integration, tiny figure, faint lines, and tiny head. Each index type for both depression and anxiety will be separately regressed with required surgeries and number of surgeries received as the main independent variables of interest, as well as standard controls which include age, gender, and birth order. The number of required surgeries and number of surgeries received are used in lieu of a single dummy variable indicating any cleft condition. This allows the regression analysis to investigate the effects of the severity of cleft as well as the effects of corrective surgery. Results will be presented with and without household fixed effects. P-values will also be converted to q-values using the Benjamini-Hochberg correction to test for false rejections.

The indices based on survey data will be used to test whether the drawing indicators as presented in the psychology literature are valid in this case. Each participant was asked to self-report their level of depression and anxiety respectively on a scale of 1 to 5, with 5 representing the most extreme cases of mental distress. These two indices are then inversely signed as before for consistency in reporting. Both indices are then regressed against the set of indicators which fall under the purview of each respective condition. These sets of indicators are identical to the ones used to construct the indices in the previous paragraph. In addition, the survey indices will be regressed against all indicators which were coded in the dataset. In addition, the regressions above will be replicated using the survey indices as the dependent variables with and without household fixed effects to test for robustness.

3.4. LASSO Model Construction

The final section of analysis will be based on cross-validated LASSO using the survey indices as the dependent variable and all drawing indicators as the independent variables. The LASSO results will serve as an additional cross-validation method on the drawing indicators themselves to assess whether these indicators as denoted by the psychology literature remain valid. The LASSO method finds an optimal weighting for each coefficient so that only those which hold predictive power remain in the model. The key difference when running this model is that all coded emotional indicators will be included for both depression and anxiety. If the indicators behave as we expect, LASSO will eliminate those which are not associated with each mental health condition. Subset regressions will then be used to construct a model for each which only includes those emotional indicators which LASSO has selected. While this holds little statistical power in terms of significance, this can provide some insight into which indicators appear to hold weight for predicting mental health.

4. Results

4.1. Principle Component Analysis

Results for the principal component analysis are reported in Table 2. For the models with anxiety as the dependent variable, with and without household fixed effects, we find that the coefficient for required surgeries is estimated to be negative and the coefficient for number of surgeries is estimated to be positive. While the direction of the coefficients is expected, these coefficients are not statistically significant. The coefficient for age in the anxiety model without household fixed effects is statistically significant at the 1% level and is negative, indicating that older children are more likely to suffer from anxiety. This coefficient remains statistically significant at the 1% level after the Benjamini-Hochberg correction is applied. The coefficient for birth order is also statistically significant at the 1% level and is estimated to be negative, indicating that younger siblings are more likely to suffer from anxiety. However, introducing household fixed effects in the anxiety model renders all coefficient estimates statistically insignificant.

The model for depression without household fixed effects does not reveal statistical significance for required surgeries or number of surgeries. Interestingly, the signs of the coefficients for required surgeries and number of surgeries are in the opposite directions than we would expect, indicating that children who require more surgeries are less likely to experience depression and children who receive more corrective surgery are more likely to experience depression. The coefficients for age, gender, and birth order are statistically significant, indicating that older children are less likely to experience depression, males are more likely to experience depression, and younger siblings are more likely to experience depression. When household fixed effects are included, the coefficients for gender and birth order remain statistically significant at the 5% level with their signs in the same direction as before.

4.2. Kling Index Analysis

Results for the Kling index analysis are reported in Table 3. For the anxiety model, neither coefficient for required surgeries or number of surgeries received are statistically significant, though the signs of the coefficient estimates are in the expected direction. The coefficients for age and birth order are statistically significant at the 1% level and have the same signs as in the principal component analysis. These indicate that older children and younger siblings are more likely to experience anxiety. Both these coefficients remain statistically significant at the 1% level after calculating their q-values. When including household fixed effects, age and birth order are no longer statistically significant, though gender is shown to be significant at the 5% level, indicating that males are more likely to experience anxiety. However, this coefficient is not significant according to its corrected q-value. Introducing household fixed effects reverses the sign for number of surgeries in the unexpected direction.

The model for depression without fixed effects does not show statistical significance for required surgeries or number of surgeries. However, unlike the principal component model, the signs for these coefficients are in the expected directions. Gender is significant at the 5% level and birth order is significant at the 1% level. The signs for both coefficients are the same as in the principal component analysis, indicating that males and younger siblings are more likely to experience depression. After calculating the corrected q-values, birth order remains statistically significant at the 1% level, though gender is only significant at the 5% level. The inclusion of household fixed effects does not change the sign for gender or birth order, though the significance of birth order decreases to the 5% level. However, the corrected q-values for these coefficients do not indicate that they are statistically significant.

4.3. Anderson Index Analysis

As in the Kling index analysis, the Anderson index analysis reveals little about the relationships between cleft, depression, and anxiety. Results for the Anderson index analysis are reported in Table 4. For the anxiety model without fixed effects, the coefficients for required surgeries and number of surgeries are not statistically significant, though the signs are in the expected direction. The coefficient for birth order is significant at the 1% level, the coefficient for age is significant at the 5% level, and the coefficient for gender is significant at the 10% level. The signs for these coefficients are negative as in the Kling index analysis. The corrected q-values indicate that birth order and age remain statistically significant at the 5% level. Introducing household fixed effects renders all coefficient estimates statistically insignificant except for gender, which remains significant at the 10% level after calculating its corrected q-value. In addition, the sign for number of surgeries received becomes negative in divergence with expectation.

The signs of the coefficients for required surgeries and number of surgeries in the depression model are both negative, though as in the anxiety models, they are not statistically significant. Birth order is shown to be significant at the 1% level from its corrected q-value, with the same sign as in the anxiety model. Including household fixed effects, gender and birth order are significant at the 5% level with the same signs as the anxiety models. However, neither coefficient is statistically significant after converting their p-values to q-values.

4.4. Survey Index Comparison

Given the evidence from the indices based on drawing indicators is inconclusive in determining a relationship between cleft, anxiety, and depression, the next natural question is whether such a relationship exists at all. As previously discussed, the survey data includes self-reported indices for both anxiety and depression which can be used to cross-validate the results from the drawings-based indices. Results from these regressions are reported in Table 5. In the anxiety model, the coefficients for required surgeries and number of surgeries are both statistically significant at the 1% level, and the signs for both are in the expected directions. The coefficient for gender is also significant at the 5% level with a positive sign as also reported in the principal component analysis. After calculating corrected q-values, required surgeries remains significant at the 1% level, and number of surgeries remains significant at the 5% level. Introducing household fixed effects does not change the coefficient signs for required surgeries or number of surgeries. However, after calculated corrected q-values, required surgeries remains statistically significant at the 1% level, while number of surgeries is significant at the 10% level.

In the depression model, required surgeries and number of surgeries are reported as statistically significant at the 1% by corrected q-values. The signs for the coefficient estimates are also in the expected direction. Gender is significant at the 10% level with the same sign as the anxiety model, though it is not significant after calculating its corrected q-value. Including household fixed effects does not change the signs of the coefficients for required surgeries and number of surgeries. Required surgeries remains statistically significant at the 1% level after calculating corrected q-values, though the significance of number of surgeries received decreases to the 5% level. Gender is no longer statistically significant when household fixed effects are included.

4.5. Drawing Indicator Analysis

The next two subsections report the results for the cross-validation procedures of the drawing indicators themselves. Results for the drawing indicator analyses based on OLS regressions are reported in Table 5. The anxiety model with only associated indicators as dictated by the psychology literature included is reported first. The results show statistical significance for smiling at the 1% level, and sketchy broken lines and single color at the 5% level. After calculating corrected q-values, smiling is significant at the 5% level, and sketchy broken lines is significant at the 10% level, with single color no longer significant. The signs for single color, sketchy broken lines, and smiling are in the expected directions. Use of a single color and sketchy broken lines in a drawing indicate a higher likelihood of anxiety, while the participant drawing themselves with a smile indicates a lower probability of experiencing anxiety.

When the anxiety index is regressed against all the drawing indicators which were coded in the dataset, smiling and tiney figure are statistically significant at the 1% level according to p-values. However, while the sign for smiling is in the direction we would expect, the sign for tiny figure is in the opposite direction from the expectation. The psychology literature indicates that a child who draws themselves very small is more likely to be experiencing high levels of anxiety, though the regression reports results to the contrary. After calculating corrected q-values, smiling is significant at the 5% level and tiny figure is significant at the 10% level. Single color, sketchy broken lines, and smaller in relative size are significant at the 5% level according to p-values, and their signs are in the expected direction. However, after calculating corrected q-values, none of these three coefficients is statistically significant.

The depression model which includes only its associated indicators reports tiny figure as statistically significant at the 1% level, though the coefficient for tiny figure is again the opposite of the expectation as in the anxiety model. After calculating q-values, tiny figure remains statistically significant at the 1% level.

When regressed against all emotional indicators, smiling and tiny figure are statistically significant at the 1% level according to p-values. The signs for these coefficients remain unchanged from the anxiety model with all indicators. After calculating q-values, tiny figure remains significant at the 1% level, though smiling is significant at the 5% level. Separated from parents is shown to be statistically significant at the 10% level according to p-value, and the positive sign indicates that a child drawing themselves apart from their parents is associated with lower probability of experiencing depression. However, this coefficient is not statistically significant when its corrected q-value is considered.

4.6. LASSO Cross-Validation

The results of the cross-validation LASSO procedures and post-estimation OLS regressions are reported in Table 6. For anxiety, LASSO selects single color, smiling, sketchy broken lines, tiny figure, and smaller in relative size as the most important indicators in predicting anxiety. When these parameters are run in post-estimation OLS, we find that the signs for single color, smiling, sketchy broken lines, and smaller in relative size are in the expected direction. However, the coefficient for tiny figure is again signed reversely from the expectation.

For depression, LASSO selects smiling, sketchy broken lines, tiny figure, separated from parents, and different drawing from prompt as the best predictors of experiencing depression. Running these parameters in post-estimation OLS, we find the coefficients for smiling, sketchy broken lines, and different drawing from prompt have the expected sign. However, the coefficient for tiny figure again has the opposite sign according to expectation. The coefficient for separated from parents is positive and is consistent with the preceding analyses.

5. Discussion

The results provide mixed evidence on the significance of the emotional indicators used in screening for anxiety and depression in this sample, as the results of the principal component analysis, Kling index analysis, and Anderson index analyses diverge from the survey data comparison. The survey data analysis shown in Table 4 indicates that children who require more surgeries are more likely to suffer from anxiety and depression, and receiving corrective surgery is expected to reduce the likelihood of anxiety and depression. Each coefficient was statistically significant except for number of surgeries in the depression model without household fixed effects when using q-values. However, when using the indices based on the drawings themselves, we find no coefficient estimates for required surgeries or number of surgeries are statistically significant, and some coefficient estimates are signed in directions that we would not expect.

The source of this inconsistency could perhaps be that the drawing indicators as shown in the psychology literature are inapplicable in this cultural setting, or perhaps from the sample method used. Regarding the former, the drawing indicator analysis shows some indicators as statistically significant in predicting levels of anxiety and depression, though there are inconsistencies with the psychology literature. For the anxiety OLS model, the best predictors for anxiety appear to be single color, sketchy broken lines, smiling, light cheery colors, tiny figure, and smaller in relative size, though the sign of the coefficient estimate for tiny figure is opposite the expected direction. LASSO selects single color, poor figural integration, smiling, light cheery colors, sketchy broken lines, tiny figure, smaller in relative size, and separated from parents as the best predictors for anxiety. For the depression OLS model, the best predictors appear to be single color, smiling, tiny figure, and separated from parents. However, the sign for tiny figure is again in the opposite expected direction. LASSO selects smiling, sketchy broken lines, tiny figure, and separated from parents as the best predictors for depression.

While HFDs have been used successfully in the field (Baird et al., 2022; Glewwe et al., 2018; Sokic et al., 2019), several studies show that the significance of certain emotional indicators can vary across cultures. For example, La Voy et al. (2001), in a study comparing HFDs of Japanese and American children, found that Japanese children drew fewer smiles, included more details in their drawings, and drew themselves larger on average. Ozer (2009), examining the validity of Koppitz's (1968) indicators for a sample in Turkey, found that there were significant differences in which indicators are deemed exceptional, i.e., occurring in less than 15% of all drawings, such as the depiction of two-dimensional feet. Ozer's (2009) study also notes that there were significant differences between socioeconomic groups within the sample. Children who attended private school tended to include more details in their drawings than children who attended public school, which could be due to more one-on-one attention from teachers and more parental involvement in education. Using a sample of children in Curacao, Vedder et al. (2000) tested the validity of three sets of HFD cognitive indicators which had been validated in the continental US, the Netherlands, and Puerto Rico. The results of the study showed that using indicators validated in Puerto Rico, which is culturally similar to Curacao as compared to the US or the Netherlands, resulted in a normal distribution of cognitive ability in the sample. However, using the indicators validated in the Netherlands would cause 50% of the sample to be diagnosed as intellectually impaired.

Specifically regarding India, Pala et al. (2016) tested the validity of the Child Drawing Hospital Scale, which is a screening method that uses HFDs to measures pain and anxiety in hospitalized children. This study used these indicators to analyze a set of drawings from a group of dental extraction patients in India. The results found that while there was a positive association with inclusion of the drawing indicators and levels of pain and anxiety experienced by the patients, the associations were not statistically significant.

Another source of inconsistency could be from the dataset itself. The data was collected by household and half of the households in the sample had one child who was born with CLP. While data was also collected from siblings without CLP as well as other households who did not have a child with CLP , the targeted nature of the sampling procedure could introduce bias into these results by having children with CLP over-represented. Previous studies on the effects of cleft palate on mental well-being have noted the difficulty in obtaining adequate data for both treatment and control groups (Wehby, 2010). Therefore, the interpretations and results of this study may not be statistically valid when applied to the broader population, and perhaps a better sampling method would need to be implemented to obtain more consistent results. One potential solution to this would be to survey classmates of children with cleft to obtain household observations which do not have a child with cleft.

6. Conclusion

This paper serves primarily as a cross-validation study of the standard emotional indicators introduced by the psychology literature and the LASSO cross-validation method. As mentioned earlier, the benefit of using machine learning in this context is that a computer algorithm could quickly analyze drawings and select those which show signs of emotional distress. The methods used in this paper, such as manually evaluating each drawing and assigning a score of 1 to each emotional indicator variable if the indicator is present, is still a time-consuming process. Therefore, an area of further research could be using machine learning to analyze the drawings themselves to determine if they can create an algorithm which can select drawings which show signs of emotional distress.

Nevertheless, HFDs are a fast and cost-effective way to screen for emotional distress in children. While the literature indicates that some emotional indicators which are exceptional for one population may not be for another, HFDs have been successfully used in international settings to assess program impacts, effects of exposure to violence, and effects of reintegration in host countries for refugees. The research indicates that HFDs would not be suitable for clinical practice or as a diagnostic tool in settings where emotional indicators have not been verified. However, many indicators still appear to be significant across cultures, and HFDs can still be an important tool for measuring the impact of social programs on mental health. As in Glewwe et al.'s (2018) study on the impact of child sponsorship on mental health, HFDs can be used to assess which programs are the most effective, which can in turn lead to better mental health outcomes for children in developing countries. This in turn can help mitigate the negative economic effects of poor mental health which carry over into adulthood. One area for further research which this paper has uncovered is validating and establishing emotional indicators which can be used diagnostically in India.

References

Ali, M. A., Abass, S. K., & Nasir, E. F. (2021). A comparative assessment of oral health-related quality of life of children born with orofacial clefts in sudan and their caregivers'. BMC Oral Health, 21(1) doi:10.1186/s12903-021-01514-2

Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. Annual Review of Economics, 11, 685-725. doi:10.1146/annurev-economics-080217-053433

Baird, Sarah; Muz, Jennifer; Panlilio, Raphael; Smith, Stephanie; Wydick, Bruce (2020): Identifying Psychological Trauma among Syrian Refugee Children for Early Intervention: Analyzing Digitized Drawings using Machine Learning. CEGA Working Paper Series No. WPS-125. Center for Effective Global Action. University of California, Berkeley. Text. https://doi.org/10.26085/C3WC7S

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32. doi:10.1023/A:1010933404324

Chen, S., Guo, Z., & Zhao, X. (2021). Predicting mortgage early delinquency with machine learning methods. European Journal of Operational Research, 290(1), 358-372. doi:10.1016/j.ejor.2020.07.058

Currie, J., & Stabile, M. (2006). Child mental health and human capital accumulation: The case of ADHD. Journal of Health Economics, 25(6), 1094-1118. doi:10.1016/j.jhealeco.2006.03.001

Currie, J., & Vogl, T. (2013). Early-life health and adult circumstance in developing countries. Annual Review of Economics, 5, 1-36. doi:10.1146/annurev-economics-081412-103704

De Quidt, J., & Haushofer, J. (2016). Depression for economists. National Bureau of Economic Research, working paper series. doi:10.3386/w22973

Eitz, M., Hays, J., & Alexa, M. (2012). How do humans sketch objects? ACM Transactions on Graphics, 31(4) doi:10.1145/2185520.2185540

Farokhi, M., & Hashemi, M. (2011). The analysis of children's drawings: Social, emotional, physical, and psychological aspects. Paper presented at the Procedia - Social and Behavioral Sciences, , 30 2219-2224. doi:10.1016/j.sbspro.2011.10.433 Retrieved from www.scopus.com

Glewwe, P., Ross, P. H., & Wydick, B. (2018). Developing hope among impoverished children: Using child self-portraits to measure poverty program impacts. Journal of Human Resources, 53(2), 330-355. doi:10.3368/jhr.53.2.0816-8112R1

Gowda, M. R., Pai, N. B., & Vella, S. -. C. (2013). A pilot study of mental health and quality-of-life of caregivers of children with cleft lip/palate in india. Indian Journal of Psychiatry, 55(2), 167-169. doi:10.4103/0019-5545.111456

Graham, Carol. Insights on Development from the Economics of Happiness, The World Bank Research Observer, Volume 20, Issue 2, Fall 2005, Pages 201-231, https://doi.org/10.1093/wbro/lki010

Jaunky, V. C., J. Jeetoo, and S. Rampersad. "Happiness and Consumption in Mauritius: An Exploratory Study of Socio-Economic Dimensions, Basic Needs, Luxuries and Personality Traits." Journal of Happiness Studies, vol. 21, no. 7, 2020, pp. 2377-2403. SCOPUS, www.scopus.com, doi:10.1007/s10902-019-00178-8.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260. doi:10.1126/science.aaa8415

Kim, J. -., & Jung, H. (2019). Predicting bid prices by using machine learning methods. Applied Economics, 51(19), 2011-2018. doi:10.1080/00036846.2018.1537477

Kim, J.-M., Xia, L., Kim, I., Lee, S., & Lee, K.-H. (2020). Finding Nemo: Predicting Movie Performances by Machine Learning Methods. Journal of Risk and Financial Management, 13(5), 93. MDPI AG. Retrieved from http://dx.doi.org/10.3390/jrfm13050093

Koppitz, E. M. (1968). Psychological evaluation of children's human figure drawings. N.Y. and London: Grune and Stratton.

La Voy, S. K., Pedersen, W. C., Reitz, J. M., Brauch, A. A., Luxenberg, T. M., & Nofsinger, C. C. (2001). Children's drawings. School Psychology International, 22(1), 53–63. https://doi.org/10.1177/0143034301221005

Monica, E.D., Davu, P., Caroline, C.P., & Jagannath, D.J. (2019). Doodle recognition using machine learning for hearing and speech-impaired people. Paper presented at the - 2019 2nd International Conference on Signal Processing and Communication (ICSPC), 109-112. doi:10.1109/ICSPC46172.2019.8976846

Nagappan, N., Madhanmohan, R., Gopinathan, N., Stephen, S., Pillai, D., & Tirupati, N. (2019). Oral health-related quality of life and dental caries status in children with orofacial cleft: An indian outlook. Journal of Pharmacy and Bioallied Sciences, 11(6), S169-S174. doi:10.4103/JPBS.JPBS_285_18

Ozer, S. (2009). Turkish children's human figure drawings: Can we borrow norms? Educational Psychology, 29(6), 701–712. https://doi.org/10.1080/01443410903210387

Pala, S. P., Nuvvula, S., & Kamatham, R. (2016). Expression of pain and distress in children during dental extractions through drawings as a projective measure: A clinical study. World journal of clinical pediatrics, 5(1), 102–111. https://doi.org/10.5409/wjcp.v5.i1.102

Patel, V. (2013). Why adolescent depression is a global health priority and what we should do about it. Journal of Adolescent Health, 52(5), 511-512. doi:10.1016/j.jadohealth.2013.03.003

Pinquart, M., & Shen, Y. (2011). Depressive symptoms in children and adolescents with chronic physical illness: An updated meta-analysis. Journal of Pediatric Psychology, 36(4), 375-384. doi:10.1093/jpepsy/jsq104

Pitukhina, M., O. Tolstoguzov, and A. Ilyin. Live, Work Or Leave? Arctic Youth Survey Results, vol. 678, 2021. SCOPUS, www.scopus.com, doi:10.1088/1755-1315/678/1/012020.

Rajaratnam, B., Roberts, S., Sparks, D., & Dalal, O. (2016). Lasso regression: Estimation and shrinkage via the limit of gibbs sampling. Journal of the Royal Statistical Society.Series B: Statistical Methodology, 78(1), 153-174. doi:10.1111/rssb.12106

Refaat, K. S., & Atiya, A. F. (2009). Hand-drawn shape recognition using the SVM'ed kernel doi:10.1007/978-3-642-04277-5_28 Retrieved from www.scopus.com

Rosselli, M., & Ardila, A. (2003). The impact of culture and education on non-verbal neuropsychological measurements: A critical review. Brain and Cognition, 52(3), 326–333. https://doi.org/10.1016/s0278-2626(03)00170-2

Skybo, T., Ryan-Wenger, N., & Su, Y.-hwa. (2007). Human figure drawings as a measure of children's emotional status: Critical Review for Practice. Journal of Pediatric Nursing, 22(1), 15–28. https://doi.org/10.1016/j.pedn.2006.05.006

Sokić, J., Đurović, D., & Biro, M. (2019). CHILDREN`S DRAWINGS AS A TRIAGE TOOL FOR THE ASSESSMENT OF NEGATIVE EMOTIONALITY IN REFUGEE CHILDREN. Primenjena Psihologija, 12(3), 245–262. https://doi.org/10.19090/pp.2019.3.245-262

Tibshirani, R. (1996), Regression Shrinkage and Selection Via the Lasso. Journal of the Royal Statistical Society: Series B (Methodological), 58: 267-288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.

Vedder, Paul; van de Vijfeijken, Katrien, & Kook, Hetty (2000). Borrowing Norms for the Human Figure Drawing Test: How to validate this practice?, Scandinavian Journal of Educational Research, 44:4, 393-404, DOI: 10.1080/713696680

Wang, H., Li, G., & Jiang, G. (2007). Robust regression shrinkage and consistent variable selection through the LAD-lasso. Journal of Business and Economic Statistics, 25(3), 347-355. doi:10.1198/073500106000000251

Wehby, G. L., & Cassell, C. H. (2010). The impact of orofacial clefts on quality of life and healthcare use and costs. Oral Diseases, 16(1), 3-10. doi:10.1111/j.1601-0825.2009.01588.x

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. Perspectives on Psychological Science, 12(6), 1100-1122. doi:10.1177/1745691617693393

Appendix

Table 1: Summary Statistics

	count	mean	sd	min	max
ShadingFace	453	.0154525	.1234804	0	1
MissingNoseMouth	453	.2406181	.4279314	0	1
FrowningCrying	453	.0750552	.2637717	0	1
DarkColors	453	.4988962	.5005516	0	1
SingleColor	453	.6732892	.469529	0	1
PoorFiguralIntegration	453	.4128035	.4928824	0	1
Smiling	453	.5739514	.4950477	0	1
LightCheeryColors	453	.218543	.4137147	0	1
SketchyBrokenLines	453	.2295806	.4210281	0	1
TinyFigure	453	.3134658	.4644148	0	1
FaintLines	453	.1942605	.3960677	0	1
TinyHead	453	.1677704	.3740754	0	1
LackofDetail	453	.5452539	.4984984	0	1
SmallerinRelativeSize	453	.0971302	.2964624	0	1
SeparatedFromParents	453	.4768212	.5000147	0	1
age	442	14.0362	3.335177	7	25
male	442	.5090498	.504995	0	2
b_order	442	1.952489	.9476098	1	6
depressed_survey	442	2.158371	1.445075	1	5
nerv_anx_survey	442	2.020362	1.384082	1	5

Table 2: Principal Component Analysis

	(1) Anxiety	(2) Anxiety with HH FE	(3) Depression	(4) Depression with HH FE
Required Surgeries	-0.0842	-0.0571	0.0190	0.00545
	(0.0515)	(0.0944)	(0.0479)	(0.0775)
Number of Surgeries	0.131	0.0293	-0.0228	-0.0863
	(0.111)	(0.185)	(0.103)	(0.152)
Age	-0.0814***	-0.0816	0.0373*	-0.0407
	(0.0208)	(0.0702)	(0.0193)	(0.0576)
Gender (1 if male)	0.0243	-0.206	-0.368***	-0.442**
	(0.134)	(0.212)	(0.125)	(0.174)
Birth Order	-0.243***	-0.249	-0.226***	-0.435**
	(0.0733)	(0.244)	(0.0682)	(0.200)
Intercept	1.670***	1.814	0.102	1.688
	(0.358)	(1.426)	(0.333)	(1.170)
N	442	442	442	442
R-sq	0.060	0.017	0.060	0.091
adj. R-sq	0.049	-1.659	0.050	-1.459

Standard errors in parentheses
="* p<0.10 ** p<0.05 *** p<0.01"

Table 2 q-values:

	(1) p-values	(1) q-values	(2) p-values	(2) q-values	(3) p-values	(3) q-values	(4) p-values	(4) q-values
Required Surgeries	0.1026	0.3078	0.5463	0.8743	0.6916	0.8251	0.9440	0.9440
Number of Surgeries	0.2372	0.4743	0.8743	0.8743	0.8251	0.8251	0.5701	0.9440
Age	0.0001	0.0005	0.2466	0.8743	0.0542	0.2170	0.4809	0.9440
Gender	0.8562	0.8562	0.3335	0.8743	0.0033	0.0166	0.0121	0.0725
Birth Order	0.0010	0.0039	0.3090	0.8743	0.0010	0.0058	0.0309	0.1547
Intercept	0.0000	0.0000	0.2052	0.8743	0.7603	0.8251	0.1509	0.6038

Table 3: Kling Index Analysis

	(1) Anxiety	(2) Anxiety with HH FE	(3) Depression	(4) Depression with HH FE
Required Surgeries	-0.0574	-0.0220	-0.0178	-0.0108
	(0.0353)	(0.0608)	(0.0357)	(0.0577)
Number of Surgeries	0.0447	-0.0147	0.00651	-0.0461
	(0.0760)	(0.119)	(0.0768)	(0.113)
Age	-0.0451***	-0.0380	0.00188	-0.0312
	(0.0143)	(0.0452)	(0.0144)	(0.0429)
Gender (1 if male)	-0.117	-0.290**	-0.217**	-0.309**
	(0.0920)	(0.137)	(0.0930)	(0.130)
Birth Order	-0.179***	-0.181	-0.179***	-0.343**
	(0.0503)	(0.157)	(0.0508)	(0.149)
Intercept	1.112***	1.087	0.463*	1.309
	(0.246)	(0.918)	(0.248)	(0.871)
N	442	442	442	442
R-sq	0.057	0.038	0.042	0.093
adj. R-sq	0.046	-1.602	0.031	-1.454

Standard errors in parentheses
="* p<0.10 ** p<0.05 *** p<0.01"

Table 3 q-values:

	(1) p-values	(1) q-values	(2) p-values	(2) q-values	(3) p-values	(3) q-values	(4) p-values	(4) q-values
Required Surgeries	0.1050	0.3149	0.7176	0.9020	0.6190	0.9325	0.8514	0.8514
Number of Surgeries	0.5565	0.5565	0.9020	0.9020	0.9325	0.9325	0.6836	0.8514
Age	0.0017	0.0067	0.4019	0.9020	0.8962	0.9325	0.4684	0.8514
Gender	0.2043	0.4086	0.0350	0.2100	0.0198	0.0991	0.0184	0.1101
Birth Order	0.0004	0.0021	0.2515	0.9020	0.0005	0.0027	0.0225	0.1126
Intercept	0.0000	0.0000	0.2383	0.9020	0.0631	0.2523	0.1350	0.5400

Table 4: Anderson Index Analysis

	(1) Anxiety	(2) Anxiety with HH FE	(3) Depression	(4) Depression with HH FE
Required Surgeries	-0.0466	-0.0100	-0.0220	-0.00886
	(0.0357)	(0.0596)	(0.0359)	(0.0584)
Number of Surgeries	0.0100	-0.0259	-0.000894	-0.0354
	(0.0767)	(0.117)	(0.0773)	(0.114)
Age	-0.0364**	-0.0329	-0.00435	-0.0176
	(0.0144)	(0.0443)	(0.0145)	(0.0434)
Gender (1 if male)	-0.176*	-0.346**	-0.183*	-0.286**
	(0.0929)	(0.134)	(0.0936)	(0.131)
Birth Order	-0.142***	-0.178	-0.162***	-0.299**
	(0.0508)	(0.154)	(0.0511)	(0.151)
Intercept	0.943***	1.022	0.507**	1.014
	(0.248)	(0.900)	(0.250)	(0.882)
N	442	442	442	442
R-sq	0.045	0.051	0.033	0.082
adj. R-sq	0.034	-1.569	0.022	-1.484

Standard errors in parentheses
="* p<0.10 ** p<0.05 *** p<0.01"

Table 4 q-values:

	(1) p-values	(1) q-values	(2) p-values	(2) q-values	(3) p-values	(3) q-values	(4) p-values	(4) q-values
Required Surgeries	0.1920	0.3841	0.8669	0.8669	0.5413	0.9908	0.8796	0.8796
Number of Surgeries	0.8962	0.8962	0.8248	0.8669	0.9908	0.9908	0.7570	0.8796
Age	0.0118	0.0471	0.4585	0.8669	0.7642	0.9908	0.6856	0.8796
Gender	0.0592	0.1775	0.0105	0.0633	0.0515	0.2061	0.0304	0.1826
Birth Order	0.0053	0.0263	0.2483	0.8669	0.0017	0.0100	0.0488	0.2440
Intercept	0.0002	0.0010	0.2576	0.8669	0.0432	0.2061	0.2518	0.8796

Table 5: Survey Index Analysis

	(1) Anxiety	(2) Anxiety with HH FE	(3) Depression	(4) Depression with HH FE
Required Surgeries	-0.256***	-0.246***	-0.325***	-0.251***
	(0.0481)	(0.0627)	(0.0497)	(0.0647)
Number of Surgeries	0.311***	0.238*	0.390***	0.322**
	(0.103)	(0.123)	(0.107)	(0.127)
Age	-0.0226	0.0279	0.00317	0.0188
	(0.0194)	(0.0466)	(0.0200)	(0.0481)
Gender (1 if male)	0.264**	0.0225	0.250*	-0.174
	(0.125)	(0.141)	(0.129)	(0.145)
Birth Order	0.0931	0.247	0.0538	0.255
	(0.0685)	(0.162)	(0.0707)	(0.167)
Intercept	-1.842***	-2.707***	-2.207***	-2.665***
	(0.335)	(0.947)	(0.346)	(0.976)
N	442	442	442	442
R-sq	0.092	0.165	0.111	0.144
adj. R-sq	0.082	-1.259	0.101	-1.317

Standard errors in parentheses
="* p<0.10 ** p<0.05 *** p<0.01"

Table 5 q-values:

	(1) p-values	(1) q-values	(2) p-values	(2) q-values	(3) p-values	(3) q-values	(4) p-values	(4) q-values
Required Surgeries	0.0000	0.0000	0.0001	0.0008	0.0000	0.0000	0.0002	0.0009
Number of Surgeries	0.0028	0.0114	0.0542	0.2169	0.0003	0.0012	0.0120	0.0479
Age	0.2454	0.2454	0.5507	0.8733	0.8743	0.8743	0.6967	0.6967
Gender	0.0354	0.1061	0.8733	0.8733	0.0536	0.1607	0.2315	0.4630
Birth Order	0.1748	0.2454	0.1287	0.3860	0.4476	0.8743	0.1279	0.3838
Intercept	0.0000	0.0000	0.0048	0.0241	0.0000	0.0000	0.0070	0.0352

Table 6: Drawing Indicator Analysis

	(1) Anxiety	(2) Anxiety	(3) Depression	(4) Depression
ShadingFace	-0.155	-0.196	-0.383	-0.313
	(0.529)	(0.533)	(0.555)	(0.557)
MissingNoseMouth	0.0705	0.0688	-0.0330	0.123
	(0.166)	(0.177)	(0.178)	(0.185)
FrowningCrying	0.268	0.216	-0.122	0.222
	(0.269)	(0.269)	(0.264)	(0.281)
DarkColors	-0.103	-0.0841	0.0733	0.124
	(0.157)	(0.157)	(0.152)	(0.164)
SingleColor	-0.390**	-0.375**	-0.151	-0.305
	(0.175)	(0.177)	(0.168)	(0.185)
PoorFiguralIntegration	0.217	0.258	0.0105	0.0303
	(0.141)	(0.170)	(0.150)	(0.177)
SketchyBrokenLines	-0.400**	-0.352**		-0.258
	(0.158)	(0.159)		(0.166)
Smiling	0.474***	0.469***		0.465***
	(0.149)	(0.149)		(0.155)
LightCheeryColors	-0.321	-0.332		-0.184
	(0.205)	(0.206)		(0.215)
TinyFigure		0.431***	0.605***	0.573***
		(0.158)	(0.159)	(0.165)
FaintLines		0.0560	-0.178	-0.151
		(0.167)	(0.176)	(0.174)
TinyHead		0.0233	0.0229	0.130
		(0.192)	(0.200)	(0.200)
LackofDetail		-0.258		-0.0405
		(0.172)		(0.179)
SmallerinRelativeSize		-0.466**		-0.337
		(0.232)		(0.242)
SeparatedFromParents		0.0842		0.253*
		(0.132)		(0.138)
Intercept	-1.942***	-1.988***	-2.234***	-2.478***
	(0.223)	(0.235)	(0.139)	(0.246)
N	442	442	442	442
R-sq	0.059	0.085	0.043	0.084
adj. R-sq	0.039	0.053	0.023	0.051

Standard errors in parentheses
="* p<0.10 ** p<0.05 *** p<0.01"

Table 6 q-values:

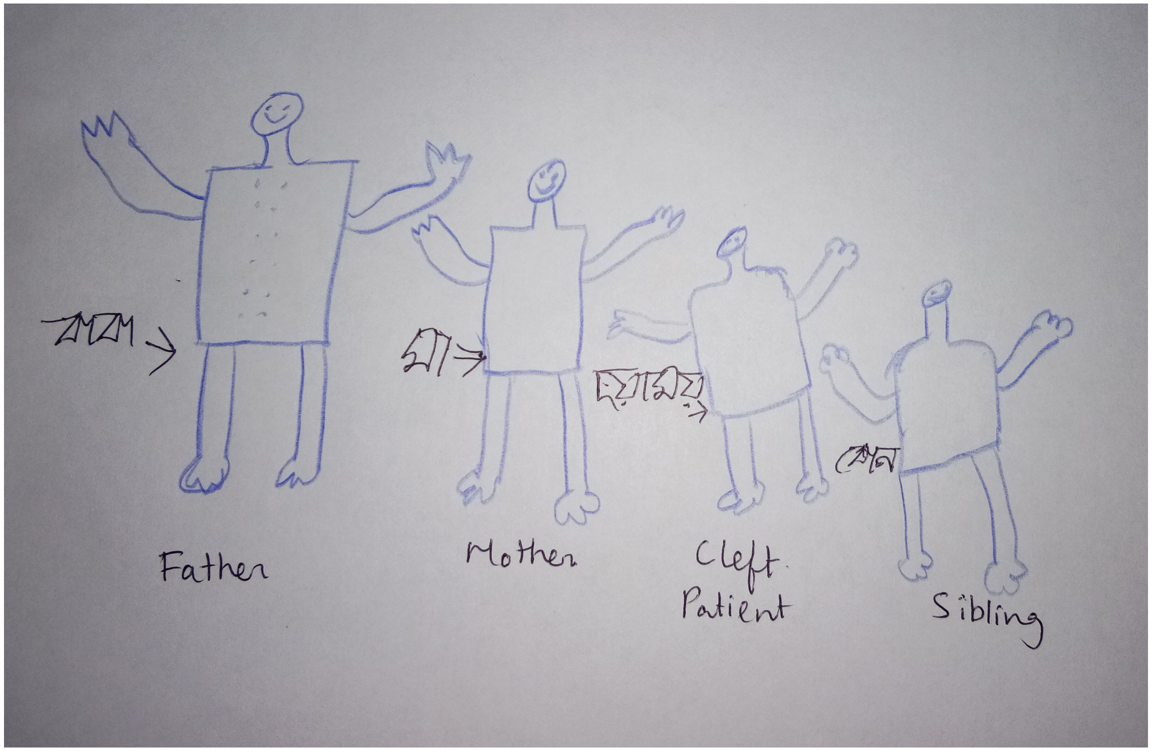
	(1) p-values	(1) q-values	(2) p-values	(2) q-values	(3) p-values	(3) q-values	(4) p-values	(4) q-values
ShadingFace	0.7703	0.7703	0.7132	0.9032	0.4909	0.9443	0.5743	0.8643
MissingNoseMouth	0.6717	0.7703	0.6984	0.9032	0.8528	0.9443	0.5065	0.8643
FrowningCrying	0.3191	0.7703	0.4220	0.9032	0.6446	0.9443	0.4285	0.8643
DarkColors	0.5128	0.7703	0.5926	0.9032	0.6297	0.9443	0.4513	0.8643
SingleColor	0.0266	0.1862	0.0349	0.4185	0.3708	0.9443	0.1008	0.8643
PoorFiguralIntegration	0.1245	0.6224	0.1287	0.9032	0.9443	0.9443	0.8643	0.8643
Smiling	0.0016	0.0146	0.0017	0.0259			0.0029	0.0407
LightCheeryColors	0.1175	0.6224	0.1069	0.9032			0.3911	0.8643
SketchyBrokenLines	0.0116	0.0930	0.0272	0.3533			0.1208	0.8643
TinyFigure			0.0066	0.0927	0.0002	0.0014	0.0006	0.0086
FaintLines			0.7375	0.9032	0.3128	0.9443	0.3878	0.8643
TinyHead			0.9032	0.9032	0.9091	0.9443	0.5179	0.8643
LackofDetail			0.1345	0.9032			0.8214	0.8643
SmallerinRelativeSize			0.0446	0.4905			0.1647	0.8643
SeparatedFromParents			0.5241	0.9032			0.0674	0.8643
Intercept	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 7: LASSO Regression

	(1) Anxiety LASSO	(2) Anxiety Post-Est OLS	(3) Depression LASSO	(4) Depression Post-Est OLS
SingleColor	-0.0788	-0.240*		
		(0.137)		
Smiling	0.254	0.406***	0.196	0.363***
		(0.131)		(0.137)
SketchyBrokenLines	-0.244	-0.342**	-0.0969	-0.266
		(0.155)		(0.161)
TinyFigure	0.268	0.488***	0.409	0.592***
		(0.146)		(0.146)
SmallerinRelativeSize	-0.0825	-0.461**		
		(0.227)		
SeparatedFromParents			0.0760	0.233*
				(0.134)
DifferentDrawingFromPrompt			-0.129	-0.949
				(0.632)
Intercept	-2.138	-2.128***	-2.416	-2.601***
		(0.153)		(0.141)
N	440	442	440	440
R-sq		0.070		0.076
adj. R-sq		0.059		0.065

Standard errors in parentheses
="* p<0.10 ** p<0.05 *** p<0.01"

Figure 1: Drawing and Coding Example:



Indicator	Present?
Shading Face	0
Missing Nose/Mouth	1
Frowning/Crying	0
Dark Colors	1
Single Color	1
Poor Figural Integration	0
Smiling	1
Light Cheery Colors	0
Sketchy Broken Lines	0
Tiny Figure	0
Faint Lines	0
Tiny Head	1
Lack of Detail	1
Smaller in Relative Size	0
Separated from Parent	0

Table 8: Indicators Associated with Risk of Anxiety and Depression:

Drawing Indicator	Anxiety	Depression
Shading Face	1	1
Missing Nose/Mouth	1	1
Frowning/Crying	1	1
Dark Colors	1	1
Single Color	1	1
Poor Figural Integration	1	1
Smiling	-1	0
Light Cheery Colors	-1	0
Sketchy Broken Lines	1	0
Tiny Figure	0	1
Faint Lines	0	1
Tiny Head	0	1