Children's Aggressiveness and Prosociality Prediction and Comparison

Ya-Yun Chen
Department of Psychology
Virginia Tech
Blacksburg, USA
yayunchen@vt.edu

Tai-Jung Chen
Department of Industrial & Systems Engineering
Virginia Tech
Blacksburg, USA
taijungchen@vt.edu

ABSTRACT

Aggression and prosociality are crucial for the well-being of individuals and society as a whole. The understanding of how humans develop aggression and prosociality will provide much help in developing more appropriate intervention and policy for preventing the ill-being. The impact of family factors and neighborhood factors have been confirmed in previous studies of aggression and prosociality. However, they have typically been studied separately and stay at the behavioral level.

This study aimed to build prediction models by using family, parental and neighborhood factors as well as neuroimaging data as the predictive attributes to predict these two directions of social behaviors. Furthermore, we compared the similarities and differences between the models of aggressive and prosocial tendencies to measure whether the relationship between aggressive behaviors and prosocial behaviors is orthogonal or simple dichotomies.

Models of decision tree demonstrated that the first three layers of the aggressive model (accuracy = .76, kappa = .49) and prosocial models (accuracy = .62, kappa = .29) shared 3 common attributes in terms of the family and neighborhood factors out of 7 nodes. Specifically, the common nodes were the family relationship, child ever received mental health or substance abuse services, and family environment conflict. For the brain connectivity as the predictive attributes, however, only the frontoparietal network to putamen connectivity is the common node. A further logistic regression analysis with both family-neighborhood and brain factors revealed that on predicting aggressive tendency the odds ratio of the family relationship was 8.24 and whether the child ever received mental health or substance abuse services was 4.29.

The current findings confirmed the previous literature in the field by demonstrating the family and neighborhood factors do predict the children's social behaviors. The results further suggest that the relationship between aggressive and prosocial tendencies is not orthogonal nor simple dichotomies.

KEYWORDS

Aggressive Behaviors, Prosocial Behaviors, Decision tree, Logistic regression, Ensemble Modeling

1 Problem Statement

With appropriate social behaviors, individuals are able to shape friendly social exchanges and to form reciprocal relations groups. Then, a high well-being society is established. Such appropriate social behaviors are called prosocial behaviors, which include helping, sharing, and comforting [1]. Conversely, inappropriate social behaviors may incur an ill reputation of individuals, bring vicious consequences to one's interpersonal relationships, and expand these maleficent qualities to the whole society. One of the notorious inappropriate social behaviors is aggressive behavior, which could be categorized as proactive aggression or reactive aggression [2,3]. Aggressive behavior and prosocial behavior serve as two disparate influences toward individuals, groups, and society, competitively. A greater understanding of how humans develop these two sorts of behaviors will offer insight into establishing a family and social safety net, preventing violence, and promoting happiness.

Previous studies have shown that parents influence both ends of children's social behaviors in the manner of parenting and through the quality of the mother-child relationship and parental interaction [4-7]. Research also suggests that social-economic state and the living neighborhood can be risk factors in shaping children's behaviors [8,9]. In addition, neural evidence suggests that empathy networks and affective and regulatory neural circuits are related to different sorts of prosocial and aggressive behaviors (e.g., proactive versus reactive aggression). These networks and circuits are vulnerable to interpersonal relationships and growing environments [10,11].

Being a crucial issue regarding the well-being of individuals and society, the research on aggressive behavior and prosocial behavior has accumulated for decades, however, separately. The direct competition, orthogonal or simple dichotomies, between these two sorts of social behaviors was explored less. The current analysis primarily aimed to use decision trees to explore the relationship between neighborhood/parents/family, social behavior related neural networks, and children's social behaviors. Specifically, (1) we used decision tree techniques to build the prediction models of aggressive behavior and a prediction model of prosocial behavior; (2) the similarity of the prediction model between the model of aggressive behavior and model of prosocial behavior will be compared, by extracting out the first 3 layers of the trees; (3) ensemble and logistic regression models were further applied to the model of aggressive behavior, trying to improve the predictability of the model.

2 Data Description

The preprocessing steps are as follows.

1. Concatenate all datasets with the inner join method.
2. Remove NaN, 777 (refuse to answer), and 999 (don't know). With the step (1) and (2) a relatively clean dataset was generated. Because the ABCD dataset is a well-controlled repository, over 7,000 samples remain.
3. Aggregate the negative social behavior scores and positive social behavior scores respectively. Specifically, for the aggressive tendency, the raw score of each survey related to children's aggression behaviors (i.e., surveys of rule break, aggression, and conduct behaviors) was normalized, and then we sum up these normalized scores to become a single score to represent the aggressive tendency (range from -2.02 to 33.37 with mean = 0.00 and median = -1.61, see figure 1(a), a higher score indicates higher aggressive tendency). For the magnitude of prosociality, because there is no a strong correlation between parent-report and child self-report prosociality (r = 0.18), we separately summed up the parent-report prosocial behavior survey to be a parent-report prosociality and summed up the child self-report prosocial behavior survey to be a child-report prosociality (range = 0-6, which 6 means highest prosociality, with mean/median = 5.25/6 for parent report and 5.03/5 for child-report, see figure 1(b).
4. Use domain knowledge to select the target cortical network to subcortical region connectivity. The cortical networks include the cingulo-opercular network, default network, dorsal and ventral attention network, frontoparietal network, and salience network, which are engaging in the functions of self-control, emotion regulation, self-related inner mental processing, and attention switching. The subcortical regions include the thalamus, caudate, putamen, brain-stem, hippocampus, amygdala, and accumbens, which are related to the strength of sensory input, reward processing, and emotion processing.

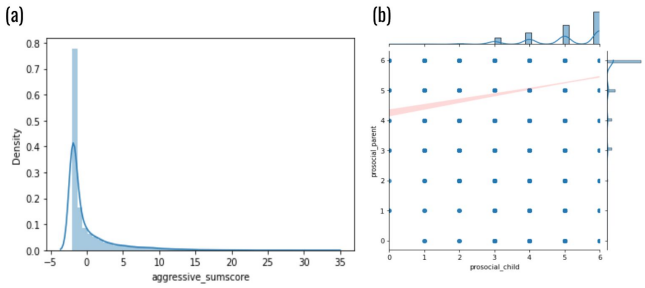

Figure 1. (a) The distribution of the children's aggressive tendency. (b) The correlation and marginal distribution of the children's prosociality.

3 Data Exploration

We used data exploration techniques to implement attribute selection and labeling. In brief, the correlation analyses were applied to select potential attributes and the descriptive statistics were utilized to group high and low aggressive groups and prosocial groups respectively.

3.1 Attribute selection

Originally there were 55 attributes representing neighborhood-parental problems-family (N-P-F) factors. The correlation between aggressive tendency/prosociality and the N-P-F factors were calculated in order to do attribute selection. The inclusion criteria are the correlation coefficient larger than 0.1 or smaller than -0.1. There were 26 attributes that survived for predicting aggressive tendency and 15 and 2 attributes survived for predicting parent-report prosociality and child-report prosociality respectively.

3.2 Labeling

For the labels of aggressive tendency, we assigned children with score = minimum score (-2.02) as the non-aggressive group (label = 0, N = 4534) and score larger than or equal to the 75 percentile (0.35-33) as the aggressive group (label = 1, N = 2903). For the labels of prosocial tendency, we assigned children with score smaller than or equal to 3 as the low prosocial group (label = 0, N = 1368 for parent report and N = 1302 for child report) and score = 6 as the high prosocial group (label = 1, N = 7328 for parent report and N = 5242 for child report). For the imbalanced data in prosociality, bootstraps with N = 1500 for each group were applied.

4 Model Building and Evaluation

The main goal of the current project is to compare the similarity of the prediction models between the aggressive model and prosocial models. As such, we used the attributes survived in correlation analysis in aggressive tendency and brain connectivities to build the aggressive model and prosocial models. If these two sorts of social behaviors are a simple dichotomy, the prediction accuracy and the structure of the decision trees should be similar. In contrast, if they are orthogonal, the prediction accuracy of the prosocial model would be around the chance level or the structure of the decision trees are different from the aggressive model. The following indices were implemented by 10-fold cross-validation.

4.1 Aggressive model predicted by the neighborhood-parental problems-family factors

The accuracy = 0.76, f1 score = 0.69, recall = 0.70, precision = 0.68, Kappa = 0.49, tree depth = 28 with binary leaf in each node. Figure 2(a) shows the first three layers of the trees of the aggressive model predicted by the N-P-F factors.

4.2 Aggressive model predicted by the brain connectivities

The accuracy = 0.58, f1 score = 0.47, recall = 0.47, precision = 0.46, Kappa = 0.12, tree depth = 28 with binary leaf in each node. Figure 3(a) shows the first three layers of the trees of the aggressive model predicted by the N-P-F factors.

4.3 Parent-report prosocial model predicted by the neighborhood-parental problems-family

The accuracy = 0.64, f1 score = 0.65, recall = 0.63, precision = 0.67, Kappa = 0.29, tree depth = 29 with binary leaf in each node. Figure 2(b) shows the first three layers of the trees of the prosocial model predicted by the N-P-F factors.

4.4 Parent-report prosocial model predicted by the brain connectivities

The accuracy = 0.50, f1 score = 0.51, recall = 0.51, precision = 0.50, Kappa = 0.01, tree depth = 27 with binary leaf in each node. Figure 3(b) shows the first three layers of the trees of the prosocial model predicted by the N-P-F factors.

4.5 Child-report prosocial model predicted by the neighborhood-parental problems-family

The accuracy = 0.61, f1 score = 0.63, recall = 0.52, precision = 0.64, Kappa = 0.22, tree depth = 23 with binary leaf in each node.

4.6 Child-report prosocial model predicted by the brain connectivities

The accuracy = 0.52, f1 score = 0.53, recall = 0.54, precision = 0.52, Kappa = 0.04, tree depth = 29 with binary leaf in each node.

Because the performance of parent-report prosocial models were better than the child-report models, the following comparisons were made between the aggressive model and the parent-report prosocial model.

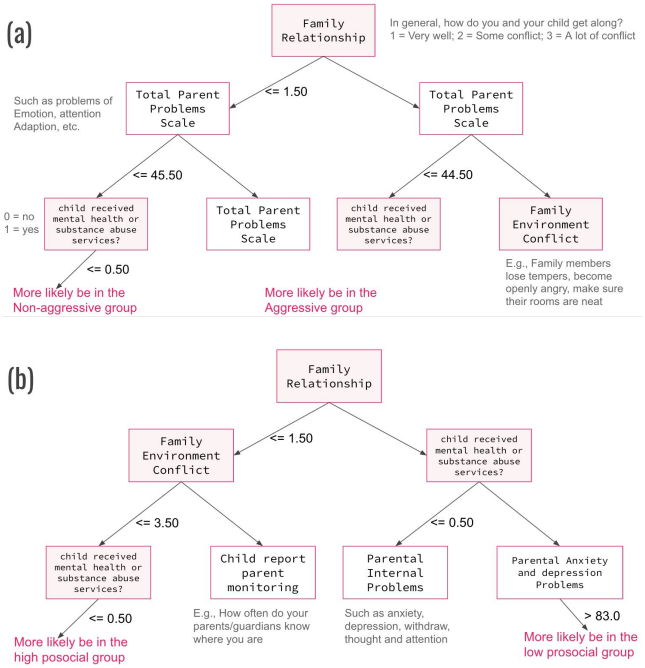

Figure 2. (a) The first three layers of the trees of the aggressive model predicted by the N-P-F factors. (b) The first three layers of the trees of the aggressive model predicted by the N-P-F factors. Boxes filled with color indicated that the factors were shared by two models.

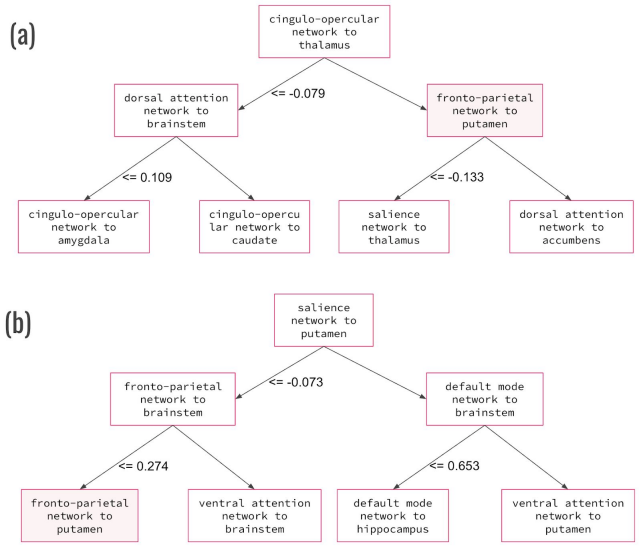

Figure 3. (a) The first three layers of the trees of the aggressive model predicted by the cortical network to subcortical region connectives. (b) The first three layers of the trees of the aggressive model predicted by the cortical network to subcortical region connectives. Boxes filled with color indicated that the factors were shared by two models.

5 Predictability Improvement

The best accuracy and kappa of the decision trees of the aggressive model were .76 and .49 respectively. We further tried to use ensemble models and logistic regression models to improve the predictability of the aggressive model. For the ensemble models, (i) the prediction results from the N-P-F factors and brain connectivities trees were fed into a logistic model to generate a final prediction (see figure 4 for illustration). Unfortunately, the accuracy and kappa of this ensemble model were decreased slightly to .75 and .46 respectively. (ii) We further tried to feed more attributes in the ensemble model as the second attempt to improve the performance. The accuracy and kappa of this second attempt model were improved slightly to .78 and .53 respectively.

Since the N-P-F factors and brain data improved the performance of the models, we tried to feed these two categories only without the prediction results from the decision trees to the logistic regression model and found that the accuracy and kappa of this logistic regression model were significantly improved to .82 and .60 respectively.

A further examination with two separated regression models found that the improvements were from the N-P-F factors (accuracy = .83; kappa = .64). There were 6 attributes that the odds ratio larger than 1.5, which were family relationship (conflict between parent and child): 8.24, whether the child ever received mental health or substance abuse services: 4.29, child have any problems with bullying at school or in the neighborhood: 2.48, family financial problems (ever could not afford somethings): 1.76, the frontoparietal network to putamen connectivity: 1.67, and the dorsal attention network to hippocampus connectivity: 1.66. Among these 6 attributes, half of them exist in the first 3 layers of the model of decision trees, i.e., family relationship, received mental health or substance abuse services, and the frontoparietal network to putamen connectivity.

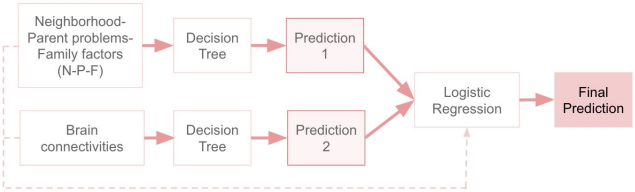

Figure 4. The illustration of the ensemble models. The path of solid lines represents the first attempt (i) of the ensemble models. The path of solid lines with a dotted line represents the second attempt (ii) of the ensemble models.

6 Real-world Insights

The current study used decision tree models trying to resolve the “orthogonal or simple dichotomies” problem between aggressive and prosocial behaviors. By comparing the first three layers of decision tree models, we found that the family relationship (conflict between parent and child) plays a pivotal role in predicting children’s aggressive and prosocial tendencies. The lower conflict between parent and child predicts lower aggressive tendency and higher prosociality. The predictability of the family relationship is further supported by the logistic regression model, in which the odds ratio was 8.24, suggesting that every increase of 1 point of the conflict increases 8 times the probability of a child being in the higher aggressive tendency group.

Other important factors shared by both aggressive and prosocial tendency predictions are: whether the child ever received mental health or substance abuse services, family environment conflict, and the frontoparietal network to putamen connectivity.

The current findings suggest that the relationship between aggressive and prosocial tendencies is not orthogonal nor simple dichotomies. In terms of the N-P-F

factors, aggressive and prosocial tendencies do have some predictive attributes in common but still have some differences among them. For example, the differences are shown in that the parental internal problems may predict children's prosocial tendency but not the aggressive tendency. In sum, further exploration is needed to dissect the latent component underlying aggressive and prosocial behaviors. In terms of brain connectivity, the differences between aggressive and prosocial tendencies are more obvious, for all 7 nodes in the first three layers of the trees, only the node of the frontoparietal network to putamen connectivity is common. However, the performance of the brain connectivity models was around chance (accuracy = .50-.58) it is restrictive to make a conclusion based on this model.

ACKNOWLEDGMENTS

Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive DevelopmentSM (ABCD) Study (https://abcdstudy.org), held in the NIMH Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children's age 9-10 and follow them over 10 years into early adulthood. The ABCD Study® is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, U24DA041147. A full list of supporters is available at https://abcdstudy.org/federal-partners.html. A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/consortium_members/. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators.

CONTRIBUTIONS

Ya-Yun Chen formed the research question, planned analyses, and did data interpretation. Tai-Jung Chen worked on Python programming, model evaluation, and parameter optimization.

REFERENCES

[1] *DUNFIELD, K. A. (2014). A construct divided: prosocial behavior as helping, sharing, and comforting subtypes. Frontiers in Psychology, 5, 958.*

[2] *DODGE, K. A., & COIE, J. D. (1987). Social-information-processing factors in reactive and proactive aggression in children's peer groups. Journal of Personality and Social Psychology, 53, 1146 –1158.*

[3] *DODGE, K. A., LOCHMAN, J. E., HARNISH, J. D., BATES, J. E., & PETTIT, G. S. (1997). Reactive and proactive aggression in school children and psychiatrically impaired chronically assaultive youth. Journal of Abnormal Psychology, 106, 37–51.*

[4] *KAWABATA, Y., ALINK, L. R., TSENG, W. L., VAN IJZENDOORN, M. H., & CRICK, N. R. (2011). Maternal and paternal parenting styles associated with relational aggression in children and adolescents: A conceptual analysis and meta-analytic review. Developmental review, 31(4), 240-278.*

[5] *BOOTH, C. L., ROSE-KRASNOR, L., MCKINNON, J., & RUBIN, K. H. (1994). Predicting social adjustment in middle childhood: The role of preschool attachment security and maternal style. Social Development, 3, 189 – 204.*

[6] *DUNFIELD, K. A. (2014). A construct divided: prosocial behavior as helping, sharing, and comforting subtypes. Frontiers in Psychology, 5, 958.*

[7] *FLOOK, L., ZAHN-WAXLER, C., & DAVIDSON, R. J. (2019). Developmental differences in prosocial behavior between preschool and late elementary school. Frontiers in psychology, 10, 876.*

[8] *GALLO, L. C. & MATTHEWS, K. A. (2003). Understanding the association between socioeconomic status and physical health: Do negative emotions play a role? Psychological Bulletin, 129(1), 10-51.*

[9] *MORRIS, A. S., SILK, J. S., STEINBERG, L., MYERS, S. S., & ROBINSON, L. R. (2007). The role of the family context in the development of emotion regulation. Social Development, 16, 361-388.*

[10] *Lee, T. H., MIERNICKI, M. E., & TELZER, E. H. (2017a). Behavioral and neural concordance in parent-child dyadic sleep patterns. Developmental cognitive neuroscience, 26, 77-83.*

[11] *LEE, T. H., MIERNICKI, M. E., & TELZER, E. H. (2017b). Families that fire together smile together: Resting state connectome similarity and daily emotional synchrony in parent-child dyads. Neuroimage, 152, 31-37*

[12] *CASEY, B. J., CANNONIER, T., CONLEY, M. I., COHEN, A. O., BARCH, D. M., HEITZEG, M. M. et al (2018). The adolescent brain cognitive development (ABCD) study: imaging acquisition across 21 sites. Developmental cognitive neuroscience, 32, 43-54.*

[13] *MCKINNEY, W. et al. (2010). Data structures for statistical computing in python. In Proceedings of the 9th Python in Science Conference (Vol. 445, pp. 51–56).*