POLITECNICO DI TORINO

Master's Degree in Mathematical Engineering

Master's Degree Thesis

Study on the effects of children's screen time activity on their mental health and brain structure

Supervisors:

Gianluca MASTRANTONIO
Department of Mathematical Sciences "Giuseppe Luigi Lagrange"
Polytechnic of Turin

Fabrizio PIZZAGALLI
Department of Neuroscience "Rita Levi Montalcini"
University of Turin

Candidate:

Marco BOTTINO

A.Y. 2020/2021

Abstract

The generation of *digital natives*, which was born between the end of the last century and the beginning of the current one, has been growing up surrounded by a screen in every part of their lives: screens are used to study, for entertainment, and social activities. Children in the age 8-18 spend, on average, 7.5 hours every day in front of a screen, 4.5 of which are spent watching TV. The effects of screen time activity on their mental health and brain structure are still not fully understood, and evidence for an impact of screentime on the health is inconsistent, with systematic reviews showing mixed findings. This may in part be due to the difficulty of separating the effects of screentime from other consequences of sedentary life, like low physical activity or poor sleep quality. Moreover, most of the literature is based on cross-sectional data, therefore it struggles to provide evidence on causal association or the direction of the relationships between screentime activity, mental health, and brain structure. This thesis aims to examine these relationships through the use of statistical models fitted on data from Adolescent Brain Cognitive Development longitudinal study, a US representative sample of more than 11000 children in age 8-12. Screentime activity is divided into four main types: TV, video (e.g., YouTube), video games, and social activities. Mental health is evaluated through the use of psychometric scales obtained from Children Behavior Checklist, whereas brain structure is described by anatomical measures like intracranial volume (ICV), cortical thickness, and cortical surface area, extracted through anatomical MRI (T1w) of the subjects' brain. The analyses showed evidence of a positive association between time spent watching TV and the scores concerning externalizing problems, like rule-breaking and aggressive behavior, whereas social activities have a negative association with internalizing problems like anxiety and withdrawal. On the other hand, time spent playing video games proved to have a significant effect on the thinning of the cerebral cortex, especially in the small frontal and temporal regions.

Ringraziamenti

Desidero innanzitutto ringraziare il professor Gianluca Mastrantonio per la prontezza e la precisione nel risolvere i miei dubbi su questo lavoro e il professor Fabrizio Pizzagalli per avermi fatto appassionare ulteriormente al mondo delle neuroscienze, fornendomi tutto il materiale e l'aiuto necessario. Senza di loro non avrei mai potuto portare a termine questa tesi.
Ringrazio inoltre tutti i miei compagni di corso che mi hanno accompagnato in questi 5 anni, sostenendoci a vicenda nei momenti di difficoltà e festeggiando insieme i successi raggiunti.
Ringrazio i miei amici, coi quali ho avuto la fortuna di crescere, viaggiare, e mordere la vita. L'elenco sarebbe innumerevole, cosa di cui non smetterò mai di ritenermi profondamente fortunato, ma vorrei fare una menzione speciale per Stewe, che mi ha insegnato quanto può essere importante la quotidianità in una relazione per crescere, e per Rara, che mi ha insegnato a perdonarmi per i miei errori e che lo ha fatto per me quando io non riuscivo.
Ringrazio anche Ilaria, protagonista delle mie giornate e dei miei pensieri per la maggior parte di questo percorso. Senza la sua dolcezza probabilmente non sarei diventato così diverso dal computer con cui lavoro.
Ringrazio inoltre esattamente quel computer, che mi ha portato attraverso l'esame di maturità, 40 esami universitari, decine di progetti e due tesi. Ora può godersi la meritata pensione.
Ringrazio il Campus e la Comala, che mi hanno permesso di mantenere un minimo di vita sociale anche nelle sessioni più dure. E' grazie a posti come questi che il percorso universitario diventa molto più divertente e lo studio molto più leggero.

Il più grosso ringraziamento va però alla mia famiglia, che da quando sono nato mi dona affetto e serenità senza chiedere nulla in cambio. Non avrei raggiunto nessun risultato senza di loro, ed è a loro che dedico tutto ciò che ho conseguito, fra cui questo lavoro.

Contents

Introduction

The screens are a symbol of the modern age, whether they are computers, televisions, tablets or smartphones. The children born between the end of the last century and the beginning of the new one are called *digital natives*, since they have grown up surrounded by digital information, social media and entertainment on-screen. According to the Centres for Disease Control and Prevention (CDC), children in the age 8-18 spend, on average, 7.5 hours in front of a screen for entertainment each day, 4.5 of which are spent watching TV [1].

The effect on well-being and psychological health of the time spent in front of a screen, which from now on will be called screentime activity, has been a recurring topic in literature in the latest years: increased screentime has been associated with unfavorable body composition, higher cardiometabolic risk, lower fitness, and lower self-esteem in children [2]. Evidence for an impact of screentime on mental health on the other hand is inconsistent, with systematic reviews showing mixing findings [3]. Moreover, most of the studies are based on cross-sectional data, which do not allow to infer the direction of causality in the associations found. A possible cause of this impact could consist in changes in brain structure, especially in the cortex: different associations between mental health and cortical abnormalities have been found in literature, especially with changes in cortical thickness and cortical surface area [4][5][6].

The purpose of this thesis is to bring those studies one step further by investigating how screentime activities affect both school-age children's mental health and brain structure. The analyses are based on a longitudinal study from the USA with data from more than 11000 children between 8 and 12 years old and seek associations between different types of screentime activities and measures of their mental health or their cortical structure. The thesis is divided into three main chapters:

- Chapter 1 contains a general analysis of the working datasets. The first section introduces the target variables, which consist in the scales from Children Behavior Checklist [7] for what concerns the children's mental health and cortical measures extracted from T1-weighted MRI images for what concerns the children's brain structure. The second section introduces all the features chosen as predictors, including the screentime for different kinds of activities. The third section describes the preprocessing of the data, through which data duplicates, outliers and redundant features were removed in order to be able to fit statistical models.

- Chapter 2 introduces the main models used, with a theoretical discussion on multiple linear regression and generalized linear models in the first section and a description of the techniques used to perform model selection and validation in the second one.

- Chapter 3 describes the results of the optimal models for each of the target variables, with a discussion on the effects of the different predictors. The first section concerns the main results on the scales from CBCL, reporting the significant effects for each scale, with a focus on the effects of different types of screentime activity. The second section does the same for what concerns the brain measures, with further analysis on the local effects of screentime activities on the cortex's regions of interest. The third section presents an analysis on the causality of the associations previously found.

Appendix A contains the tables with all the estimated coefficients from the optimal models fitted in this work.
All the plots in the study have been obtained through the use of the programming language Python 3.0 and its open-source libraries Matplotlib and Seaborn. The statistical analyses were conducted in R, a language and environment for statistical computing [8].

Chapter 1

Dataset Exploration

The work is based on longitudinal data from the Adolescent Brain Cognitive Development (ABCD) Study® [9][10], which is an ongoing observational study exploring the development and health among children from an age of 8 years through early adulthood, with a focus on brain health and cognition. Data for this study are collected on a biennial-to-annual basis over a 10-year period across 21 sites throughout the United States. They include information about demographics, physical health, mental health, brain imaging, culture and much more. The ABCD data repository grows and changes over time: the data used in this report came from the fast track data release. The raw data are available at `https://nda.nih.gov/edit_collection.html?id=2573`, whereas instructions on how to create an NDA study are available at `https://nda.nih.gov/training/modules/study.html`. NDA has granted access to the ABCD data to Dr. Fabrizio Pizzagalli based on the proposed project.
The starting dataset for this work has been created by merging multiple datasets containing different features of interest through the subject IDs and the timestamp of the sample. It contains data from the first three years of the study (2016-2018), which are composed of 29703 samples divided into 11877 from the first year, or baseline, 11248 from the second year or 1-year follow-up and 6578 from the third year or 2-year follow-up. The following sections first describe the variables chosen as target and predictors, then the preprocessing made to clean the dataset and get it ready to be fit in a model. All the histograms in Section 1.1 and Section 1.2 represent the distribution of the features after the cleaning process.

1.1 Target variables

The targets of the following analyses are the children's mental health and their brain structure. The first one is evaluated in this study through the psychometric scales provided by Children Behavior Checklist, whereas the second one is described by the anatomical measures of the children's intracranial volume and their cortical thickness and surface. These measures were obtained after the preprocessing of anatomical MRI (T1-weighted) from their brains.

	cbcl_score	anx_depr	with_depr	som_comp	social_pr	thought_pr	att_pr	rule_br_bh	agg_bh
mean	15.37	2.53	1.06	1.47	1.53	1.61	2.89	1.14	3.13
std	15.84	3.06	1.72	1.94	2.20	2.19	3.44	1.81	4.23
min	0	0	0	0	0	0	0	0	0
25%	4	0	0	0	0	0	0	0	0
50%	10	2	0	1	1	1	2	0	2
75%	21	4	1	2	2	2	5	2	4
93%	42	8	4	5	5	5	9	4	10
97%	56	10	6	6	7	7	12	6	15
max	122	25	15	18	17	20	20	20	36

Table 1.1: Statistics of CBCL scores. The first column represents the total score, which is the sum of all the syndrome scale scores, described by the following eight columns.

Children Behavior Checklist (CBCL)

The Child Behavior Checklist 6-18 is a widely used caregiver report form identifying problem behavior in school-age children. It is a component in the Achenbach System of Empirically Based Assessment (ASEBA) developed by Thomas M. Achenbach [7], and it is based on 112 problem behavior questions. Responses are recorded on a Likert scale with the following codification: 0 = Not True, 1 = Somewhat or Sometimes True, 2 = Very True or Often True. Similar questions are grouped into a number of syndrome scale scores, and their scores are summed to produce a raw score for that syndrome. The 8 empirically-based syndrome scales are: aggressive behavior, anxious/depressed, attention problems, rule-breaking behavior, somatic complaints, social problems, thought problems, withdrawn/depressed. For each syndrome and the total score, their values are used to classify each child in the *normal*, *borderline*, or *clinical* behavior, using the following thresholds: any score that falls below the 93^{rd} percentile is considered normal, scores between the $93 - 97^{th}$ percentile are borderline clinical, and any score above the 97^{th} percentile are in the clinical range.

Figure 1.1 shows the distribution of the total score and the 8 scales. They are all highly asymmetrically distributed towards zero, which will be important in the choice of the right model. Tab.1.1 on the other hand shows some basic statistics from all the scales: the total scores considered as normal fall under 42 in the total score, whereas they start to be clinical when greater than 56.

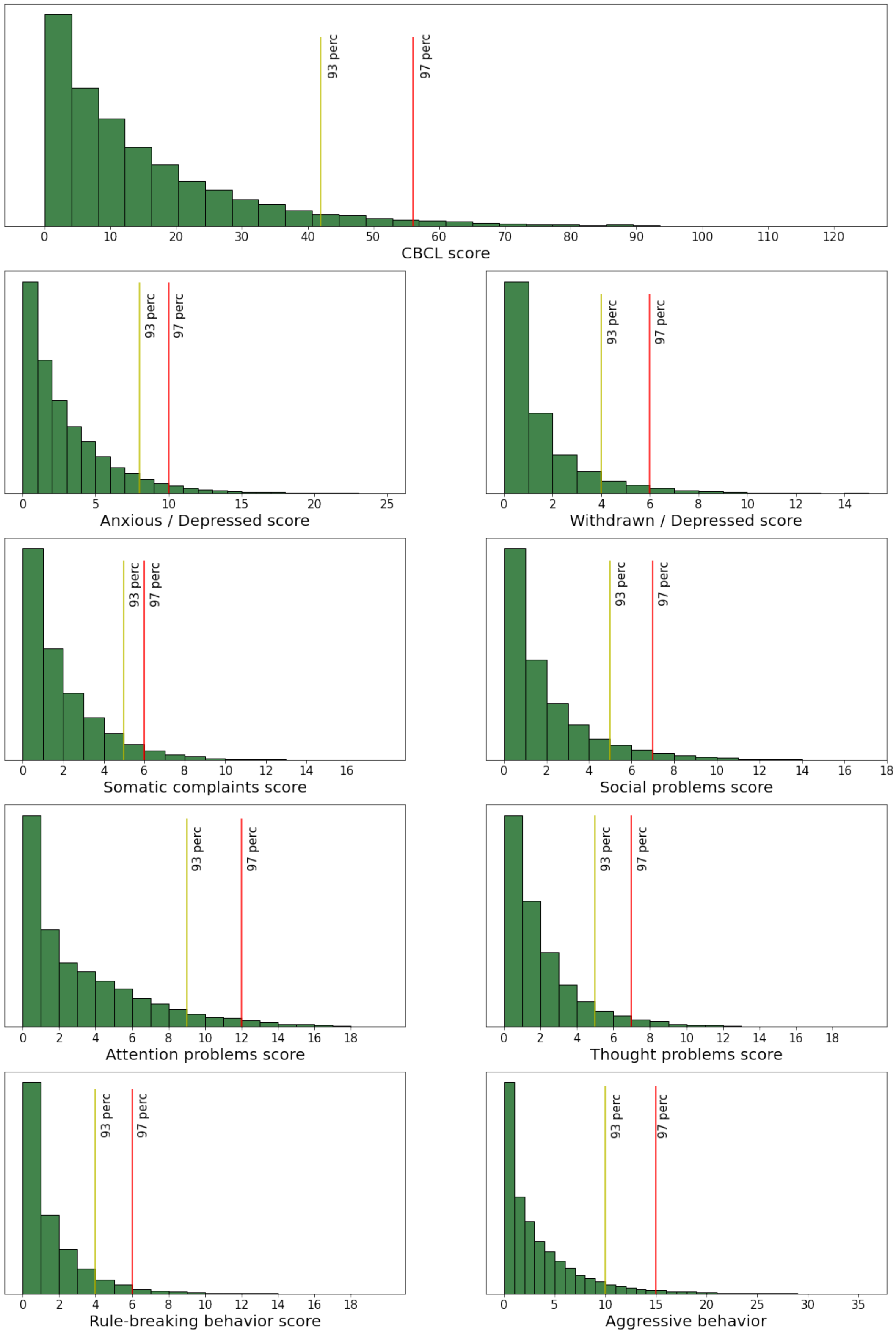


Figure 1.1: Distribution of CBCL target variables. The yellow lines represent the threshold from *normal* to *borderline* behavior, whereas the red ones represent the threshold from *borderline* to *clinical* behavior.

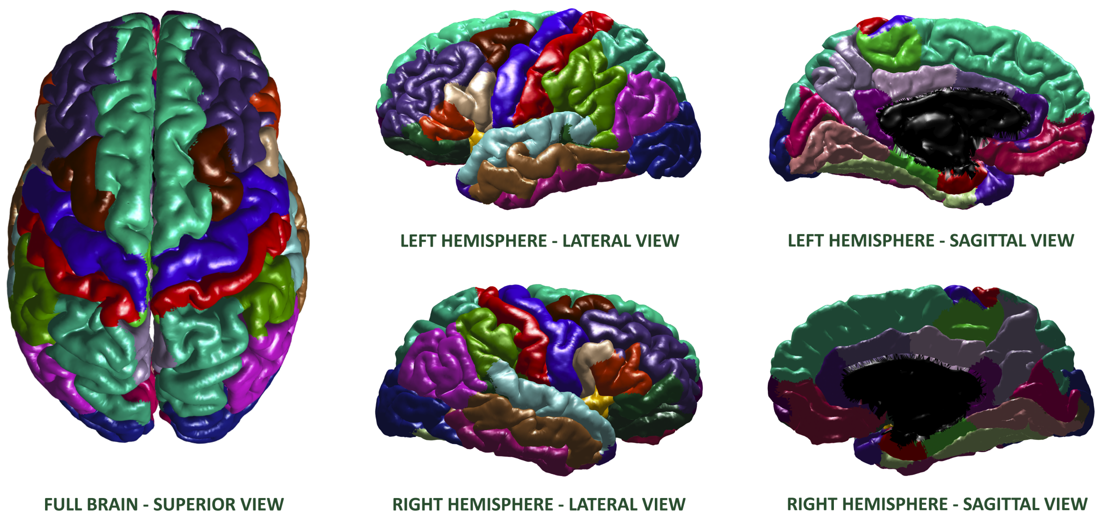


Figure 1.2: Cortical parcellation from Desikan–Killiany atlas.

Brain measures

A subset of the subjects provided anatomical MRI images (T1-weighted) of their brains. These images were corrected for intensity and cortical segmentations and parcellations were created with the freely available and validated segmentation software, FreeSurfer [13]. Segmentations of 68 (34 left and 34 right) cortical gray matter regions were created based on the Desikan–Killiany atlas [14], as shown in Figure 1.2. ROI-based thickness and surface area were extracted, as well as the hemispheric total surface area, thickness and intracranial cortical volume (ICV). All these measures were then merged to the dataset as target variables to be modeled: the histograms in Figure 1.3 show the distributions of the total measures, which have real non-negative values and seem to follow a Gaussian distribution, whereas Table 1.2 describes the main feature statistics. The ICV is measured in mm^3, the cortical surface area in mm^2 and the cortical thickness in mm.

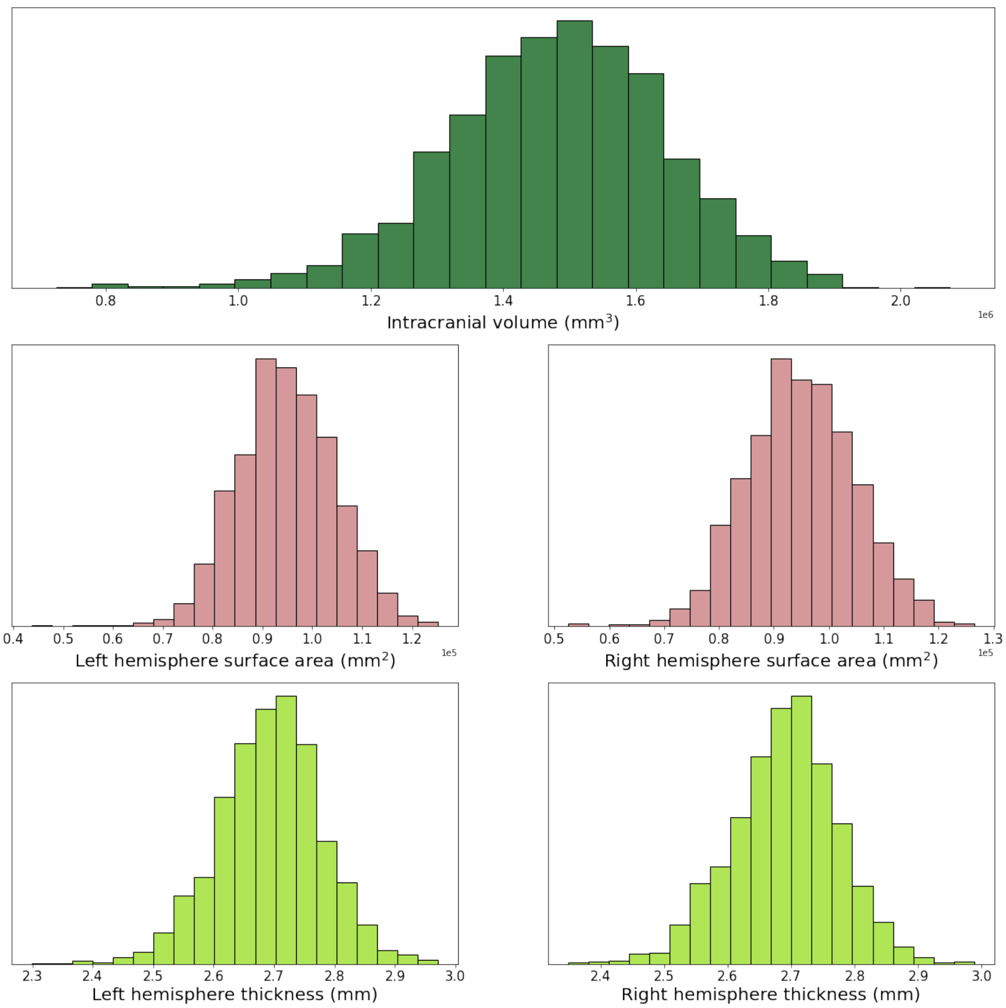


Figure 1.3: Distribution of brain measures.

	ICV	Left thickness	Right thickness	Left surface area	Right surface area
mean	1480709.56	2.69	2.69	94737.65	94973.95
std	168481.34	0.09	0.09	9739.07	9828.59
min	724791	2.30	2.34	43685	52467
25%	839567	2.64	2.64	88379	88400
50%	942605	2.70	2.70	94410	94825
75%	1003344	2.75	2.74	101454	101807
max	2074240	2.97	2.99	125207	126504

Table 1.2: Statistics of brain measures. From the left the columns represent statistical details of intracranial volume (ICV) in mm^3, cortical thickness left and right hemispheres in mm and cortical surface area of left and right hemispheres in mm^2.

1.2 Predictor variables

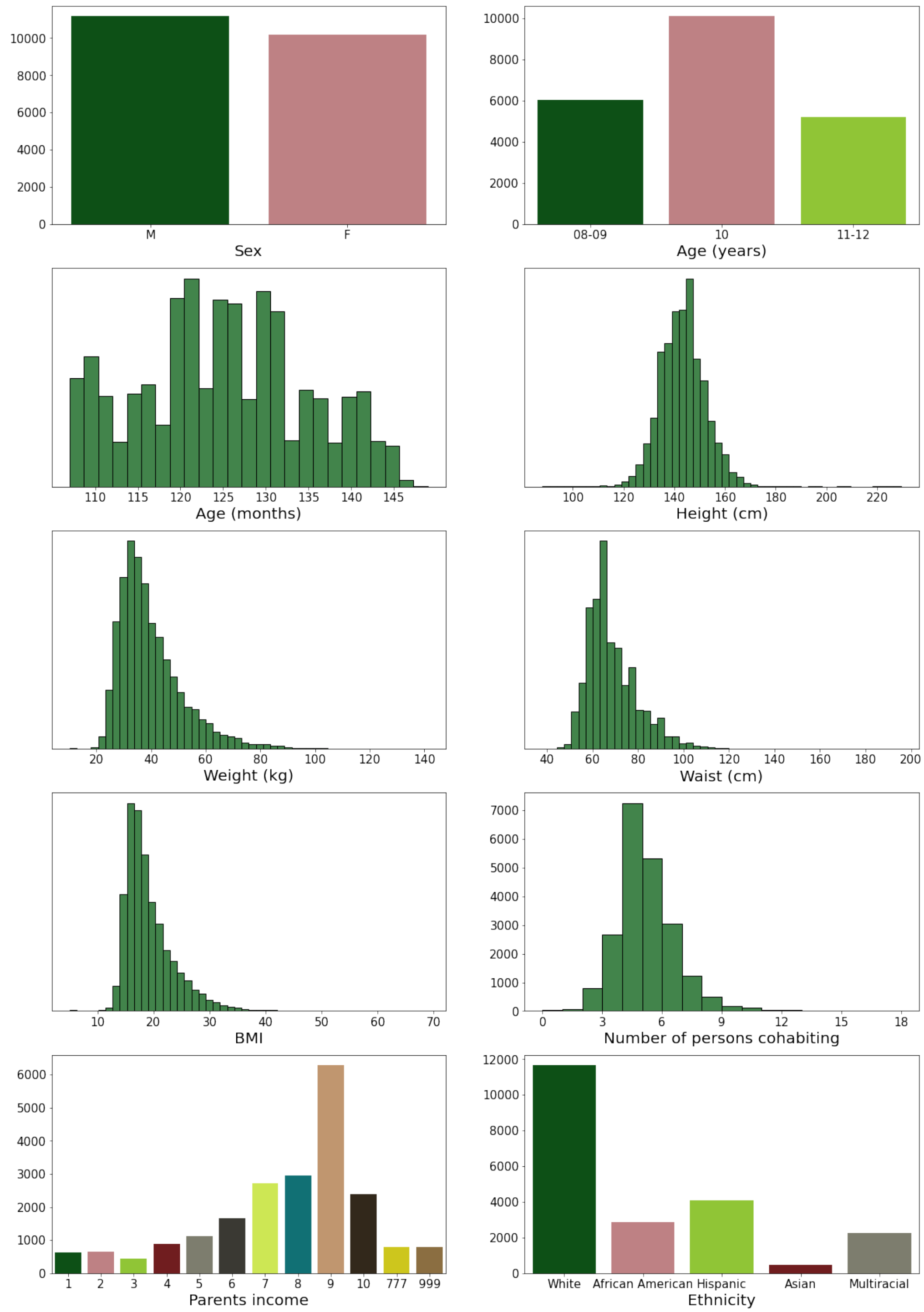


Figure 1.4: Distributions of the demographic features. The bar plot for "Parents income" represent the classes of yearly combined income of children's parents, from 1 (< 5000$) to 10 (> 200,000$). Class 9, which is the most frequent, corresponds to 100,000$ – 200,000$, whereas classes 777 and 999 represent the answers *Don't know* and *Refuses to answer*.

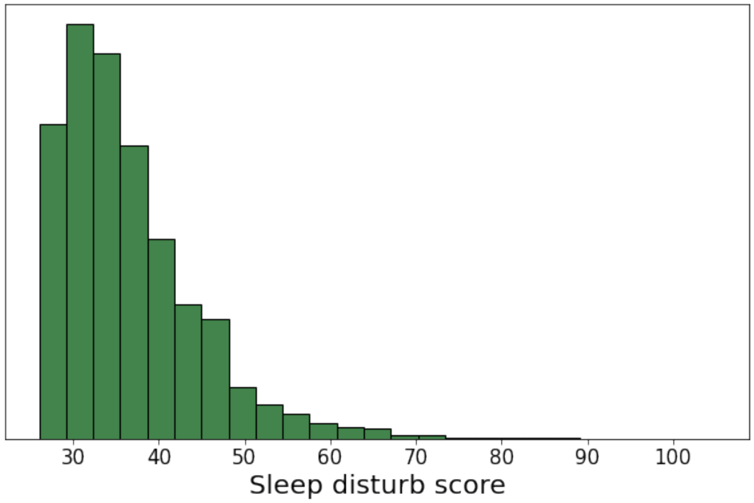


Figure 1.5: Distribution of total sleep disturb score

The predictor variables were chosen in order to take account of the main factors which can affect the targets defined in the previous section, in order to limit the confounding effects. They can be divided into three main groups: the demographic features, the scores from Sleep Disturbance Scale for Children and the daily screentime for different types of activity.

Demographic features

Demographic features help to obtain a background of the children in the dataset. They can be a useful tool to understand how the sample is composed and which observations may represent outliers that can affect the performance of the models. Moreover, features like sex and age often have an effect on biological response variables, therefore they need to be taken into account to avoid confounding. The demographics selected for this study, coherently to what was done for similar works on ABCD [11][18], were sex, age, ethnicity, physiological measures (height, weight, waist dimension and BMI), parents income and number of people cohabiting. Figure 1.4 describes the distributions of the selected features: males and females are quite balanced ($M : 11167$, $F : 10169$), whereas *White* is by far the most common ethnicity with ($> 54\%$ of the samples). Most children live in wealthy families (higher classes of income are the most common) and cohabit on average with 4 people.

Sleep Disturbance Scale for Children (SDSC)

The association between poor sleep quality and mental health has been proven repeatedly in literature [11][19], therefore sleep disorders need to be considered through some metrics in the analysis. Sleep Disturbance Scale for Children (SDSC) [12] is a 26 items Likert-type rating that attempts a categorization of sleep disorders in children, where higher scores suggest more acute sleep disturbances. The first two questions investigate the duration of sleep of the child and the time needed to fall asleep, whereas all the other questions are divided in a 5-points scale with the following codification: 1 =

	sleep_disturb	DIMS	SBD	DA	SWTD	DOES	SHY
mean	36.55	11.84	3.73	3.39	8.11	7.06	2.41
std	7.99	3.75	1.20	0.83	2.53	2.49	1.12
min	26	7	3	3	6	5	2
25%	31	9	3	3	6	5	2
50%	35	11	3	3	7	6	2
75%	40	14	4	4	9	8	2
max	105	35	15	15	29	25	10

Table 1.3: Statistics of sleep disturb scores. The first column represents the total score for sleep disturb, which is the sum of all the scores described by the following eight columns, respectively disorders of initiating and maintaining sleep (DIMS), sleep breathing disorders (SBD), disorders of arousal (DA), sleep-Wake transition disorders (STWD), disorders of excessive somnolence (DOES) and sleep hyperhidrosis (SHY).

Never, 2 = Occasionally (once or twice per month or less), 3 = Sometimes (once or twice per week), 4 = Often (3 or 5 times per week), 5 = Always (daily). As well as giving an overall score the instrument uses six subcategories: disorders of initiating and maintaining sleep, sleep breathing disorders, disorders of arousal, sleep-Wake transition disorders, disorders of excessive somnolence and sleep hyperhidrosis. Figure 1.5 and Table 1.3 show that the minimum value of the total score is 26 since every answer gives at least 1 point, and that once again the histogram is asymmetrical towards low values.

Screentime activities

Recreational screentime was measured using the Youth Screen Time Survey (14 items). Children were asked to report the number of hours spent on six screentime activities, on a typical weekday and weekend day. The activities corresponded to watching television/movies, watching videos (e.g., YouTube), playing video games, texting, visiting social networking sites (e.g., Facebook, Instagram) and using video chat (e.g., Skype). They were also asked the total daily screentime on a weekday and on a weekend day. Children responded to each question using a 7-point Likert-type scale coded as follows: 0 = none, $0.25 \leq 30$min, $0.50 = 30$min, $1 = 1$h, $2 = 2$h, $3 = 3$h, and $4 =\geq 4$h. The average daily screentime was evaluated as a weighted sum between weekdays and weekend days:

$$((\text{weekday screentime} \times 5) + (\text{weekend day screentime} \times 2))/7.$$

Figure 1.6 shows the distribution of the total daily screentime. It is interesting to note that the healthy threshold advised by pediatrists for screentime activity corresponds to a maximum of two hours per day [16], which in this dataset is only the 30^{th} percentile. This means that 70% of the children spends more time than advised in front of a screen.

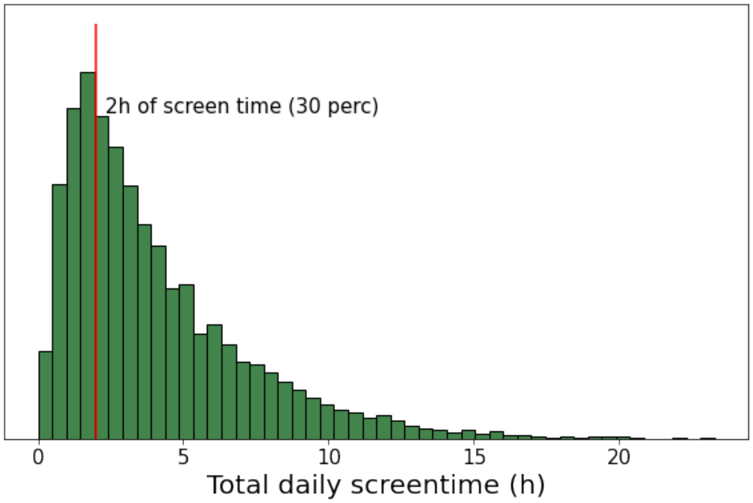


Figure 1.6: Distribution of total daily screentime.

	tot_screen_time	tv_time	video_time	videogames_time	texting_time	videochat_time	social_time
mean	4.07	1.24	1.06	1.08	0.30	0.23	0.16
std	3.25	1.03	1.16	1.14	0.61	0.55	0.50
min	0.00	0.00	0.00	0.00	0.00	0.00	0.00
25%	1.75	0.50	0.25	0.25	0.00	0.00	0.00
50%	3.11	1.00	0.64	0.64	0.00	0.00	0.00
75%	5.46	1.71	1.57	1.57	0.32	0.25	0.00
max	23.29	4.00	4.00	4.00	4.00	4.00	4.00

Table 1.4: Statistics of screentime features. Every column represent statistical details of the average daily screen time for different types of activities. The first column represents the sum of the time spent in all the activities.

Table 1.4, on the other hand, shows some basic statistics about the different screentime activities of the children in the dataset. The most common activities are individual activities like watching TV, videos and playing video games, while social activities like texting, spending time on a social network and video-chatting are much less frequent at this age. It may be interesting to sum the social activities in a new feature due to the low use of these behaviors, so that its values can be compared with the ones of the other activities.

Summing up, every sample represents an observation of a child in a single year, with the following features:

- **subjectkey**: a unique identifier for the subject;
- **eventname**: distinction between baseline, 1-year follow-up and 2-year follow-up;
- **age_months**: age in months of the subject;
- **age_years**: age in years of the subject;
- **sex**: sex of the subject;
- **parents_income**: combined parents income in past 12 months from all sources before taxes and deductions on a scale of 1 =< \$5000; 2 = \$5000 – \$11,999; 3 = \$12,000–\$15,999; 4 = \$16,000–\$24,999; 5 = \$25,000–\$34,999; 6 = \$35,000–\$49,999; 7 = \$50,000 – \$74,999; 8 = \$75,000 – \$99,999 ; 9 = \$100,000 – \$199,999 ; 10 => \$200,000 ; 777 = Don't know; 999 = Refuse to answer;
- **people_cohabiting**: number of people cohabiting with the subject;
- **height_cm**: height of the subject (cm);
- **weight_kg**: weight of the subject (kg);
- **waist_cm**: waist of the subject (cm);
- **bmi**: BMI index of the subject;
- **race_ethnicity**: ethnicity of the subject, divided in White, Black / African American, Hispanic, Asian, Other / Mixed;
- **physical_activity**: days per week when the subject performs more than one hour of intense physical activity;
- **Sleep disturb scores**: scores from the Sleep Disturbance Scale for children;
  - **DIMS**: Disorders of initiating and maintaining sleep;
  - **SBD**: Sleep breathing disorders;
  - **DA**: Disorders of arousal;
  - **SWTD**: Sleep-Wake transition disorders;
  - **DOES**: Disorders of excessive somnolence;
  - **SHY**: Sleep hyperhydrosis;
  - **sleep_disturb**: Sum of the previous scores;
- **CBCL scores**: scores from Child Behavior Checklist;
  - **anx_depr**: Anxious / Depressed;
  - **with_depr**: Withdrawn / Depressed;
  - **som_comp**: Somatic complaints;
  - **social_pr**: Social problems;
  - **thought_pr**: Thought problems;
  - **att_pr**: Attention problems;
  - **rule_br_bh**: Rule-breaking behavior;
  - **agg_bh**: Aggressive behavior;

  - **cbcl_score**: Sum of the previous scores;

- **Screentime activities**: screen time spent by the child, divided according to the activity (h);
  - **tv_time**: Time spent watching TV;
  - **video_time**: Time spent watching videos (e.g. YouTube);
  - **videogames_time**: Time spent playing video games;
  - **texting_time**: Time spent texting;
  - **social_time**: Time spent on social media;
  - **videochat_time**: Time spent on video chats;
  - **tot_screen_time**: Sum of the previous times.

The subset of samples with data from brain structure, on the other hand, adds the following features to previous ones:

- **ICV**: intracranial volume of the subject (mm^3);
- **LSurfArea**: left hemisphere cortical surface area of the subject (mm^2);
- **RSurfArea**: right hemisphere cortical surface area of the subject (mm^2);
- **LThickness**: left hemisphere cortical thickness of the subject (mm);
- **RThickness**: right hemisphere cortical thickness of the subject (mm);
- measures of cortical thickness and surface area of 70 regions from DKT atlas cortical parcellation of the subject's brain.

1.3 Preprocessing data

Data cleaning

The starting dataset was composed of 29703 samples and 38 features, with 11877 samples from the baseline observation, 11248 from the 1-year follow-up and 6578 from the 2-year follow-up. The first step in the cleaning process is removing the duplicates (N = 6). Since the items from screentime activity survey changed in the 2-year follow-up, the observations from this event were eliminated to keep more consistency in the analysis. Afterward all the rows containing missing values in at least one feature were eliminated, which did not produce a significant change in the dataset. The following step was performing outlier detection in the numerical features, especially in the demographics, in order to remove samples with absurd values. The accepted samples were chosen according to the following filters:

- 30 cm < **height** < 250 cm
- 10 kg < **weight** < 200 kg
- 5 < **BMI** < 100
- 20 cm < **waist** < 200 cm
- **people cohabiting** < 20
- **tot_screen_time** < 24 h

The filters were decided by looking at the boxplots of the samples' distribution and the usual demographic data of a child between 8 and 12. Only samples with impossible values (e.g., a child taller than the tallest man on Earth) were removed.

Finally, all the samples for which it was present only the baseline or the 1-year follow-up were removed. This allows studying the effects of baseline data on the follow-up. The resulting dataset has 21348 observations from 10674 children, whereas the subset with anatomical data has 1816 observations from 908 children.

Analysis of correlation

The analysis of correlation among the numerical features of the dataset highlights any linear relationships between the features, which is useful to understand how they affect one another and can be an index of multicollinearity in the models which are going to be fitted. The heatmap shown in Figure 1.7 describes the correlation between all the numerical features: the largest correlations are within the scales from the same survey, e.g. the scales from CBCL, within the different screentime activities and within physiological measures, e.g. height or weight. It is interesting to notice the correlation between mental health problems and sleep disturbs, which is an association that has been widely proven in the literature. A paper by Michel D. Guerrero [11], for example, showed that greater sleep duration predicts a $8.8-16.6\%$ decrease in problem behaviors by working with ABCD study. This is one of the reasons why the effect of poor sleep quality cannot be ignored when studying how screentime activity affects mental health.

Figure 1.8, on the other hand, shows how the brain measures correlate with the numerical features in the smaller dataset. There is almost perfect correlation between

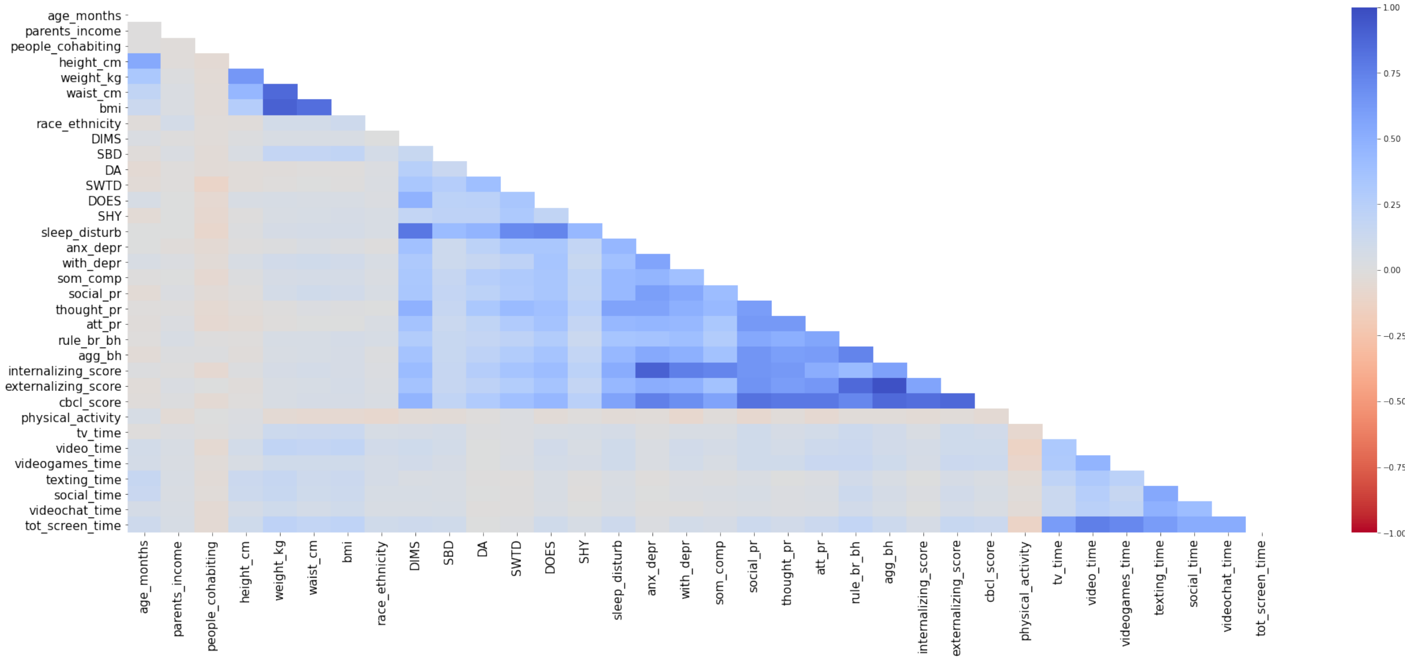


Figure 1.7: Heatmap of correlation among all the numerical features

left and right cortical measures, and there is also high correlation between surface area and intracranial volume, as expected. All the other features seem to have very low correlation, with the most significant values between 0.1 and 0.2, coming from physiological measures and the measures of screentime.

Feature selection

A high number of predictor variables, some of which are very similar or highly correlated, can lead to problems of overfitting or multicollinearity. Multicollinearity is a phenomenon where some predictor variables can be predicted by the others with a good degree of accuracy, which can lead to lower reliability of the coefficients estimates in the model. Small changes in the data would bring big changes in the individual predictors' coefficients, making any inference on the effects of the features meaningless. This problem can be monitored by evaluating each predictor's variance inflation factor (VIF) in the model, and it can be prevented with some precaution in the choice of the initial features to be fit, trying to remove any redundancies:

- *age_months* and *age_years* measure the same thing. The latter one can be useful for data visualization, but it brings less information, therefore it can be dropped

- *tot_screen_time* and *sleep_disturb* are linear combination of other predictors. It could be interesting to build models based on only these features, but in this case they can be dropped in order to analyse the effects of the different activities or sleep problems.

- the features *texting_time*, *social_time* and *videochat_time* have much lower values than the other screentime activities, and they are correlated since they all

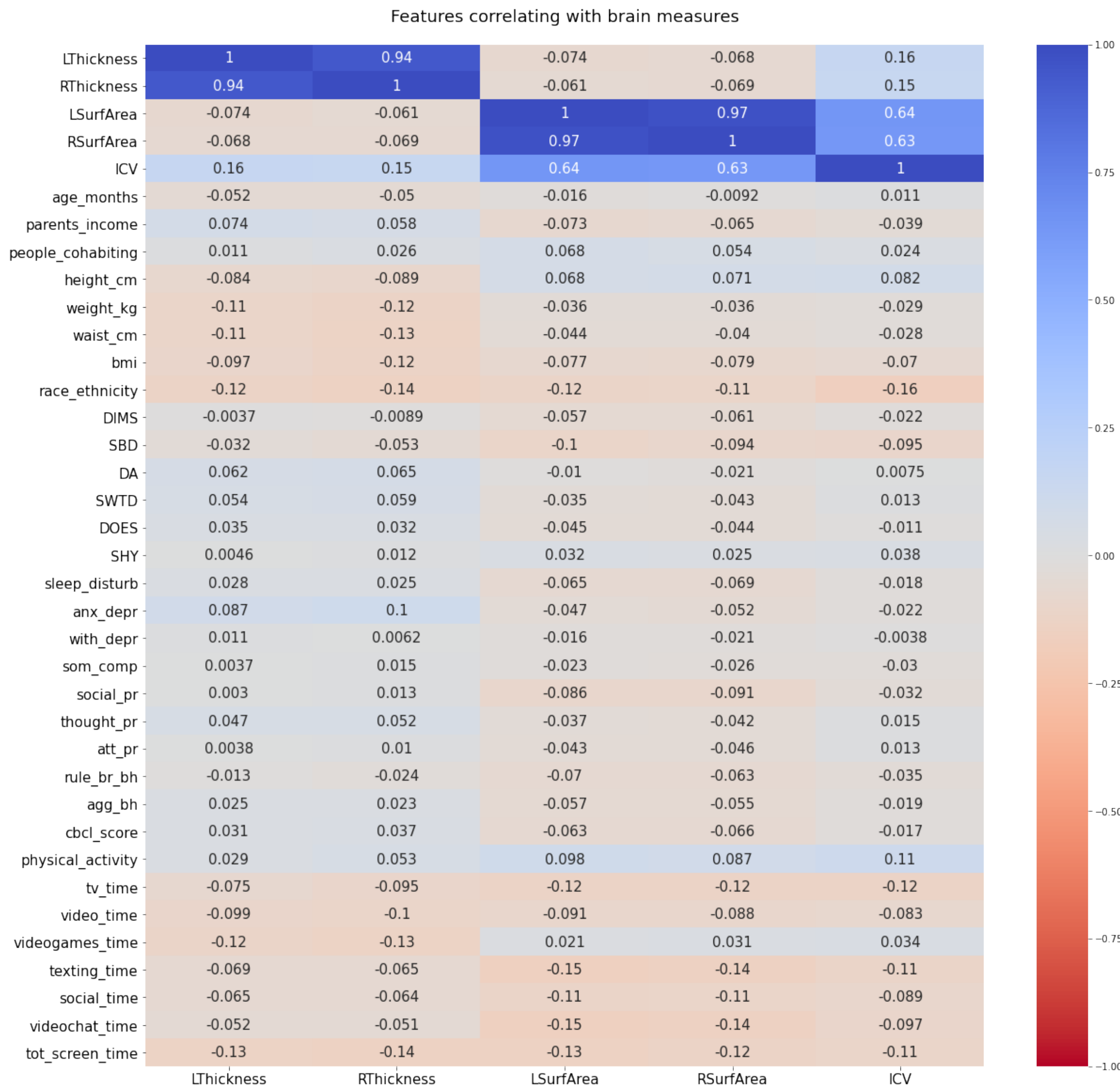


Figure 1.8: Correlation of the numerical features with the main brain measures

belong to the social sphere of the child, therefore they can be summed up into a new feature *social_activities_time*;

- the physiological measures *height_cm*, *weight_kg*, *bmi* and *waist_cm* are highly correlated and in some first trials they proved to be the ones with the highest VIF for every model, therefore they need to be transformed into new independent features which bring approximately the same information. This can be done through Principal Component Analysis (PCA), as shown in the next section.

Transformation of physiological features: PCA

The Principal Component Analysis (PCA) seeks the most accurate data representation of a group of features in a lower-dimensional space. The principal components are the orthogonal directions of the largest variance along which one can project the data. Applying PCA means computing the eigenvalues of the correlation matrix of the data,

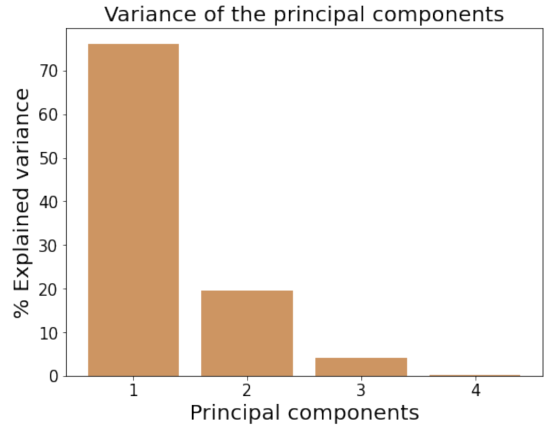


Figure 1.9: Variance explained by the principal components in decreasing order.

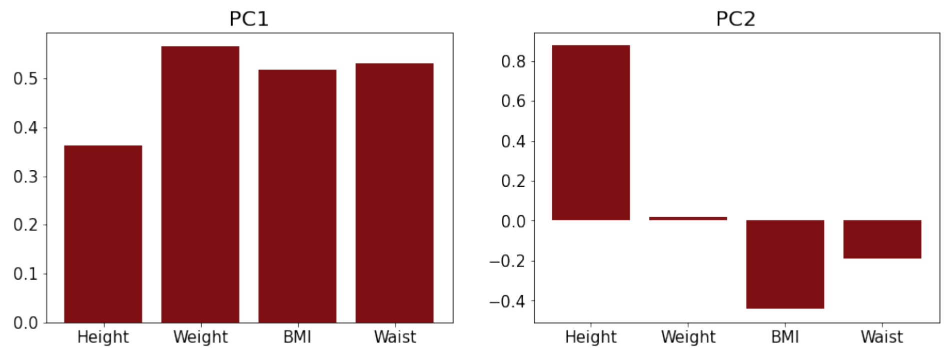


Figure 1.10: Loading scores of the first two principal components.

sorting them from the largest to the smallest, evaluating their eigenvectors and then projecting the data along the first k eigenvectors: those are the principal components, and represent the first k orthogonal directions of largest variance.
Before performing PCA it is advisable to scale the variables so that each one has standard deviations equal to 1 and to center them to have mean equal to 0 so that they have the same weight in the evaluation of the principal components. Figure 1.9 shows which percentage of the total variance of the original features is explained by the principal components. The first two components are sufficient to describe more than 90% of the original variance present in the physiological features.
An interpretation of what each component represents can be obtained analysing the loading scores of these components, i.e. the coefficients of the linear transformation from the original features to the new components. Figure 1.10 shows that the first component is a balanced combination of all the original features, with positive weights, therefore it probably describes the general *size* of the child. The second component, on the other hand, increases when the child is higher and decreases when it is thinner. For this reason it was called *slenderness*. These new two features have mean 0, standard deviation (σ) < 2 and can replace efficiently the original physiological features since they are totally independent.

Chapter 2

Models and methods

Different target variables require different statistical models to be analysed. This chapter describes the main methods used to model their values and to understand which covariates have a significant effect on them. The first section introduces the main regression models used, with a brief description of multiple linear regression models and generalized linear models. The second section describes the main evaluation metrics considered and the protocols of p-value correction for multiple testing.

2.1 Regression models

Multiple linear regression models

Multiple linear regression (MLR) is a statistical technique that uses several explanatory variables, or predictors, to predict the outcome of a response variable, which is then modeled as

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip} + \epsilon_i \tag{2.1}$$

where β_0 represents the *intercept*, β_j the *slope* coefficient for each predictor and ϵ_i the model's error term, also called *residual*. Relation (2.1) can be rewritten in matrix form:

$$\mathbf{y} = \mathbf{X}\beta + \epsilon$$

The estimates for the β_j coefficients can be found with the Least Squares method (LS), which allows to find the vector $\hat{\beta}$ which minimizes the sum of square errors $G(\beta)$ defined as

$$G(\beta) = \mathbf{e}'\mathbf{e} = (\mathbf{y} - \mathbf{X}\beta)'(\mathbf{y} - \mathbf{X}\beta).$$

The resulting vector can therefore be evaluated as

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

The LS estimators for linear regression are the most efficient unbiased estimators for Gauss-Markov theorem and coincide with the maximum-likelihood estimators.
Summing up, MLR is based on different assumptions:

- there is a linear relationship between the predictors and the response variable;

- the predictors are not too highly correlated with each other. When this assumption is not met, multicollinearity causes matrix X to be singular or nearly singular;
- y_i observations are selected independently and randomly from the population;
- residuals should be normally distributed with a mean of 0 and variance σ;

Generalized Linear Models (GLM)

Generalized linear models (GLM) are a generalization of the linear regression models that allow for the response variable to have a distribution other than the normal distribution. They are composed of three main parts: (i) the distribution of the response variable, which belongs to the exponential family (see Equation (2.2)), (ii) a linear predictor η_i defined in term of the explanatory features and (iii) a link function g(μ_i) between the mean μ_i of the response variable and the linear predictor.

$$f\left(y_i;\theta_i,\phi_i\right)=\exp\left(\frac{y_i\theta_i-b\left(\theta_i\right)}{a\left(\phi_i\right)}+c\left(y_i,\phi_i\right)\right) \tag{2.2}$$

A GLM can be fit to understand which are the factors that influence the scores in the CBCL survey, and therefore the mental health of the children. This could allow to highlight how screentime activity can affect the mental health independently from other triggering factors, for example the sleep quality.
According to the distribution of the response variable and the link function defined, various GLMs with different assumptions can be applied. When the response variable can take on real numeric values, for example with anatomical brain measures, a Gaussian distribution can be chosen with the identity link to obtain the simple linear regression. The scores of CBCL, though, are based on integer numbers (see Section 1.1) and all of the scales can only assume discrete non-negative values. This suggests the use of count models, such as those based on the Poisson family or on the Negative Binomial family.

Poisson model If the response variable has non-negative discrete values, it is reasonable to treat it as a count variable, and fit a Poisson model:

$$y_i\sim P\left(\lambda_i\right)\longrightarrow f\left(y_i;\lambda_i\right)=\frac{e^{-\lambda_i}{\lambda_i}^{y_i}}{y_i!}$$
$$\eta_i=\beta_0+\sum_{j=1}^{p}\beta_j x_{ij}$$
$$Link:\log\left(\mu_i\right)=\eta_i$$

Poisson regression models are generalized linear models with the Poisson distribution function as the assumed probability distribution of the response and usually have the logarithm link. A characteristic of the Poisson distribution is that its mean is equal to its variance. In certain circumstances, it will be found that the observed variance is greater than the mean; this is known as overdispersion. This can cause the presence of patterns in the residual plots and can be caused by high variance of the data or by the presence of too many zeros in the distribution. The Poisson model can be fit accounting for overdispersion.

Poisson model with overdispersion The difference between the Poisson GLM and the Poisson GLM with overdispersion is that the Poisson distribution is no longer explicitly specified, but only a relationship between the mean and variance of Y, which in this case are given by $\mathrm{E}(Y) = \mu$ and $\mathrm{var}(Y) = \phi \times \mu$. Although the Poisson distribution is not specified, the same type of model structure is used in terms of the link function and predictor function. If the dispersion parameter $\phi = 1$, the results are the same (in terms of estimated parameters and standard errors) as the Poisson GLM. If $\phi > 1$ there is overdispersion,. whereas if $\phi < 1$ there is underdispersion.

Negative binomial model Negative binomial regression is a popular generalization of Poisson regression because it loosens the highly restrictive assumption, made by the Poisson model, that the variance is equal to the mean. The traditional negative binomial regression model is based on the Poisson-gamma mixture distribution, i.e. it models the Poisson parameter with a gamma distribution. The negative binomial model has the following density function:

$$f(y; k, \mu) = \frac{\Gamma(y+k)}{\Gamma(k)\Gamma(y+1)} \left(\frac{k}{\mu+k}\right)^k \left(1 - \frac{k}{\mu+k}\right)^y$$

It has $E(Y) = \mu$ and $\mathrm{var}(Y) = \mu + \frac{\mu^2}{k}$, where the parameter k accounts for the overdispersion of the model. Its link function is the logarithm and it can be used when the overdispersion is too high to be modeled efficiently by the Poisson model with overdispersion.

Zero-inflated models Figure 1.1 represents the distribution of one of the response variables to be modeled., the cause of overdispersion in the model can probably be the high number of zeros, which seems to be unbalanced with respect to the rest of the distribution. In this case it can be caused by understatement of the question in the survey or by the subjectivity of the parents report, which is one of the main problems in psychometrics survey [17]. This problem can be dealt by using zero-inflated models, i. e. mixture models composed of a binomial part which measures the probability of finding a zero and a count part, which can be, for example, a Poisson or a Negative Binomial model, that counts the value to be predicted. The two models can have different coefficients and can be built on different features. After fitting a zero-inflated Poisson (ZIP) model and a zero-inflated negative binomial (ZINB) model, which accounts for overdispersion in two different ways, we can compare them by performing a χ^2 test on their respective likelihoods. Both ZIP and ZINB are characterized by:

- a binomial model which evaluates the probability π_i of finding zeros as:

$$\pi_i = \frac{e^{\nu+\gamma_1 Z_{i1}+\cdots\gamma_q Z_{iq}}}{1+e^{\nu+\gamma_1 Z_{i1}+\cdots\gamma_q Z_{iq}}} \tag{2.3}$$

  where ν is the intercept, γ_j are the coefficients and Z_{ij} are the covariates;

- a count model whose mean μ_i is evaluated as:

$$\mu_i = e^{\alpha+\beta_1 X_{i1}+\cdots+\beta_q X_{iq}}$$

  where α is the intercept, β_j are the coefficients and X_{ij} are the covariates, which are labelled differently since they may not be the same of Equation (2.3)

The expected value of a zero-inflated model can be therefore expressed as $E(Y_i) = \mu_i(1-\pi_i)$. The growth by 1 of a predictor j of the count model (or, in case of a categorical variable, the change from from 0 to 1) is therefore equivalent to multiplying the expected value by e^{β_j}. On the other hand, if the predictor is also present in the binomial model, the growth of γ_j can affect the slope of $\frac{\partial \pi_i}{\partial Z_{ij}}$ if the predictor is numerical, and the log odds ratio if it is categorical, meaning in few words that it would make the probability of finding zeros increase.

2.2 Model selection and validation

The choice of the optimal model can be performed at two different levels: first, the most significant features are selected for each linear model, then the different models can be compared through different metrics.
For each of the regression models previously discussed and for each of the response variables to be modeled the first step is to fit a model on all the features chosen in the previous chapter, then the least significant one is removed iteratively until all the remaining features are significant. The removal is performed through the use of a backward approach that tests for each feature if the deviance of the model (or the residual sum of squares in case of simple linear regression) is different from that of the nested model that does not consider that feature. In the case of linear regression, an F-test is performed on the difference between the RSS of the models, whereas in the case of GLMs the command performs χ^2-tests on the difference between the two deviances. The process is iterated until only significant features are left, i.e. when all features have p-value under the threshold $\alpha < 0.05$.

According to the model fitted, different metrics can be used to evaluate its performance:

- **Coefficient of determination $\mathbf{R^2}$**: $1 - \frac{RSS}{TSS} = 1 - \frac{\sum_i \epsilon_i^2}{\sum_i (y_i - \bar{y})^2}$, where RSS is the residual sum of squares and TSS is the total sum of squares. Can be used only when the predictors coefficients are calculated by ordinary least-squares regression;
- **Adjusted $\mathbf{R^2}$**: 1-$(1-R^2)\frac{n-1}{n-p-1}$, where n is the number of samples and p the number of predictors. Variant of R^2 which adjust it for the number of predictors;
- **Akaike information criterion(AIC)**: $2k - 2\ln(\hat{L})$, where k is the number of predictors and $\hat{L}$ is the maximum value of the likelihood function of the model.
- **Pearson correlation r**: $\frac{cov\left(\hat{Y},Y\right)}{\sigma_{\hat{Y}}\sigma_Y}$, where $cov\left(\hat{Y},Y\right)$ is the covariance between the predicted and the true output and $\sigma_Y/\sigma_{\hat{Y}}$ are their standard deviations;
- **Spearman correlation $\boldsymbol{\rho}$**: it corresponds to the Pearson correlation between the rank variables of the predicted and the true output;
- **Mean squared error (MSE)**: $\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2$, where n is the number of samples;
- **Mean absolute error (MAE)**: $\frac{1}{n}\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right|$;

The models described in the next chapter have been mainly chosen according to the AIC for GLMs and both AIC and Adjusted R^2 for MLRs, since they use a high a number of predictors and these metrics favor simpler models. The other metrics can help the choice when the values of AIC are similar or not available: this is the case of Poisson models with overdispersion, where no maximum likelihood method is defined, and therefore neither the AIC.

Multiple testing correction

In biological studies, it often happens that a set of hypotheses need to be tested simultaneously. This happens in genomic or in proteomic studies, but also in neuroscience, where many voxels or many regions of the brain need to be tested simultaneously for a particular effect, e.g. for being active in response to a particular task. In practice, to determine whether an observed score is deemed statistically significant, the corresponding statistical confidence measure (the p-value) must be compared to a confidence threshold α. However, the higher is the number of the tests, the higher is the probability of making at least one type I error, and therefore declaring a false positive, as shown in Figure 2.1. For this reason, we need a *multiple testing correction* procedure to adjust our statistical confidence measures based on the number of tests performed [20]. The most simple and common one is *Bonferroni correction*: if you are using a significance threshold of α, but you perform n separate tests, then the Bonferroni correction deems a score significant only if the corresponding p-value is $\leq \alpha/n$. Bonferroni correction though can be too strict, since it is an adjustment that ensures that the probability that one or more scores were drawn according to the null distribution is α. This kind of adjustments control what is called the *family-wise error rate*. Rather than saying that we want to be 95 sure that none of the observed scores is drawn according to the null, it is frequently sufficient to identify a set of scores for which a specified percentage of scores are drawn according to the null. This is the basis of multiple testing correction via *false discovery rate* (FDR) estimation. The FDR can be computed from the p-values using the Benjamini-Hochberg (BH) procedure, which works as follows:

1. For a given α, find the largest k such that $P_{(k)} \leq \frac{k}{m}\alpha$;
2. Reject the null hypothesis (i.e., declare discoveries) for all $H_{(i)}$ for $i = 1, \ldots, k$.

The BH procedure relies on the p-values being uniformly distributed under the null hypothesis and controls the FDR under a threshold α.

Multiple testing corrections were needed in this work to study the effects of screen-time activities on the 68 cortical regions obtained from the parcellation described in Section 1.1, which required 68 tests for each activity.

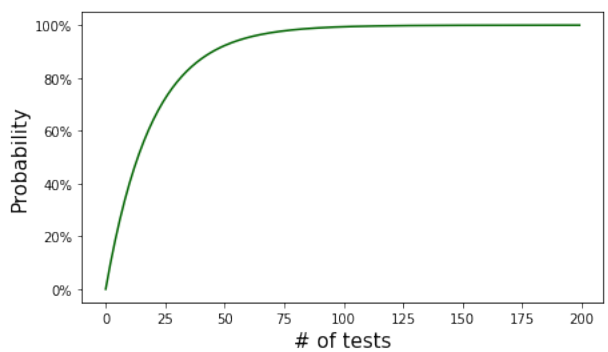


Figure 2.1: Probability of making at least one type I error with $\alpha = 0.05$ as threshold.

Chapter 3

Results of the models

This chapter presents the results obtained by fitting the models previously described on the target variables. Only the main effects will be discussed by reporting the estimates β of the coefficients and their confidence intervals (CI). A confidence interval is a range of values that's likely to include the effect coefficient with a certain degree of confidence, called confidence level. The confidence level used in this work is 95%. The tables containing all the model coefficients can be found in Appendix A.
The first section discusses the findings on associations with CBCL scales, with a subsection describing every scale and a final summary. The second section describes the associations with brain structure measures, with a focus on ROI-based analysis when significant effects are found. The third and final section presents further analysis on the causality of the effects previously found.

3.1 Results on CBCL scores

The first set of targets to be modeled are the scores from CBCL. The goal is to understand how the different predictors affect the scales and therefore the mental health of the children, with a focus on screentime activities.

CBCL score

The first score being modeled is the total score of CBCL, which from now on will be called CBCL score. It is based on summing all the answers to the 112 items in the survey, where respondents answered questions on a 3-point scale: not true (0), somewhat/sometimes true (1), or very true/often true (2). The variable has discrete values which can range from 0 to 224, even if the maximum in the studied samples is 122. Since it can only take on non-negative discrete values, it makes sense to treat it as a count variable, and therefore use count models like Poisson or Negative Binomial. These models represent an approximation since one of their hypotheses is that the response variable belongs to $\mathbb{N}$, whereas in this case the output is limited to [0,224], but they still are the most suitable for this kind of analysis. The table in Figure 3.1 shows how different count models performed in fitting the score with the best combination of features. Zero-Inflated Negative Binomial (ZINB) is the best according to AIC and Spearman correlation, and one the of best in MSE and MAE. The coefficients of the fitted model can be analysed only if no multicollinearity is present: as a rule of thumb,

Model	AIC	Pearson corr.	Sperman corr.	MSE	MAE	Df
Poisson	284994.595	0.594	0.558	405.706	13.199	32
Poisson with overdispersion	NaN	0.593	0.558	405.713	13.199	28
Negative Binomial	153610.938	0.598	0.568	404.427	13.197	29
Zero-Inflated Poisson	265472.698	0.552	0.567	177.710	9.231	45
Zero-Inflated Negative Binomial	152754.270	0.474	0.572	297.000	9.636	42

Table 3.1: Evaluation metrics for the different models for CBCL score

a VIF value that exceeds 5 or 10 indicates a problematic amount of collinearity [21]. None of the features' VIF exceeds 2, therefore there should not be multicollinearity.
The estimated coefficients are shown in Figure A.1 and Figure A.2 in Appendix A: they can be interpreted to obtain information about the different predictors. The categorical variables are treated in different ways: the β for *sex* feature weighs how the score changes between a female and a male, whereas the coefficients for *parents_income* weigh the contrast between an income class and the following one. If, for example, the coefficient of *parents_income 75-100 vs. 50-75* is positive, then there is a growth in the outcome variable when comparing the lower class with the higher one. Since *race_ethnicity* is a multi-categorical variable and it is not ordinal, all comparisons between the different ethnicities need to check, therefore it is more readable to sum them up in the table shown in Figure A.1.
According to the results, the score seems to drop slightly with age and physical activity and boys seem to obtain higher scores than the girls. The income of the family affects significantly the score only for high classes: it seems that the richer are the parents, the lower is the child's score. The number of people cohabiting did not affect the score significantly, whereas the child slenderness does ($\beta = -0.030$, $CI = [-0.044, -0.015]$). The ethnicities with the highest score are White and Other/Mixed, whereas the Asians have the lowest scores in the count model. On the other hand, being Asian seems to increase significantly the probability of finding a zero, in fact all the significant comparisons with other ethnicities led to negative values for γ. All the scores from Sleep Disturbance Scale affect positively the CBCL score, which confirms what has already been found in literature [11][19], and negatively the probability of finding zeros. This would mean that the more disturbed is a child's sleep, the less is probable that a score is equal to zero on the CBCL scale. Since screentime activities are the focus of this study, their effects on the CBCL score are highlighted by Table 3.2, which shows how the CBCL score is affected in percentage by the daily hours of screentime for each type of activity. The time spent in front of TV has the largest effect, followed by time spent watching videos and playing video games. Social activities involving screentime do not seem to be significant for this score.

The origin of these percentages is shown through an example: we consider the feature *tv_time*, whose coefficient $\beta \approx 0.021$ in this model. Since π_i is not significantly affected by *tv_time*, we can write

$$E(Y_i) = e^{\alpha+\beta_1 X_{i1}+\cdots+\beta_q X_{iq}+0.021 tv\_time}(1-\pi_i)$$

Activity	1 hour	2 hours	3 hours	4 hours
Tv	2.1%	4.4%	6.7%	9.0%
Video	1.8%	3.7%	5.5%	7.4%
Video games	1.4%	2.8%	4.2%	5.7%
Social activities	*n.s.*	*n.s.*	*n.s.*	*n.s.*

Table 3.2: Percent change of CBCL total score according to the type of activity and the number of hours. If the activity doesn't affect the score significantly, it's noted as not significant (n.s.).

Activity	1 h	2 h	3 h	4 h
Tv	*n.s.*	*n.s.*	*n.s.*	*n.s.*
Video	*n.s.*	*n.s.*	*n.s.*	*n.s.*
Video games	*n.s.*	*n.s.*	*n.s.*	*n.s.*
Social activities	−1.6%	−3.2%	−4.7%	−6.2%

Table 3.3: Percent change of *anx_depr* score according to the type of activity and the number of hours.

Then:

$$E(Y_i) = e^{\alpha+\beta_1 X_{i1}+\cdots+\beta_q X_{iq}}(1-\pi_i)e^{0.021tv\_time} = E(Y_i)_{-tv\_time}e^{0.021tv\_time}$$

where $E(Y_i)_{-tv\_time}$ corresponds to the expected value without considering the feature tv_time. If for instance we have $tv\_time = 2$ then $e^{0.0212} \sim 1.044$, from which a percentage increase of 4.4% with respect to the case with $tv\_time = 0$.
All the previous considerations on the model choice have been checked for all the following scales from CBCL, and they always led to the choice of ZINB as optimal model. Since the results are very similar, only the effects of the predictors will be discussed in the following sections.

Problems of anxiety / depression

The *anx_depr* scale is evaluated as the sum to the answers to the following items:

- Cries a lot
- Fears certain animals, situations, or places, other than school
- Fears going to school
- Fears he/she might think or do something bad
- Feels he/she has to be perfect
- Feels or complains that no one loves him/her
- Feels worthless or inferior
- Nervous, highstrung, or tense
- Too fearful or anxious
- Feels too guilty
- Self-conscious or easily embarassed
- Talks about killing himself
- Worries

It ranges from 0 to 26 and evaluates the anxiety of the children and their self-insecurity. No features are significant for the count model other than the sleep disturb scale, the physical activity, the social activities screentime and some comparison in the ethnicities. It is interesting to notice that Sleep breathing disorders (SBD) is the

Activity	1 hour	2 hours	3 hours	4 hours
Tv	$n.s.$	$n.s.$	$n.s.$	$n.s.$
Video	4.1%	8.4%	12.8%	17.5%
Videogames	3.5%	7.0%	10.7%	14.6%
Social activities	−4.9%	−9.6%	−14.0%	−18.3%

Table 3.4: Percent change of *with_depr* score according to the type of activity and the number of hours.

only sleep disturb which does not affect the score significantly in the count model and that physical activity seems to lower significantly the level of anxiety of the children ($\beta = -0.011$, $CI = [-0.017, -0.004]$). Being Black/Afr.Amer. or Asian contribute to lower the count model and increase the probability of obtaining zero as result. The zero model shows that being a male increases the probability of finding zeros ($\gamma = 0.239$, $CI = [0.092, 0.386]$), as well as cohabiting with more people.
Concerning screentime activities, Table 3.3 shows that the only ones that have a significant effect are the social activities, which seem to lower significantly ($\beta = -0.016$, $CI = [-0.031, -0.001]$) the anxiety of the child. This makes sense, since social relations can help with feelings of depression and loneliness [22].

Problems of withdrawal / depression

The *with_depr* scale is evaluated as the sum to the answers to the following items:

- There is very little he/she enjoys
- Would rather be alone than with others
- Refuses to talk
- Secretive, keeps things to self
- Too shy or timid
- Underactive, slow moving, or lacks energy
- Unhappy, sad, or depressed
- Withdrawn, doesn't get involved with others

It ranges from 0 to 16 and evaluates the social withdrawal of the children and how their shyness can cause them depression. It is interesting to notice in the model coefficients that size plays a role both in the count and in the zero model, by raising both the values. Age and sex affect significantly the score of the count model: older kids obtain on average higher scores, as well as males in comparison to females ($\beta = 0.083$, $CI = [0.040, 0.127]$). Ethnicity does not have significant effects for the zero model and has few significant effects in the count model, which shows low differences between the different ethnicities. The sleep scales and physical activity behave similarly to the previous cases.
Concerning screentime activities, Table 3.4 shows that different activities can have opposite effects: videos and video games rise significantly the feeling of withdrawal of the child ($\beta_{video} = 0.040$, $CI_{video} = [0.020, 0.060]$ and $\beta_{videogames} = 0.034$, $CI_{videogames} = [0.014, 0.054]$), whereas social activities, even if on screen, seem to help by lowering the score ($\beta = -0.050$, $CI = [-0.073, -0.028]$).

Activity	1 hour	2 hours	3 hours	4 hours
Tv	$n.s.$	$n.s.$	$n.s.$	$n.s.$
Video	2.3%	4.7%	7.2%	9.7%
Videogames	$n.s.$	$n.s.$	$n.s.$	$n.s.$
Social activities	$n.s.$	$n.s.$	$n.s.$	$n.s.$

Table 3.5: Percent change of *som_comp* score according to the type of activity and the number of hours.

Somatic complaints

The *som_comp* scale is evaluated as the sum to the answers to the following items:

- Physical problems **without known medical cause**:
  - Aches or pains (not stomach or headaches)
  - Headaches
  - Nausea, feels sick
  - Problems with eyes (not if corrected by glasses)
  - Rashes or other skin problems
  - Stomachaches
  - Vomiting, throwing up
- Nightmares
- Constipated, doesn't move bowels
- Feels dizzy or lightheaded
- Overtired without good reason

It ranges from 0 to 22 and evaluates the abnormal somatic complaints of the child. By looking at the model coefficients it stands out that males have lower scores than females on average ($\beta = -0.141$, $CI = [-0.172, -0.108]$) and that physical features like size and slenderness are significant: since size increases the score and slenderness lowers it, we can think that being fitter can indeed help with this kind of somatic problems. The number of people cohabiting with the child seems to be significant both in the count and in the zero model, by lowering the score and increasing the probability of finding zeros. Disorder of arousal (DA) is the highest between the coefficients of sleep disturb ($\beta = 0.088$, $CI = [0.0700.107]$). This makes sense since it is the scale linked to nightmares and sleepwalking. Once again Black/Afr.Amer. and Asians have lower scores compared to the other ethnicities, wheres in the zero model the higher effects come from Black/Afr.Amer. and Hispanics.
Concerning screentime activities, Table 3.5 shows that the only significant effects come from the time spent watching videos (e.g. YouTube), which slightly increases the score ($\beta = 0.023$, $CI = [0.009, 0.037]$.

Social problems

The *social_pr* scale is evaluated as the sum to the answers to the following items:

Activity	1 hour	2 hours	3 hours	4 hours
Tv	5.0%	10.3%	15.9%	21.8%
Video	*n.s.*	*n.s.*	*n.s.*	*n.s.*
Videogames	1.8%	3.7%	5.5%	7.5%
Social activities	*n.s.*	*n.s.*	*n.s.*	*n.s.*

Table 3.6: Percent change of *social_pr* score according to the type of activity and the number of hours.

Activity	1 hour	2 hours	3 hours	4 hours
Tv	3.4%	6.9%	10.5%	14.2%
Video	*n.s.*	*n.s.*	*n.s.*	*n.s.*
Videogames	3.0%	6.2%	9.4%	12.7%
Social activities	*n.s.*	*n.s.*	*n.s.*	*n.s.*

Table 3.7: Percent change of *thought_pr* score according to the type of activity and the number of hours.

- Clings to adults or too dependent
- Complains of loneliness
- Doesn't get along with other kids
- Easily jealous
- Feels others are out to get him/her
- Gets hurt a lot, accident-prone
- Gets teased a lot
- Not liked by other kids
- Poorly coordinated or clumsy
- Prefers being with younger kids
- Speech problem

It ranges from 0 to 22 and evaluates the problems in social relationships of the child. By looking at the model coefficients it is interesting to notice that age increases significantly the probability of finding zeros ($\gamma = 0.014$, $CI = [0.0070.022]$) and that higher classes of parents income bring on average to lower scores. Males have higher scores than females on average and the only ethnicity with significant differences is Asian, which is once again the one with the lowest scores. The effects of physical features and sleep scales are coherent with the previous cases.
Concerning screentime activities, we can see from Table 3.6 that the time spent watching TV increases significantly the score, up to a $\approx 5\%$ increase for every hour ($\beta = 0.049$, $CI = [0.032, 0.066]$). Video games also contribute to increasing the score, even if to a lesser extent ($\beta = 0.018$, $CI = [0.002, 0.034]$).

Thought problems

The *thought_pr* scale is evaluated as the sum to the answers to the following items:

Activity	1 hour	2 hours	3 hours	4 hours
Tv	3.0%	6.1%	9.3%	12.6%
Video	3.0%	6.0%	9.1%	12.3%
Videogames	2.6%	5.4%	8.1%	11.0%
Social activities	*n.s.*	*n.s.*	*n.s.*	*n.s.*

Table 3.8: Percent change of *att_pr* score according to the type of activity and the number of hours.

- Can't get his/her mind off certain thoughts / obsessions
- Deliberately harms self or attempts suicide
- Hears sound or voices that aren't there
- Nervous movements or twitching
- Picks nose, skin, or other parts of body
- Plays with own sex parts in public
- Plays with own sex parts too much
- Repeats certain acts over and over / compulsions
- Sees things that aren't there
- Sleeps less than most kids
- Stores up too many things he/she doesn't need
- Strange behavior
- Strange ideas
- Talks or walks in sleep
- Trouble sleeping

It ranges from 0 to 28 and evaluates the mental problems of the child, like schizophrenia, OCD or troubles sleeping. By looking at the model coefficients it stands out that the sleep disturbs scales, most of all DIMS, DA and SWTD, have higher effects when compared to the other scales of CBCL. This makes sense since *thoght_pr* is a scale linked to sleep problems. Males have once again higher scores than females on average, and the number of people cohabiting affect positively the probability of finding zeros ($\gamma = 0.073$, $CI = [0.0320.114]$). White and Other/Mixed are the ethnicities with the highest scores, which may also be caused by a lower probability in the zero model compared with the other ethnicities.
Concerning screentime activities, the results are similar to those of the previous scale, with only TV and video games screentime affecting positively the score.

Attention problems

The *att_pr* scale is evaluated as the sum to the answers to the following items:

- Acts too young for his/her age
- Fails to finish things he/she starts
- Can't concentrate, can't pay attention for long
- Can't sit still, restless, or hyperactive
- Confused or seems to be in a fog
- Daydreams or gets lost in his/her thoughts
- Impulsive or acts without thinking
- Poor school work
- Inattentive or easily distracted
- Stares blankly

It ranges from 0 to 20 and evaluates problems of attention, hyperactivity and impulsivity of the child. The effect of the covariates is coherent with most of the previous cases for what concern sleep disturbs, physical activity and sex. The number of people

Activity	1 hour	2 hours	3 hours	4 hours
Tv	7.1%	14.7%	22.8%	31.5%
Video	*n.s.*	*n.s.*	*n.s.*	*n.s.*
Videogames	*n.s.*	*n.s.*	*n.s.*	*n.s.*
Social activities	8.8%	18.3%	28.7%	40.0%

Table 3.9: Percent change of *rule_br_bh* score according to the type of activity and the number of hours.

cohabiting, size and slenderness affect negatively the score, and Asians have on average the lowest scores and the highest probability of finding zeros.
Concerning screentime activities, social activities are the only ones which do not affect significantly the score, whereas all the others make it increase. These effects, even if smaller than in other scales, must not be underestimated since screentime activity is usually quite heterogeneous, therefore the three different effects can sum up and be associated with bigger increases in the score.

Rule-breaking behavior

The *rule_br_bh* scale is evaluated as the sum to the answers to the following items:

- Drinks alcohol without parents' approval
- Doesn't seem to feel guilty after misbehaving
- Breaks rules at home, school, or elsewhere
- Hangs around with others who get in trouble
- Lying or cheating
- Prefers being with older kids
- Runs away from home
- Sets fires
- Sexual problems
- Steals at home
- Steals outside the home
- Swearing or obscene language
- Thinks about sex too much
- Smokes, chews, or sniffs tobacco
- Truancy, skips school
- Uses drugs for nonmedical purposes (don't include alcohol or tobacco)
- Vandalism

It ranges from 0 to 34 and evaluates problems of rule-breaking behavior and rebellion of the child. By looking at the model coefficients it is interesting to notice that parents' income affects negatively the score in the most of the steps between high classes. The sleep disturb scales increase the score, whereas physical activity does not seem to affect it significantly; Males have higher scores than females on average ($\beta = 0.300$, $CI = [0.2550.347]$). There are no large differences between ethnicities in the count model, whereas in the zero one Asians have by far the highest probability of finding zeros.
Concerning screentime activities, we can see that watching TV and especially social activities increase significantly the score ($\beta_{TV} = 0.068$, $CI_{TV} = [0.051, 0.086]$, $\beta_{social} = 0.084$, $CI_{social} = [0.065, 0.103]$): with just 4 hours of one of these two activities the score can rise up to 40% more than the same score without that activity.

Activity	1 hour	2 hours	3 hours	4 hours
Tv	3.4%	6.8%	10.4%	14.1%
Video	*n.s.*	*n.s.*	*n.s.*	*n.s.*
Videogames	*n.s.*	*n.s.*	*n.s.*	*n.s.*
Social activities	3.1%	6.2%	9.5%	12.8%

Table 3.10: Percent change of *agg_bh* score according to the type of activity and the number of hours.

Aggressive behavior

The *agg_bh* scale is evaluated as the sum to the answers to the following items:

- Argues a lot
- Cruelty, bullying, or meanness to others
- Demands a lot of attention
- Destroys his/her own things
- Destroys things belonging to his/her family or others
- Disobedient at home
- Disobedient at school
- Gets in many fights
- Physically attacks people
- Screams a lot
- Stubborn, sullen, or irritable
- Sudden changes in mood or feelings
- Sulks a lot
- Suspicious
- Teases a lot
- Temper tantrums or hot temper
- Threatens people
- Unusually loud

It ranges from 0 to 36 and evaluates problems of aggressiveness and anger of the child. By looking at the model coefficients it is interesting to notice that Disorders of Excessive Somnolence (DOES) increases the score with a larger effect than the previous results ($\beta = 0.058$, $CI = [0.0510.065]$). The other sleep disturb scales on the other hand are coherent with other scales, as well as sex and parents' income. Size, slenderness and number of people cohabiting are significant but seem to produce minor effects, whereas physical activity is not significant. The White ethnicity has the highest scores on average, whereas the Asian the lowest.
Concerning screentime activities, the results shown in Table 3.10 are similar to the ones of rule-breaking behavior: the only significant effects are produced by TV time and social activities, even it is much smaller than the previous case ($\beta_{TV} = 0.033$, $CI_{TV} = [0.018, 0.048]$, $\beta_{social} = 0.030$, $CI_{social} = [0.013, 0.047]$).

Summary and discussion

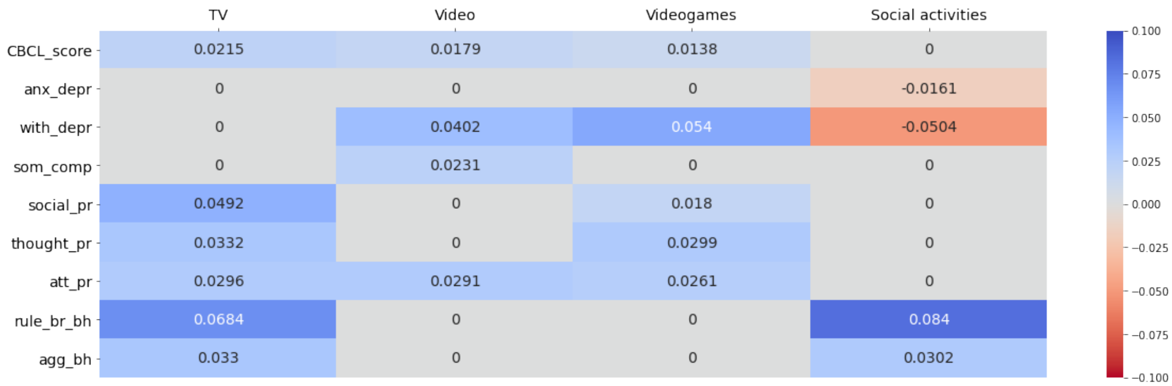


Figure 3.1: βs of the different screentime activities in the best models for each score from . The cells are color-coded from red to blue according to the value of the effect.

Summing up the results of the previous models, evidence shows that for what concerns the covariates males have on average higher scores, apart from somatic complaints, just like the White ethnicity when compared to the others, whereas the Asian is usually the one with the lowest scores. The scores from SDSC have always positive coefficients, which means that sleep disturbs tend to increase the scores of CBCL, whereas physical activity has usually the opposite effect. The number of people cohabiting and the parents' income are rarely significant, but when comparing further classes of income (the models previously shown only tested the contrast between a class and the following one) the results show that the children from richer families tend to have lower scores. The effects of the different screentime activities are shown by the heatmap in Figure 3.1. The spent watching TV increases significantly the scores concerning externalizing problems, like rule-breaking behavior and aggressive behavior, whereas social activities have opposite effects according to the score considered. This can be interesting since according to the children problems different approaches can be advised to try improving their mental health. The larger effects are shown in the problems of withdrawal and rule-breaking behavior: the next chapter will investigate what happens on an anatomical level, which may shed some light on what causes these effects.

3.2 Results on brain measures

The second set of targets to be modeled are the brain measures of ICV, cortical thickness and cortical surface area. The goal is to understand how the different predictors affect brain structure through the use of multiple linear regression models. The choice of using this kind of models is a consequence of the response variables distribution: as opposed to the scores from CBCL, the anatomical measures take on real values, therefore it makes sense to use linear regression. Moreover, the residual plots of the models (see Figure 3.2) showed normality and homoschedasticity of the residuals and all the values of VIF were once again < 2, making it reasonable to assume low multicollinearity. In MLR the analysis of the model coefficients on the target is much simpler than ZINB GLMs: the coefficient value signifies how much the mean of the dependent variable changes given a one-unit shift in the independent variable while holding other predictors in the model constant. This property of holding the other variables constant is crucial because it allows assessing the effect of each variable in isolation from the others. The tables describing the models can be found in the Appendix.

ICV

The measures of ICV in the model are expressed in mm^3, and the coefficients in Table A.10 follow that scale, but for the sake of discussion the results will be presented according to cm^3 scale. The intracranial volume (ICV) is defined as the volume within the cranium, including the brain, meninges and cerebrospinal fluid and it is evaluated by measuring the volume of the voxels within the brain mask. The ICV, demarcating the maximum size of the brain, is commonly utilized as a biomarker, especially as a normalization measure for studies on regional brain volumes [23]. Recently, it was also found that the ICV correlates with several psychiatric conditions. For example, a small but significant correlation was found between schizophrenia disorder and a reduced ICV [24]. Similarly, the ICV was the only brain volumetric biomarker that shows a genetic correlation with ADHD [25]. Lower gray matter volumes were also found for major depressive disorders and anxiety disorders [26]. The ICV is also strongly correlated with sex and head size. Thus, females, on average, have a smaller ICV, and people with bigger heads have, on average, larger ICV [23]. The difference according to sex is also present in the dataset: males ICV measures on average 1542.075 cm^3 ($std = 157.699$), whereas the females' one measures 1410.661 cm^3 ($std = 153.549$). This is also reflected in the coefficient of the model, with $\beta = 125.615$ and $CI = [112.032, 139.198]$. The dataset does not include a measure for head size, but the general size of the children proved to be significant affecting positively the value of ICV ($\beta = 9.818$ and $CI = [5.619, 14.017]$). It is also interesting to notice that there are significant differences between the ethnicities, making it an important factor to be considered in the regression, which is coherent with the results in literature [27]. An interesting effect comes from age($\beta = -1.728$ $CI = [-2.557, -0.900]$) which seems to contradict previous finding literature, according to which the brain volume does not decrease until the fourth decade of life [23]. This model coefficient, on the other hand, would imply that on average every month the children's ICV decreases of ≈ 1.7 cm^3. By taking into account all these effects in the models, no significant effects were found concerning screentime activities. The optimal model scored $R^2 = 0.267$ and $Adj R^2 = 0.257$, meaning that more than a quarter of the

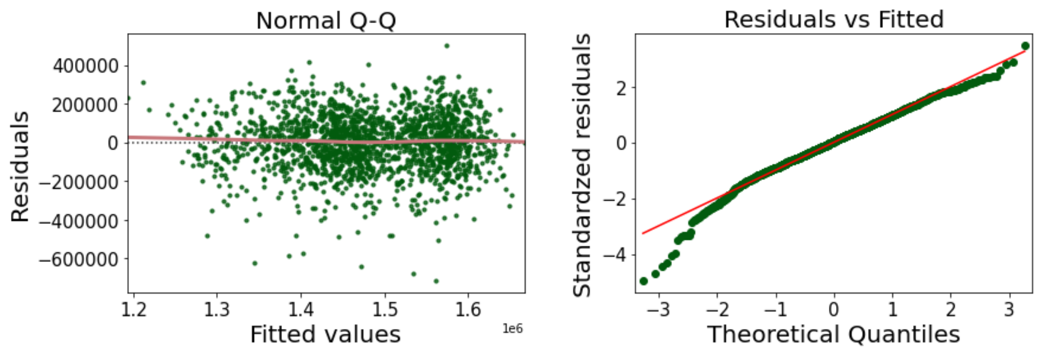


Figure 3.2: Residual plots of MLR model for ICV. The plot on the left is a scatter plot of residuals on the y-axis and fitted values (estimated responses) on the x-axis. The plot is used to detect non-linearity, unequal error variances, and outliers, which are not present in this case. The plot on the right, on the other hand, shows that hypothesis of normality of residuals is met by comparing the theoretical quantiles with the standardized residuals. The results from all the other models were very similar to these ones, shown as example.

variation in the output variable is predictable from the independent variables. The use of more features would probably increase this value, but it would also increase the risk of overfitting.

Cortical surface area

The cortical surface area is usually evaluated on the middle cortical surface, which lies at the geometric center between the inner and outer cortical surfaces. This provides a relatively unbiased representation of sulcal versus gyral regions. Table 1.2 seems to highlight a difference in the area of left and right hemispheres, but a Student's t-test showed that the difference between the means is not significant (p-value = 0.467). Similarly to the ICV it is strongly affected by the sex of the subject ($\beta_{left} = 7588.686$, $CI_{left} = [6801.597, 8375.774]$, $\beta_{right} = 8060.042$, $CI_{right} = [7266.781, 8853.302]$), with the males having an average of 98573.791 mm^2 ($std = 90611.68$) and the females having an average of 90611.683 mm^2 ($std = 8244.399$). Size affects positively the measure in both the lobes, as well as physical activity. A much bigger effect though comes from the slenderness of the child, with a coefficient 4 times larger than that of size ($\beta_{left} = 2017.942$, $CI_{left} = [1485.354, 2550.530]$, $\beta_{right} = 2043.441$, $CI_{right} = [1505.545, 2581.337]$).
The most interesting effects, and also the main difference between the two lobes, come from screentime activities: every hour of social activity online is associated on average with a decrease of -601.276 mm^2 ($CI = [-1001.299, -201.253]$) in left hemisphere cortical surface area, whereas in the right side a similar effect comes from watching videos ($\beta_{right} = 536.178$, $CI_{right} = [-859.513, -212.844]$). This would imply that 4 hours per day of these activities are associated with cortical surface area on average 2 cm^2 smaller. Similarly to the case of ICV, the optimal model for the left hemisphere scored $R^2 = 0.259$ and $AdjR^2 = 0.255$, whereas the optimal model for the right one scored $R^2 = 0.257$ and $AdjR^2 = 0.253$: it's interesting that the results are almost the same even if one of the predictors is completely different.

Region	Estimate	Std. Error	2.5 %	97.5 %	p-value	Adj p (Bonferroni)	Adj p (FDR)
lh_cuneus	-18.113	5.712	-29.316	-6.910	0.002	0.105	0.072
lh_lateraloccipital	-50.638	17.228	-84.427	-16.850	0.003	0.226	0.072
lh_pericalcarine	-15.485	6.406	-28.048	-2.921	0.016	1.000	0.214
lh_rostralmiddlefrontal	-43.136	21.218	-84.751	-1.521	0.042	1.000	0.261
rh_bankssts	-8.332	3.823	-15.830	-0.834	0.029	1.000	0.222
rh_cuneus	-17.556	5.988	-29.299	-5.812	0.003	0.232	0.072
rh_lateraloccipital	-38.310	17.139	-71.925	-4.695	0.026	1.000	0.222
rh_parsorbitalis	-6.441	2.821	-11.974	-0.908	0.023	1.000	0.222
rh_parstriangularis	-13.584	6.654	-26.633	-0.534	0.041	1.000	0.261
rh_pericalcarine	-14.859	6.746	-28.089	-1.629	0.028	1.000	0.222
rh_superiortemporal	-31.361	10.946	-52.829	-9.894	0.004	0.287	0.072

Table 3.11: Effects of *video_time* on regions of interest surface area with p-value < 0.05

Region	Estimate	Std. Error	2.5 %	97.5 %	p-value	Adj p (Bonferroni)	Adj p (FDR)
lh_caudalmiddlefrontal	-32.409	12.141	-56.220	-8.598	0.008	0.521	0.261
rh_caudalanteriorcingulate	-10.929	4.085	-18.942	-2.917	0.008	0.512	0.261
rh_superiorfrontal	-59.336	26.210	-110.741	-7.930	0.024	1.000	0.449

Table 3.12: Effects of *social_activities_time* on regions of interest surface area with p-value < 0.05

Regions of interest analysis The following step is to understand if the effects can be localized to some of the regions of interest of the cortex. A new MLR model was fitted for each of the 68 regions, and whenever any of the four screentime activities had an effect on the region surface area it was saved in a table. Tables 3.11 and 3.12 sum up all the regions where video or social activities screentime have an effect with p-value < 0.05. Due to the high number of tests, though, these effects need to be adjusted. The last two columns of the tables show the adjusted values with Bonferroni and FDR correction: none of the regions is significantly affected by screentime activities after either of the corrections.

Cortical thickness

The cortical thickness is defined as the average Euclidean distance between the inner and outer cortical surfaces. It is evaluated separately for both the hemispheres and for each of the regions of interest previously introduced. Only a few predictors affected significantly this measure: in both the hemispheres the size seems to lower slightly the outcome ($\beta = -0.003$, $CI_{left} = [-0.005, -0.001]$ and $CI_{left} = [-0.006, -0.001]$), and in the right one sex has an effect too, with the females having on average the cortex ≈ 0.01 mm thicker. This last effect is very small though, with p-value$= 0.045$.
The largest effect come from the video games screentime, which seem to be associated with a slight cortical thinning ($\beta = -0.006$, $CI_{left} = [-0.009, -0.003]$ and $CI_{right} = [-0.009, -0.002]$). This result was also confirmed by comparing with a Student's t-test the children who usually do not play video games ($videogames\_time = 0$) with children who play video games 4 or more hours per day ($videogames\_time = 4$): the difference between the population means is significant with p-value$= 2.508 \times 10^{-4}$. Research has produced mixed findings on this association [4][28]. The evaluation metrics for the optimal models are much lower than the previous cases: the best model for left hemisphere scored $R^2 = 0.053$ and $AdjR^2 = 0.0503$, whereas the best model for the right one scored $R^2 = 0.074$ and $AdjR^2 = 0.069$. This may be a consequence of the lower number of significant predictors found, which in this case explain a smaller proportion of the target variation.

Regions of interest analysis The analysis of local effects can help to understand which regions of the cortex associate significantly decrease thickness with video games screentime. The protocol was the same described in the previous section, with a new MLR model being fitted for each region. In 19 regions the estimate of $videogames\_time$ β for cortical thickness has p-value < 0.05, but after the adjustment only 7 have significant effects with FDR correction and 3 with Bonferroni correction. Table 3.13 sums up all the regions where video games screentime has an effect with p-value < 0.05 after FDR correction. 4 of these regions belong to the right hemisphere, whereas the remaining 3 belong to the left. Most of the regions are located in the temporal lobe, as shown in Figure 3.3. None of the other screentime activities has significant effects after either of the corrections. As a further step to strengthen the results, the same protocol was repeated after normalization of the local measure. The normalization was performed in two ways, dividing the value of each region's cortical thickness by the average thickness in the hemisphere and by the ICV. The models fitted after the normalization did not show any significant results after either of the corrections, for both kinds of normalization.

Region	Estimate	Std. Error	2.5 %	97.5 %	p-value	Adj p (Bonferroni)	Adj p (FDR)
lh_fusiform	-0.010	0.003	-0.015	-0.004	0.000	0.020	0.007
lh_inferiortemporal	-0.009	0.003	-0.015	-0.003	0.003	0.232	0.036
lh_transversetemporal	-0.013	0.004	-0.020	-0.005	0.001	0.058	0.014
rh_fusiform	-0.008	0.003	-0.014	-0.003	0.002	0.166	0.033
rh_middletemporal	-0.011	0.003	-0.017	-0.006	0.000	0.008	0.007
rh_precentral	-0.009	0.003	-0.014	-0.003	0.004	0.252	0.036
rh_superiortemporal	-0.011	0.003	-0.016	-0.005	0.000	0.021	0.007

Table 3.13: Effects of *videogames_time* on regions of interest with FDR adjusted p-value < 0.05

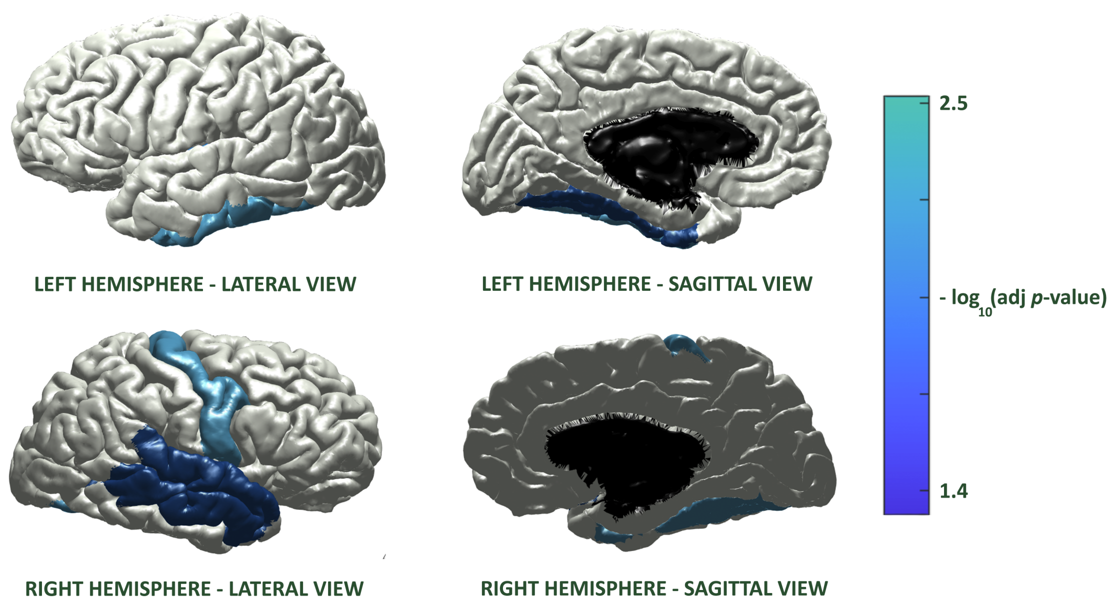


Figure 3.3: Parcellation of the brain with the regions of interest colored according to the FDR adjusted p-value of the effect of video games screentime on cortical thickness.

3.3 Analysis on the causality of the effects

	Diff_ICV	Diff_LSurfArea	Diff_RSurfArea	Diff_LThickness	Diff_RThickness
mean	-0.012	0.003	0.003	0.002	0.002
std	0.073	0.055	0.057	0.025	0.023
min	-0.355	-0.584	-0.408	-0.110	-0.080
25%	-0.041	-0.009	-0.008	-0.012	-0.011
50%	-0.014	0.009	0.009	0.001	0.001
75%	0.012	0.024	0.026	0.014	0.013
max	0.528	0.547	0.514	0.164	0.176

Table 3.14: Statistics of the percentage changes in brain measures between follow-up and baseline.

All the results obtained up until this point highlight associations between screentime activities and differences in brain measures, but none of them can prove the direction of causality in these effects. This would be impossible in a cross-sectional study since there would not be any temporal information. This dataset, on the other hand, provides information about two runs, a baseline and a follow-up. Even if many more time points would be required to highlight a trend that could prove some form of causality, it can be interesting to analyse whether some of the predictors from baseline data has an effect on the relative difference between the target variables from the follow-up and the ones from the baseline. Proving an association between screentime activities and eventual changes in the brain structure would give more credit to the hypothesis of causality. This kind of analysis could not be done for CBCL scores, since the difference between the scores of two following years is a discrete variable with both positive and negative values, which is much more difficult to model.
If BM_0 is the value of a brain measure from a subject in baseline run and BM_1 the value of the same subject's brain measure from follow-up run, then the target variable to be modeled is evaluated as $\frac{BM_1 - BM_0}{BM_0}$. The brain measures considered are once again ICV and cortical surface and thickness from both the hemispheres, and the new target variables represent the percentage change of these measures. Table 3.14 shows the main statistics of these new targets: the changes are very small, with ICV decreasing on average of $\approx 1.2\%$ and the other measures increasing of $0.2 - 0.3\%$. The variance of the features is also low: the changes in the first and third quartiles never exceed 5%. MLR models were fitted for each of these target variables, but none of them produced significant results. For each model an F-test was performed to test whether any of the independent variables could be significant, i.e. if any of the β is different from 0, but the null hypothesis was always accepted.

Conclusion

The starting question of this work was: how does screentime activity affect the mental health and the brain structure of the digital natives? Thanks to the data provided by the ABCD study, it was possible to show the effect that different kinds of activities have on some of the most common mental disorders and brain measures like ICV, cortical surface area and cortical thickness.
For what concerns the time spent watching television is was shown that it is positively associated with social, thought and attention problems, but the largest effects were observed with externalizing problems like rule-breaking and aggressive behaviors. The models on brain structure, on the other hand, did not highlight any significant effects of this activity on ICV, cortical surface area or cortical thickness.
The time spent watching videos was associated with increased scores in attention problems, somatic complaints and especially feelings of withdrawal and depression. Moreover, it was shown an effect on brain structure, with more hours spent in front of videos being associated with the right hemisphere cortical surface area.
Similarly, social activities screentime is associated with a decreased cortical surface area in the left hemisphere. For what concerns CBCL scores, social activities contribute to lowering the values of internalizing scores like feelings of anxiety and withdrawal and significantly increase externalizing problems like rule-breaking and aggressive behaviors. Previous findings in literature [5] show that irritability is associated with decreased cortical surface area, therefore there may be a connection between the two effects.
The most interesting results, though, come from video games activity. Kevin Dabbs et al. wrote a paper on the patterns between cortical thickness and CBCL scores [29] where they found out that higher scores across several CBCL syndrome scales including Withdrawn/ Depressed, Thought Problems and Attention Problems were associated with decreased cortical thickness. In particular, higher thought problems were located in right superior temporal gyrus and left transverse temporal region. The effects of video games screentime highlighted in this thesis overlap almost perfectly with these findings: video games activity was associated with increased scores especially in Withdrawn/ Depressed, Thought Problems and Attention Problems scales, as well as with decreased cortical thickness, especially in temporal regions which include the ones cited in the paper. The main limitation in this work is in its reliance on parents' report on sleep quality and problem behaviors as well as reliance on children to report their screentime behavior. It is possible that children underestimated their time spent on the different screen types or that parents misjudged their children's behaviors, therefore the data may be biased. Conducting longitudinal research using objective measures of screentime and sleep can confirm or refute the findings of this thesis. Another limitation is in the short time span considered: despite this work is not based on cross-sectional data,

more accurate findings may come out with longitudinal data throughout the children school age. A natural continuation of the study would focus more on the relationship between the changes in brain structure and CBCL scores in order to better understand the role of screentime activity. Moreover, it would be interesting to study the effects of different kinds of video game activity, as well as the effects of different platforms (e.g., Netflix, Youtube, Tik Tok...) and different types of screens (e.g. tablets, laptops and especially smartphones). Finally, thanks to the fMRI data provided by ABCD the children's resting state brain activity could be analysed to deepen the knowledge about how the different regions of interest interact with each other and therefore if and how the brain functions are affected by screentime activity.

Appendix A

Further materials

The third chapter of this thesis discussed the results highlighted by the optimal regression models used to predict the target of interest for this study. This section reports the models in detail, with each model being described by two tables. The first one contains all the significant estimates β of the predictors' coefficients and their confidence intervals, alongside the the values of empirical mean and standard deviation for the numerical variables, which allow to better understand the scale of the predictor effect. The second table, on the other hand, shows the effects of all the comparisons between the ethnicities: when the effect is not significant, and therefore there are no differences in the outcome between the two ethnicities, the cell will be noted as not significant (*n.s.*). For what concerns the CBCL scores, the optimal models are Zero-Inflated Negative Binomial and the output is predicted by a mixture of two models, therefore four tables will be needed, with two tables for the count model and two for the binomial model. On the other hand the brain measures are fitted through multiple linear regression models, therefore only two tables will be needed.

A.1 Models for CBCL scores

Feature	mean (σ)	β (count)	[2.5% 97.5%] (count)
Intercept		0.864	[0.674 1.054]
Age (months)	125.04 (9.73)	-0.002	[-0.003 -0.001]
Sex: M vs. F		0.164	[0.139 0.188]
Parents yearly income (x1000 $): 75-100 vs. 50-75		-0.060	[-0.104 -0.017]
Parents yearly income (x1000 $): 100-200 vs. 75-100		-0.043	[-0.080 -0.006]
Parents yearly income (x1000 $): >200 vs. 100-200		-0.070	[-0.110 -0.029]
Slenderness	0.00 (0.88)	-0.030	[-0.044 -0.015]
Disorders of initiating and maintaining sleep (DIMS)	11.84 (3.75)	0.061	[0.058 0.064]
Sleep Breathing Disorders (SBD)	3.74 (1.20)	0.017	[0.007 0.027]
Disorders of arousal (DA)	3.39 (0.83)	0.073	[0.058 0.088]
Sleep-Wake Transition Disorders (SWTD)	8.11 (2.53)	0.052	[0.047 0.058]
Disorders of excessive somnolence (DOES)	7.06 (2.49)	0.072	[0.066 0.077]
Sleep Hyperhydrosis (SHY)	2.41 (1.12)	0.036	[0.025 0.046]
Physical activity (days/week)	3.50 (2.18)	-0.013	[-0.019 -0.008]
Screentime: TV (h)	1.24 (1.03)	0.021	[0.010 0.033]
Screentime: Video (h)	1.06 (1.16)	0.018	[0.007 0.029]
Screentime: Videogames (h)	1.08 (1.14)	0.014	0.002 0.026]

Table A.1: Coefficients of the count part of ZINB model for CBCL score

Feature	mean (σ)	γ (zero)	[2.5% 97.5%] (zero)
Intercept		13.720	[10.787 16.652]
# people cohabiting	4.70 (1.52)	0.072	[0.017 0.127]
Disorders of initiating and maintaining sleep (DIMS)	11.84 (3.75)	-0.201	[-0.255 -0.148]
Sleep Breathing Disorders (SBD)	3.74 (1.20)	-0.185	[-0.340 -0.032]
Disorders of arousal (DA)	3.39 (0.83)	-1.325	[-2.096 -0.554]
Sleep-Wake Transition Disorders (SWTD)	8.11 (2.53)	-0.654	[-0.863 -0.444]
Disorders of excessive somnolence (DOES)	7.06 (2.49)	-0.869	[-1.099 -0.640]

Table A.2: Coefficients of the binomial part of ZINB model for CBCL score

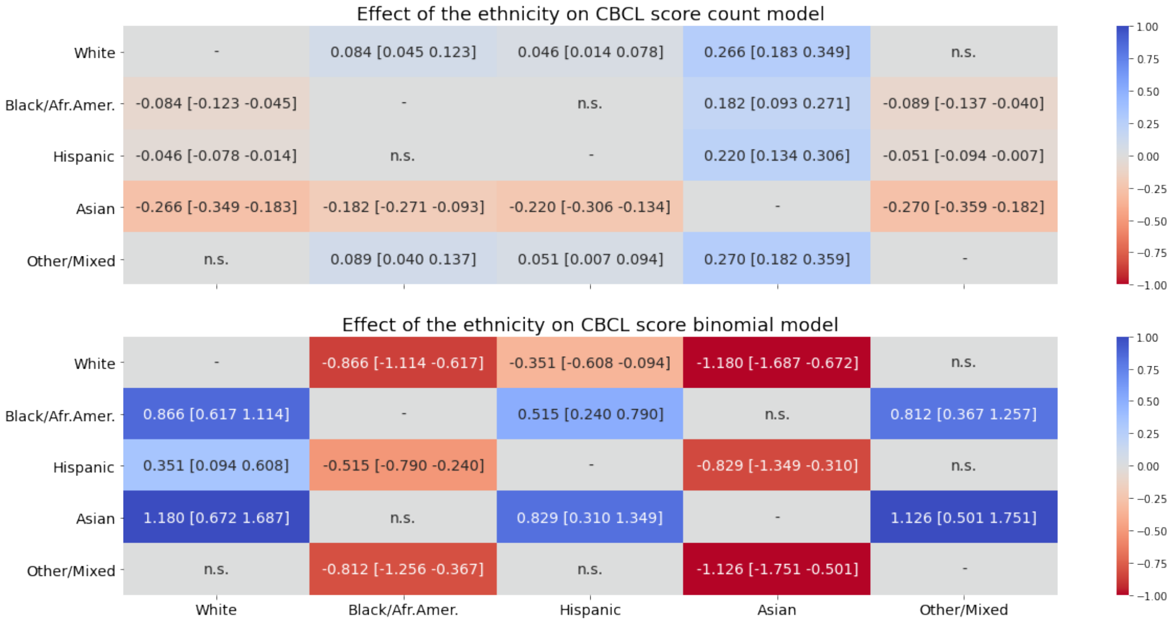


Figure A.1: Effect of the ethnicity on CBCL score

Feature	mean (σ)	β (count)	[2.5% 97.5%] (count)
Intercept		-0.511	[-0.597 -0.425]
Disorders of initiating and maintaining sleep (DIMS)	11.84 (3.75)	0.052	[0.048 0.057]
Disorders of arousal (DA)	3.39 (0.83)	0.047	[0.029 0.065]
Sleep-Wake Transition Disorders (SWTD)	8.11 (2.53)	0.048	[0.042 0.055]
Disorders of excessive somnolence (DOES)	7.06 (2.49)	0.042	[0.035 0.048]
Sleep Hyperhydrosis (SHY)	2.41 (1.12)	0.034	[0.021 0.046]
Physical activity (days/week)	3.50 (2.18)	-0.011	[-0.017 -0.004]
Screentime: Social activities (h)	0.60 (0.95)	-0.016	[-0.032 -0.001]

Table A.3: Coefficients of the count part of ZINB model for *anx_depr* score

Feature	mean (σ)	γ(zero)	[2.5% 97.5%] (zero)
Intercept		7.081	[5.889 8.273]
Sex: M vs. F		0.239	[0.092 0.386]
# people cohabiting	4.70 (1.52)	0.084	[0.040 0.128]
Disorders of initiating and maintaining sleep (DIMS)	11.84 (3.75)	-0.168	[-0.204 -0.131]
Sleep Breathing Disorders (SBD)	3.74 (1.20)	-0.270	[-0.400 -0.141]
Disorders of arousal (DA)	3.39 (0.83)	-0.455	[-0.719 -0.191]
Sleep-Wake Transition Disorders (SWTD)	8.11 (2.53)	-0.307	[-0.394 -0.220]
Disorders of excessive somnolence (DOES)	7.06 (2.49)	-0.442	[-0.549 -0.334]
Sleep Hyperhydrosis (SHY)	2.41 (1.12)	-0.246	[-0.471 -0.021]

Table A.4: Coefficients of the binomial part of ZINB model for *anx_depr* score

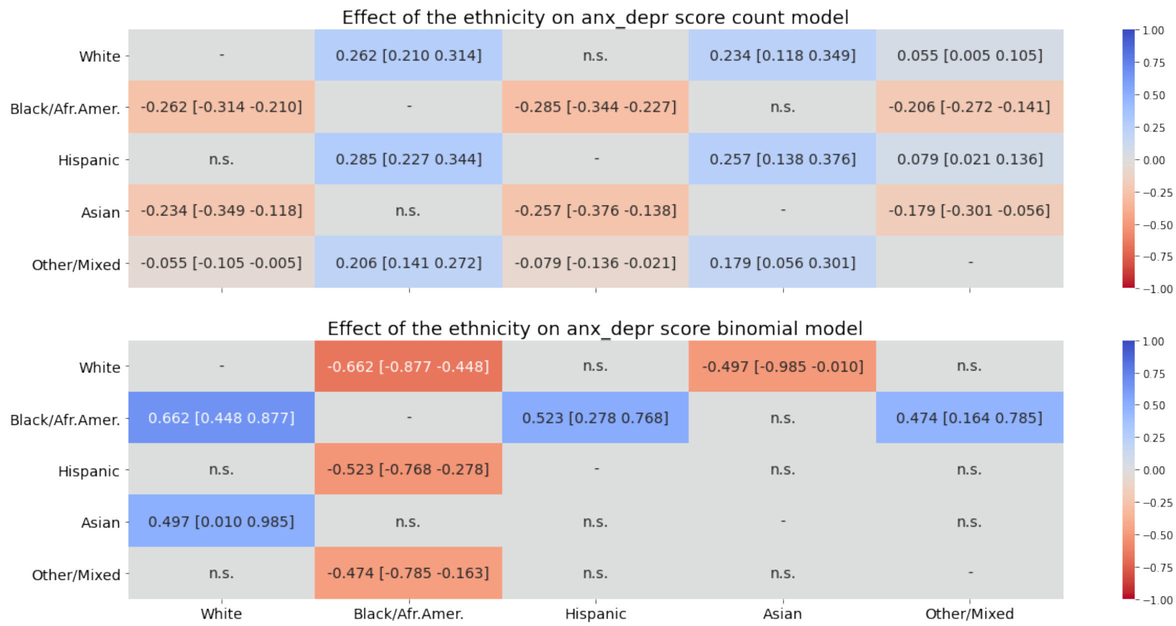


Figure A.2: Effect of the ethnicity on *anx_depr* score

Feature	mean (σ)	β (count)	[2.5% 97.5%] (count)
Intercept		-1.873	[-2.179 -1.567]
Age (months)	125.04 (9.73)	0.003	[0.001 0.005]
Sex: M vs. F		0.083	[0.039 0.127]
Size	0.00 (1.74)	0.052	[0.039 0.065]
Disorders of initiating and maintaining sleep (DIMS)	11.84 (3.75)	0.053	[0.047 0.059]
Disorders of arousal (DA)	3.39 (0.83)	0.042	[0.018 0.066]
Sleep-Wake Transition Disorders (SWTD)	8.11 (2.53)	0.020	[0.011 0.030]
Disorders of excessive somnolence (DOES)	7.06 (2.49)	0.077	[0.068 0.086]
Sleep Hyperhydrosis (SHY)	2.41 (1.12)	0.046	[0.029 0.062]
Physical activity (days/week)	3.50 (2.18)	-0.036	[-0.047 -0.025]
Screentime: Video (h)	1.06 (1.16)	0.040	[0.020 0.060]
Screentime: Videogames (h)	1.08 (1.14)	0.034	[0.014 0.054]
Screentime: Social activities (h)	0.60 (0.95)	-0.050	[-0.073 -0.028]

Table A.5: Coefficients of the count part of ZINB model for *with_depr* score

Feature	mean (σ)	γ (zero)	(2.5% 97.5%) (zero)
Intercept		4.259	[3.303 5.216]
Size	0.00 (1.74)	0.059	[0.005 0.112]
Sleep Breathing Disorders (SBD)	3.39 (0.83)	-0.229	[-0.368 -0.089]
Sleep-Wake Transition Disorders (SWTD)	8.11 (2.53)	-0.154	[-0.230 -0.078]
Disorders of excessive somnolence (DOES)	7.06 (2.49)	-0.683	[-0.826 -0.539]
Physical activity (days/week)	3.50 (2.18)	0.091	[0.040 0.142]

Table A.6: Coefficients of the binomial part of ZINB model for *with_depr* score

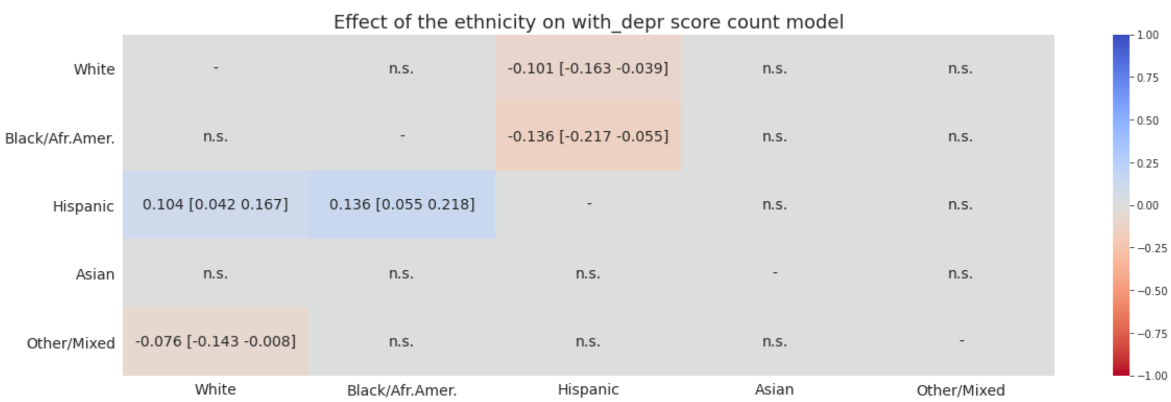


Figure A.3: Effect of the ethnicity on *with_depr* score

Feature	mean (σ)	ß (count)	[2.5%97.5%] (count)
Intercept		-0.630	[-0.933 -0.328]
Sex: M vs. F		-0.141	[-0.173 -0.109]
Parents yearly income (x1000 $): 75-100 vs. 50-75		-0.129	[-0.189 -0.069]
Parents yearly income (x1000 $):>200 vs. 100-200		-0.010	[-0.158 -0.041]
# people cohabiting	4.70 (1.52)	-0.018	[-0.030 -0.006]
Size	0.00 (1.74)	0.023	[0.013 0.033]
Slenderness	0.00 (0.88)	-0.030	[-0.050 -0.010]
Disorders of initiating and maintaining sleep (DIMS)	11.84 (3.75)	0.037	[0.032 0.042]
Sleep Breathing Disorders (SBD)	3.74 (1.20)	0.011	[-0.002 0.025]
Disorders of arousal (DA)	3.39 (0.83)	0.088	[0.069 0.107]
Sleep-Wake Transition Disorders (SWTD)	8.11 (2.53)	0.030	[0.022 0.037]
Disorders of excessive somnolence (DOES)	7.06 (2.49)	0.048	[0.041 0.055]
Sleep Hyperhydrosis (SHY)	2.41 (1.12)	0.046	[0.033 0.060]
Screentime: Video (h)	1.06 (1.16)	0.023	[0.009 0.037]

Table A.7: Coefficients of the count part of ZINB model for *som_comp* score

Feature	mean (σ)	γ (zero)	[2.5% 97.5%] (zero)
Intercept		8.958	[7.127 10.789]
Age (months)	125.04 (9.73)	-0.008	[-0.016 0.001]
# people cohabiting	4.70 (1.52)	0.071	[0.025 0.117]
Disorders of initiating and maintaining sleep (DIMS)	11.84(3.75)	-0.081	[-0.112 -0.049]
Sleep Breathing Disorders (SBD)	3.74 (1.20)	-0.173	[-0.267 -0.079]
Disorders of arousal (DA)	3.39 (0.83)	-1.346	[-1.758 -0.932]
Sleep-Wake Transition Disorders (SWTD)	8.11 (2.53)	-0.214	[-0.288 -0.141]
Disorders of excessive somnolence (DOES)	7.06 (2.49)	-0.313	[-0.404 -0.222]
Sleep Hyperhydrosis (SHY)	2.41 (1.12)	-0.204	[-0.364 -0.043]
Physical activity (days/week)	3.50(2.18)	-0.037	[-0.069 -0.005]

Table A.8: Coefficients of the binomial part of ZINB model for *som_comp* score

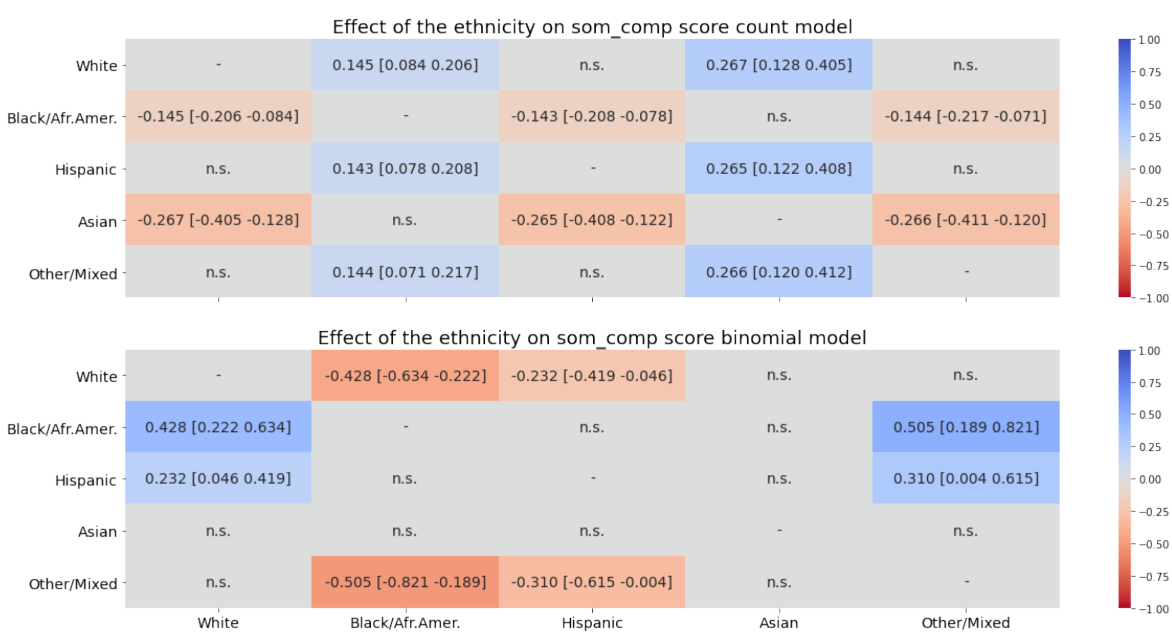


Figure A.4: Effect of the ethnicity on *som_comp* score

Feature	mean (σ)	ß (count)	(2.5% 97.5%) (count)
Age (months)	125.04 (9.73)	-0.005	[-0.007 -0.003]
Sex: M vs. F		0.089	[0.053 0.126]
Parents yearly income (x1000 $): 50-75 vs. 35-50		-0.091	[-0.166 -0.016]
Parents yearly income (x1000 $): 75-100 vs. 50-75		-0.078	[-0.144 -0.013
Parents yearly income (x1000 $): 100-200 vs. 75-100		-0.068	[-0.125 -0.011]
Parents yearly income (x1000 $): >200 vs. 100-200		-0.074	[-0.140 -0.008]
Size	0.00 (1.74)	0.037	[0.026 0.047]
Slenderness	0.00 (0.88)	-0.036	[-0.057 -0.014]
Disorders of initiating and maintaining sleep (DIMS)	11.84 (3.75)	0.039	[0.034 0.045]
Disorders of arousal (DA)	3.39 (0.83)	0.048	[0.027 0.069]
sleep-Wake Transition Disorders (SWTD)	8.11 (2.53)	0.046	[0.039 0.054]
Disorders of excessive somnolence (DOES)	7.06 (2.49)	0.050	[0.043 0.058]
Sleep Hyperhydrosis (SHY)	2.41 (1.12)	0.033	[0.019 0.048]
Physical activity (days/week)	3.50 (2.18)	-0.020	[-0.028 -0.011]
Screentime: TV (h)	1.24 (1.03)	0.049	[0.032 0.066]
Screentime: Videogames (h)	1.08 (1.14)	0.018	[0.002 0.034]

Table A.9: Coefficients of the count part of ZINB model for *social_pr* score

Feature	mean (σ)	γ (zero)	(2.5% 97.5%) (zero)
Intercept		4.423	[3.129 5.717]
Age (months)	125.04 (9.73)	1.448	[0.007 0.022]
Disorders of initiating and maintaining sleep (DIMS)	11.84 (3.75)	-0.154	[-0.188 -0.120]
Sleep Breathing Disorders (SBD)	3.74 (1.20)	-2.587	[-0.365 -0.152]
Disorders of arousal (DA}	3.39 {0.83)	-4.641	[-0.685 -0.243]
Sleep-Wake Transiti on Disorders {SWTD)	8.11(2.53)	-0.145	[-0.200 -0.090]
Disorders of excesslve somnolence (DOES)	7.06 (2.49)	-0.318	[-0.386 -0.249]
Sleep Hyperhydrosls (SHV)	2.41(1.12)	-2.206	[-0.399 -0.042]
Physical activity (days/week)	3.50 {2.18)	4.693	[0.014 0.080]

Table A.10: Coefficients of the binomial part of ZINB model for *social_pr* score

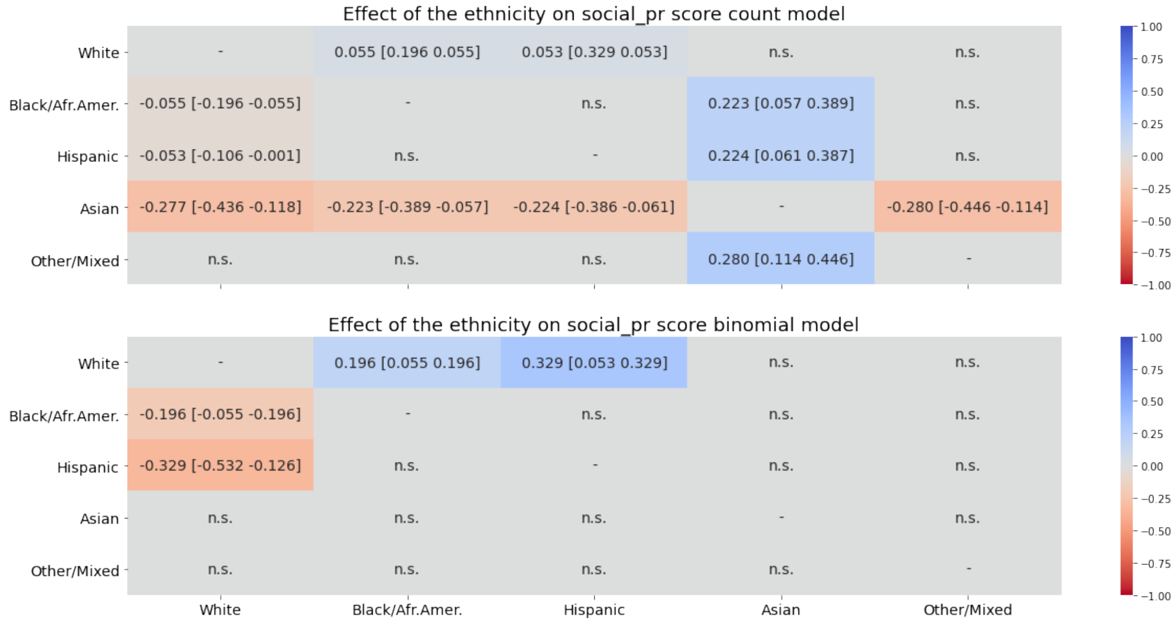


Figure A.5: Effect of the ethnicity on *social_pr* score

Feature	mean (σ)	β (count)	[2.5% 97.5%] (count)
Intercept		-1.608	[-1.700 -1.517]
Sex: M vs. F		0.168	[0.135 0.201]
Disorders of initiating and maintaining sleep (DIMS)	11.84 (3.75)	0.074	[0.070 0.078]
Disorders of arousal (DA)	3.39 (0.83)	0.078	[0.062 0.094]
Sleep-Wake Transition Disorders (SWTD)	8.11 (2.53)	0.066	[0.060 0.072]
Disorders of excessive somnolence (DOES)	7.06 (2.49)	0.036	[0.030 0.041]
Sleep Hyperhydrosis (SHY)	2.41 (1.12)	0.019	[0.008 0.031]
Physical activity (days/week)	3.50 (2.18)	-0.011	[-0.018 -0.004]
Screentime: TV (h)	1.24 (1.03)	0.033	[0.019 0.048]
Screentime: Videogames (h)	1.08 (1.14)	0.030	[0.016 0.044]

Table A.11: Coefficients of the count part of ZINB model for *thought_pr* score

Feature	mean (σ)	γ (zero)	[2.5% 97.5%] (zero)
Intercept		15.035	[10.334 19.737]
Sex:M vs.F		-0.368	[-0.515 -0.221]
# people cohabiting	4.70 (1.52)	0.073	[0.032 0.114]
Disorders of initiating and maintaining sleep (DIMS)	11.84 (3.75)	-0.178	[-0.206 -0.149]
Sleep Breathing Disorders (SBD)	3.74 (1.20)	-0.127	[-0.211-0.043]
Disorders of Arousal (DA)	3.39 (0.83)	-2.999	[-4.557-1.441]
Sleep-Wake Transition Disorders (SWTD)	8.11(2.53)	-0.346	[-0.426 -0.265]
Disorders of excessive somnolence (DOES)	7.06 (2.49)	-0.310	[-0.375 -0.245]
Sleep Hyperhydrosis (SHY)	2.41 (1.12)	-0.371	

Table A.12: Coefficients of the binomial part of ZINB model for *thought_pr* score

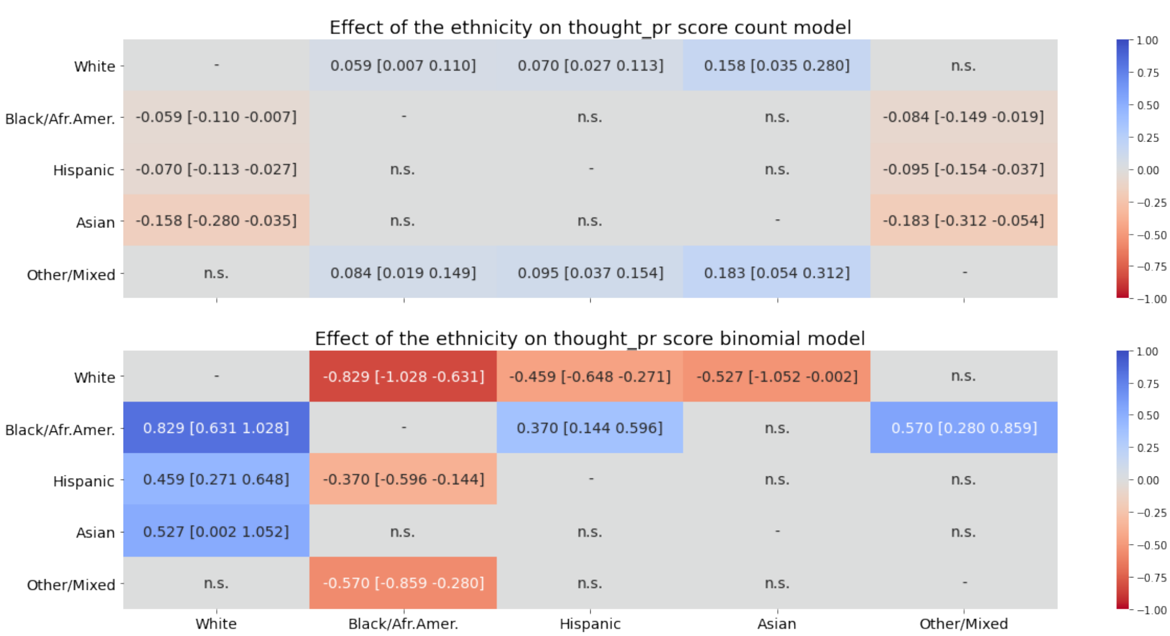


Figure A.6: Effect of the ethnicity on *thought_pr* score

Feature	mean (σ)	β (count)	[2.5% 97.5%] (count)
Intercept		-0.055	
Sex: M vs. F		0.270	[0.239 0.302]
# people cohabiting	4.70 (1.52)	-0.019	[-0.029 -0.009]
Size	0.00 (1.74)	-0.029	[-0.037 -0.021]
Slenderness	0.00 (0.88)	-0.025	[-0.042 -0.009]
Disorders of initiating and maintaining sleep (DIMS)	11.84 (3.75)	0.036	[0.032 0.040]
Disorders of arousal (DA)	3.39 (0.83)	0.027	[0.010 0.044]
Sleep-Wake Transition Disorders (SWTD)	8.11 (2.53)	0.032	[0.026 0.038]
Disorders of excessive somnolence (DOES)	7.06 (2.49)	0.048	[0.042 0.054]
Sleep Hyperhydrosis (SHY)	2.41 (1.12)	0.017	[0.005 0.029]
Physical activity (days/week)	3.50 (2.18)	-0.024	[-0.030 -0.016]
Screentime: TV (h)	1.24 (1.03)	0.030	[0.016 0.044]
Screentime: Video (h)	1.06 (1.16)	0.029	[0.015 0.043]
Screentime: Videogames (h)	1.08 (1.14)	0.026	[0.012 0.040]

Table A.13: Coefficients of the count part of ZINB model for *att_pr* score

Feature	mean (σ)	γ (zero)	[2.5% 97.5%] (zero)
Intercept		3.563	[2.728 4.398]
Age (months)	125.04 (9.73)	0.011	[0.006 0.015]
Sex: M vs. F		-0.582	[-0.680 -0.483]
# people cohabiting	4.70 (1.52)	0.057	(0.026 0.879]
Disorders of initiating and maintaining sleep (DIMS)	11.84 (3.75)	-0.154	[0.177 0.132]
Sleep Breathing Disorders (SBD)	3.74 (1.20)	-0.105	[0.164 0.046]
Disorders of arousal (DA)	3.39 (0.83)	0.162	[-0.266 -0.058]
Sleep-Wake Transition Disorders (SWTD)	8.11 (2.53)	-0.164	[0.201 0.128]
Disorders of excessive somnolence (DOES)	7.06 (2.49)	-0.319	[0.362 0.275]
Sleep Hyperhydrosis (SHY)	2.41 (1.12)	0.126	[0.210 -0.042]
Physical activity (days/week)	3.50 (2.18)	0.032	[0.010 0.054]

Table A.14: Coefficients of the binomial part of ZINB model for *att_pr* score

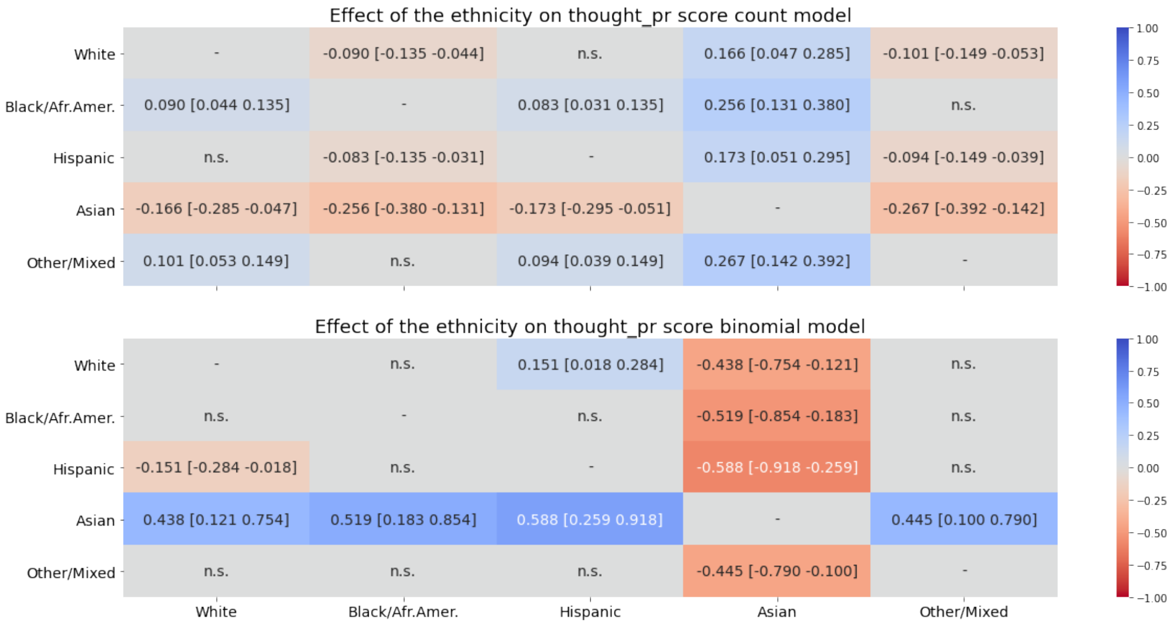


Figure A.7: Effect of the ethnicity on *att_pr* score

Feature	mean (σ)	β (count)	[2.5% 97.5%] (count)
Intercept		-0.533	[-0.860 -0.205]
Age (months)	125.04 (9.73)	-0.003	[-0.005 -0.001]
Sex: M vs. F		0.301	[0.255 0.346]
Parents yearly income (x1000 $): 35-50 vs. 25-35		-0.088	[-0.186 -0.009]
Parents yearly income (x1000 $): 50-75 vs. 35-50		-0.1222	[-0.204 -0.040]
Parents yearly income (x1000 $): 100-200 vs. 75-100		-0.141	[-0.205 -0.077]
# people cohabiting	4.70 (1.52)	0.038	[0.024 0.051]
Disorders of initiating and maintaining sleep (DIMS)	11.84 (3.75)	0.038	[0.033 0.044]
Disorders of arousal (DA)	3.39 (0.83)	0.40	[0.018 0.063]
Sleep-Wake Transition Disorders (SWTD)	8.11 (2.53)	0.025	[0.017 0.034]
Disorders of excessive somnolence (DOES)	7.06 (2.49)	0.039	[0.031 0.047]
Screentime: TV(h)	1.24(1.03)	0.068	[0.051 0.086]
Screentime: Social activities (h)	0.60 (0.95)	0.084	[0.065 0.103]

Table A.15: Coefficients of the count part of ZINB model for *rule_br_bh* score

Feature	mean (σ)	γ (zero)	[2.5% 97.5%] (zero)
Intercept		4.122	[2.763 5.481]
Age (months)	125.04 (9.73)	0.013	[0.005 0.020]
Sex: M vs. F		-0.629	[-0.778 -0.480]
# people cohabiting	4.70 (1.52)	0.087	[0.045 0.128]
Disorders of initiating and maintaining sleep (DIMS)	11.84 (3.75)	-0.123	[-0.154 -0.092]
Sleep Breathing Disorders (SBD)	3.74 (1.20)	-0.230	[-0.315 -0.145]
Disorders of arousal (DA)	3.39 (0.83)	-0.336	[-0.523 -0.148]
Sleep-Wake Transition Disorders (SWTD)	8.11 (2.53)	-0.113	[-0.166 -0.059]
Disorders of excessive somnolence (DOES)	7.06 (2.49)	-0.295	[-0.361 -0.229]
Sleep Hyperhydrosis (SHY)	2.41 (1.12)	-0.353	[-0.533 -0.174]

Table A.16: Coefficients of the binomial part of ZINB model for *rule_br_bh* score

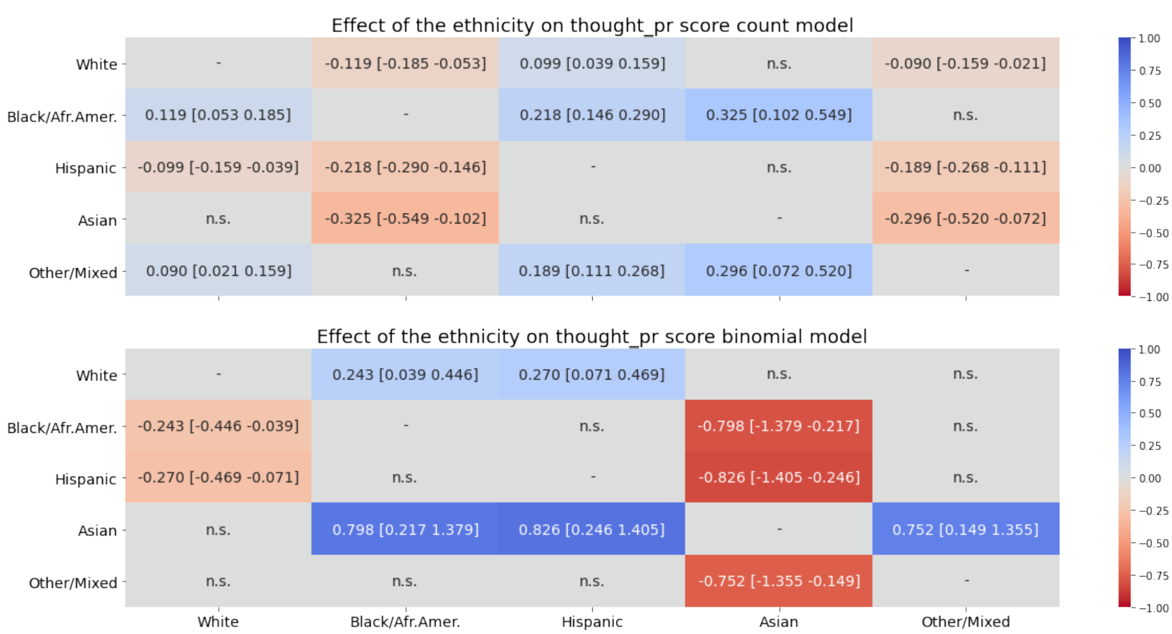


Figure A.8: Effect of the ethnicity on *rule_br_bh* score

Feature	mean (σ)	β (count)	[2.5% 97.5%] (count)
Intercept		0.439	[0.159 0.718]
Age (months)	125.04 (9.73)	-0.005	[-0.007 -0.003]
Sex: M vs. F		0.197	[0.163 0.231]
Parents yearly income (x1000 $): 75-100 vs. 50-75		-0.076	[-0.136 -0.015]
Parents yearly income (x1000 $): 100-200 vs. 75-100		-0.059	[-0.111 -0.007]
# people cohabiting	4.70 (1.52)	0.033	[0.022 0.043]
Size	0.00 (1.74)	0.013	[0.003 0.023]
Slenderness	0.00 (0.88)	-0.032	[-0.052 -0.012]
Disorders of initiating and maintaining sleep (DIMS)	11.84 (3.75)	0.046	[0.042 0.051]
Disorders of arousal (DA)	3.39 (0.83)	0.046	[0.027 0,066]
Sleep-Wake Transition Disorders (SWTD)	8.11 (2.53)	0.036	[0.028 0.042]
Disorders of excessive somnolence (DOES)	7.06 (2.49)	0.058	[0.051 0.065]
Sleep Hyperhydrosis (SHY)	2.41 (1.12)	0.0245	[0.010 0.037]
Screentime: TV (h)	1.24 (1.03)	0.033	[0.018 0.048]
Screentime: Social activities (h)	0.60 (0.95)	0.030	[0.013 0.047]

Table A.17: Coefficients of the count part of ZINB model for *agg_bh* score

Feature	mean (σ)	γ (zero)	[2.5% 97.5%] (zero)
Intercept		6.183	[5.4544 6.913]
Sex: M vs. F		-0.321	[-0.445 -0.197]
Disorders of initiating and maintaining sleep (DIMS)	11.84 (3.75)	-0.203	[-0.235 -0.171]
Sleep Breathing Disorders (SBD)	3.74 (1.20)	-0.147	[-0.224 -0.070]
Disorders of arousal (DA)	3.39 (0.83)	-0.222	[-0.366 -0.078]
Sleep-Wake Transition Disorders (SWTD)	8.11 (2.53)	-0.156	[-0.205 -0.107]
Disorders of excessive somnolence (DOES)	7.06 (2.49)	-0.378	[-0.439 -0.317]
Sleep Hyperhydrosis (SHY)	2.41 (1.12)	-0.287	[-0.435 -0.139]

Table A.18: Coefficients of the binomial part of ZINB model for *agg_bh* score

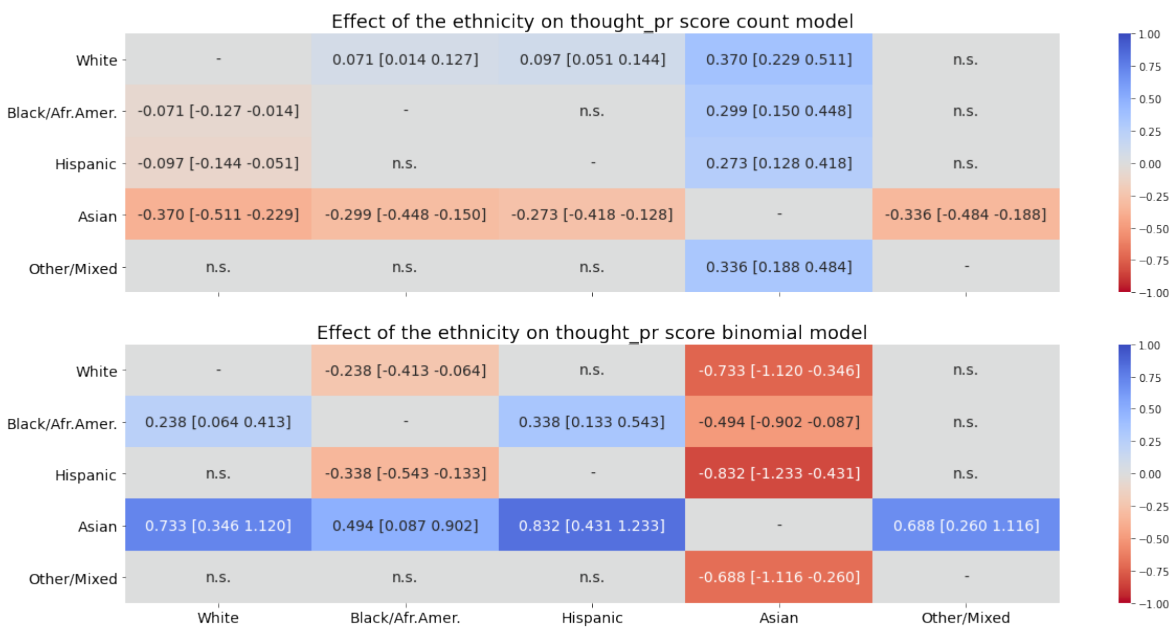


Figure A.9: Effect of the ethnicity on *agg_bh* score

A.2 Models for brain measures

Feature	mean (σ)	β	[2.5% 97.5%]
Intercept		1643337.715	[1526632.889 1760042.542]
Age (months)	125.04 (9.73)	-1728.947	[-2557.282 -900.611]
Sex: M vs. F		125615.475	[112032.456 139198.494]
Parents yearly income (x1000 $): 16-25 vs. 12-16		64391.080	[3733.162 125048.997]
Parents yearly income (x1000 $): 35-50 vs. 25-35		45948.199	[9839.096 82057.303]
Parents yearly income (x1000 $): 75-50 vs. 35-50		38672.804	[8950.022 68395.585]
Size	0.00 (1.74)	9818.461	[5619.892 14017.031]
Slenderness	0.00 (0.88)	27471.888	[18175.931 36767.846]
Sleep Breathing Disorders (SBD)	3.74 (1.20)	-7122.421	[-13102.316 -1142.526]
Physical activity (days/week)	3.50 (2.18)	3974.077	[742.253 7205.901]

Figure A.10: Coefficients of the linear model for ICV

Effect of the ethnicity on ICV

	White	Black/Afr.Amer.	Hispanic	Asian	Other/Mixed
White	-	72675.960 [47528.835 97823.085]	40557.130 [19765.347 61348.913]	55658.727 [495.757 110821.698]	29587.350 [6186.055 52988.645]
Black/Afr.Amer.	-72675.960 [-97823.085 -47528.835]	-	-32118.830 [-59641.402 -4596.258]	n.s.	-43088.610 [-73599.839 -12577.381]
Hispanic	-40557.130 [-61348.913 -19765.347]	32118.830 [4596.258 59641.402]	-	n.s.	n.s.
Asian	-55658.727 [-110821.698 -495.757]	n.s.	n.s.	-	n.s.
Other/Mixed	-29587.350 [-52988.645 -6186.055]	43088.610 [12577.381 73599.839]	n.s.	n.s.	-

Table A.19: Effect of the ethnicity on ICV

Feature	mean (σ)	β	[2.5% 97.5%]
Intercept		100676.162	[94410.826 106941.498]
Age (months)	125.04 (9.73)	-118.472	[-166.720 -70.225]
Sex: M vs. F		7588.686	[6801.597 8375.774]
Size	0.00 (1.74)	495.079	[257.175 732.983]
Slenderness	0.00 (0.88)	2017.942	[1485.354 2550.530]
Sleep-Wake Transition Disorders (SWTD)	8.11 (2.53)	-255.818	[-412.520 -99.116]
Physical activity (days/week)	3.50 (2.18)	243.376	[57.287 429.464]
Screentime: Social activities (h)	0.60 (0.95)	-601.276	[-1001.299 -201.253]

Figure A.11: Coefficients of the linear model for left hemisphere cortical surface area

Effect of the ethnicity on left hemisphere cortical surface area

	White	Black/Afr.Amer.	Hispanic	Asian	Other/Mixed
White	-	7741.966 [6357.031 9126.900]	2249.836 [1121.655 3378.018]	n.s.	1632.153 [298.812 2965.494]
Black/Afr.Amer.	-7741.966 [-9126.900 -6357.031]	-	-5492.129 [-7088.384 -3895.875]	-7915.434 [-11321.606 -4509.262]	-6109.812 [-7860.159 -4359.465]
Hispanic	-2249.836 [-3378.018 -1121.655]	5492.129 [3895.875 7088.384]	-	n.s.	n.s.
Asian	n.s.	7915.434 [4509.262 11321.606]	n.s.	-	n.s.
Other/Mixed	-1632.153 [-2965.494 -298.812]	6109.812 [4359.465 7860.159]	n.s.	n.s.	-

Table A.20: Effect of the ethnicity on left hemisphere cortical surface area

Feature	mean (σ)	β	[2.5% 97.5%]
Intercept		102355.635	[96064.077 108647.193]
Age (months)	125.04 (9.73)	-120.225	[-168.540 -71.910]
Sex: M vs. F		8060.042	[7266.781 8853.302]
Size	0.00 (1.74)	500.228	[258.124 742.332]
Slenderness	0.00 (0.88)	2043.441	[1505.545 2581.337]
Sleep-Wake Transition Disorders (SWTD)	8.11 (2.53)	-286.486	[-444.696 -128.277]
Screentime: Video (h)	1.06 (1.16)	-536.178	[-859.513 -212.844]

Figure A.12: Coefficients of the linear model for right hemisphere cortical surface area

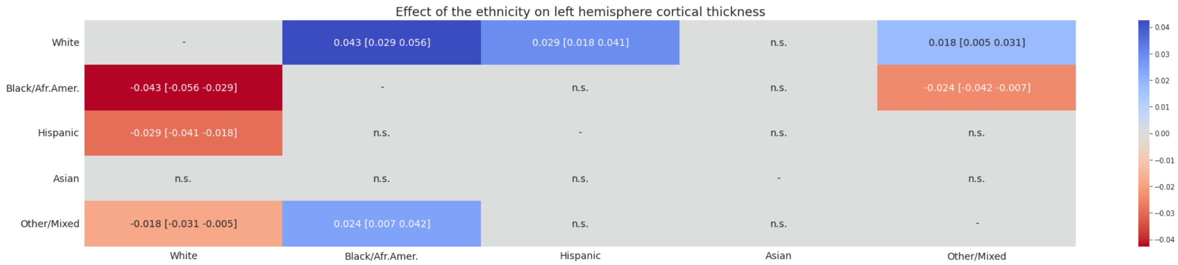
Effect of the ethnicity on right hemisphere cortical surface area

	White	Black/Afr.Amer.	Hispanic	Asian	Other/Mixed
White	-	7566.479 [6174.690 8958.269]	2361.776 [1225.943 3497.608]	n.s.	1479.185 [135.155 2823.215]
Black/Afr.Amer.	-7566.479 [-8958.269 -6174.690]	-	-5204.704 [-6812.144 -3597.264]	-7701.875 [-11147.454 -4256.295]	-6087.295 [-7857.792 -4316.797]
Hispanic	-2361.776 [-3497.608 -1225.943]	5204.704 [3597.264 6812.144]	-	n.s.	n.s.
Asian	n.s.	7701.875 [4256.295 11147.454]	n.s.	-	n.s.
Other/Mixed	-1479.185 [-2823.215 -135.155]	6087.295 [4316.797 7857.792]	n.s.	n.s.	-

Table A.21: Effect of the ethnicity on right hemisphere cortical surface

Feature	mean (σ)	β	[2.5% 97.5%]
Intercept		2.711	[2.705 2.717]
Size	0.00 (1.74)	-0.003	[-0.005 -0.001]
Screentime: Videogames (h)	1.08 (1.14)	-0.006	[-0.009 -0.003]

Figure A.13: Coefficients of the linear model for left hemisphere cortical thickness



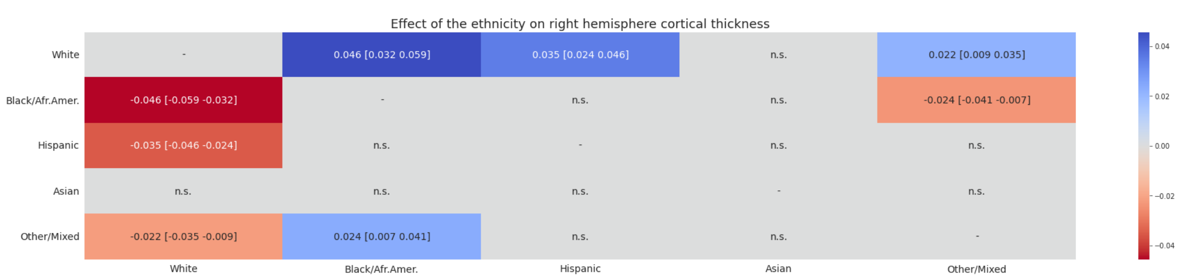
Effect of the ethnicity on left hemisphere cortical thickness

	White	Black/Afr.Amer.	Hispanic	Asian	Other/Mixed
White	-	0.043 [0.029 0.056]	0.029 [0.018 0.041]	n.s.	0.018 [0.005 0.031]
Black/Afr.Amer.	-0.043 [-0.056 -0.029]	-	n.s.	n.s.	-0.024 [-0.042 -0.007]
Hispanic	-0.029 [-0.041 -0.018]	n.s.	-	n.s.	n.s.
Asian	n.s.	n.s.	n.s.	-	n.s.
Other/Mixed	-0.018 [-0.031 -0.005]	0.024 [0.007 0.042]	n.s.	n.s.	-

Table A.22: Effect of the ethnicity on left hemisphere cortical thickness

Feature	mean (σ)	β	[2.5% 97.5%]
Intercept		2.649	[2.632 2.666]
Sex: M vs. F		-0.008	[-0.016 0.000]
Size	0.00 (1.74)	-0.003	[-0.006 -0.001]
Sleep-Wake Transition Disorders (SWTD)	8.11 (2.53)	0.002	[0.001 0.004]
Screentime: Videogames (h)	1.08 (1.14)	-0.006	[-0.009 -0.002]

Figure A.14: Coefficients of the linear model for right hemisphere cortical thickness



Effect of the ethnicity on right hemisphere cortical thickness

	White	Black/Afr.Amer.	Hispanic	Asian	Other/Mixed
White	-	0.046 [0.032 0.059]	0.035 [0.024 0.046]	n.s.	0.022 [0.009 0.035]
Black/Afr.Amer.	-0.046 [-0.059 -0.032]	-	n.s.	n.s.	-0.024 [-0.041 -0.007]
Hispanic	-0.035 [-0.046 -0.024]	n.s.	-	n.s.	n.s.
Asian	n.s.	n.s.	n.s.	-	n.s.
Other/Mixed	-0.022 [-0.035 -0.009]	0.024 [0.007 0.041]	n.s.	n.s.	-

Table A.23: Effect of the ethnicity on right hemisphere cortical thickness

References

[1] Centres for Disease Control and Prevention,*Infographic on Screen Time vs. Lean Time*, https://www.cdc.gov/nccdphp/dnpao/multimedia/infographics/getmoving.html

[2] Carson V., Hunter S., Kuzik N., Gray C.E., Poitras V.J., Chaput J.-P., et al., *Systematic review of sedentary behaviour and health indicators in school-aged children and youth: an update.* Appl Physiol Nutr Metab. 2016;41:S240–65.

[3] Stiglic N., Viner R.M., *Effects of screentime on the health and well-being of children and adolescents: a systematic review of reviews*, BMJ Open 2019.

[4] Paulus M.P., Squeglia L.M., Bagot K., Jacobus J., Kuplicki R., Breslin F.J., Bodurka J., Morris A.S., Thompson W.K., Bartsch H., Tapert S.F., *Screen media activity and brain structure in youth: Evidence for diverse structural correlation networks from the ABCD study.* Neuroimage. 2019 Jan 15;185:140-153.

[5] Piguet C., Mihailov A., Grigis A., Laidi C., Duchesnay E. and Houenou J., *Irritability Is Associated With Decreased Cortical Surface Area and Anxiety With Decreased Gyrification During Brain Development.* (2021) Front. Psychiatry 12:744419.

[6] Ducharme S., Hudziak J.J., Botteron K.N., Albaugh M.D. et al., *Brain Development Cooperative Group. Decreased regional cortical thickness and thinning rate are associated with inattention symptoms in healthy children.*, J. Am. Acad. Child Adolesc. Psychiatry. 2012 Jan;51(1):18-27.e2.

[7] Achenbach T.M., Resorta L.A. *Manual for ASEBA school-age forms and profiles [Internet]*, Research Centre for Children, Youth and Families, 2001. Available from: https://aseba.org/school-age/

[8] R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

[9] *ABCD Study®, Teen Brains. Today's Science. Brighter Future.®, El cerebro adolescente. La ciencia de hoy. Un futuro más brillante.®*, (https://abcdstudy.org)

[10] Barch D.M., Albaugh M.D., Avenevoli S., Chang L., Clark D.B., Glantz M.D., Hudziak J.J., Jernigan T.L., Tapert S.F., Yurgelun-Todd D., Alia-Klein N., Potter

A.S., Paulus M.P., Prouty D., Zucker R.A., Sher K.J., *Demographic, physical and mental health assessments in the adolescent brain and cognitive development study: Rationale and description*, Dev Cogn Neurosci., August 2018

[11] Guerrero M.D., Barnes J.D., Chaput JP. et al. *Screen time and problem behaviors in children: exploring the mediating role of sleep duration*. Int J Behav Nutr Phys Act 16, 105 (2019).

[12] Bruni O., Ottaviano S., Guidetti V., Romoli M., Innocenzi M., Cortesi F., Giannotti F. *The Sleep Disturbance Scale for Children (SDSC). Construction and validation of an instrument to evaluate sleep disturbances in childhood and adolescence.* J Sleep Res. 1996 Dec;5(4):251-61.

[13] Fischl B., Salat D.H., Busa E., Albert M., Dieterich M., Haselgrove C. et al. *Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain.* Neuron 2002; 33: 341–355.

[14] Desikan R.S., Segonne F., Fischl B., Quinn B.T., Dickerson B.C., Blacker D. et al. *An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest.* Neuroimage 2006; 31: 968–980.

[15] ABCD Youth Screen Time Survey - STQ (`https://nda.nih.gov/data_structure.html?short_name=abcd_stq01`)

[16] Children, Adolescents, and Television Committee on Public Education Pediatrics Feb 2001, 107 (2) 423-426.

[17] Glynis M Breakwell, Hammond S., Fife-Schaw C., Smith J.A.,*Research Methods in Psychology*, Chapters 10-11, SAGE Publications, 2006

[18] Nagata J.M., Iyer P., Chu J., Baker F.C., Pettee Gabriel K., Garber A.K., Murray S.B., Bibbins-Domingo K., Ganson K.T., *Contemporary screen time modalities among children 9-10years old and binge-eating disorder at one-year follow-up: A prospective cohort study*. Int J Eat Disord. 2021 May; 54(5):887-892.

[19] Cheng W., Rolls E., Gong W., Du J., Zhang J., Zhang X.Y., Li F., Feng J., *Sleep duration, brain structure, and psychiatric and cognitive problems in children.* Mol Psychiatry. 2020 Feb 3.

[20] Noble W. S., *How does multiple testing correction work?*, Nature biotechnology, 27(12), 1135–1137., 2009.

[21] Gareth J., Daniela W., Trevor H., and Robert T., *An Introduction to Statistical Learning: With Applications in R.*, 2014 Springer Publishing Company, Incorporated, Chapter 3.

[22] Nowland R., Necka E. A. et al., *Loneliness and Social Internet Use: Pathways to*

Reconnection in a Digital World?, Perspectives on Psychological Science, vol. 13, no. 1, Jan. 2018, pp. 70–87.

[23] Caspi Y., Brouwer R. M. , Schnack H. G., van de Nieuwenhuijzen M. E. , Cahn W., Kahn S. R., et. al, *Changes in the intracranial volume from early adulthood to the sixth decade of life: A longitudinal study*, NeuroImage, Volume 220, 2020, 116842,ISSN 1053-8119.

[24] Haijma S.V., Van Haren N., Cahn W., Koolschijn P.C., Hulshoff Pol H.E., Kahn R.S. *Brain volumes in schizophrenia: a meta-analysis in over 18 000 subjects.*, Schizophr Bull. 2013 Sep;39(5):1129-38.

[25] Klein M., Walters R.K., Demontis D. et al., *Genetic Markers of ADHD-Related Variations in Intracranial Volume. Am J Psychiatry.*, 2019 Mar 1;176(3):228-238.

[26] van Tol M., van der Wee N.J.A., van den Heuvel O.A., et al., *Regional Brain Volume in Depression and Anxiety Disorders*, Arch Gen Psychiatry. 2010; 67(10) :1002–1011.

[27] Choi Y. Y., Kim H., Lee K. H., *Race or ethnicity is a factor that should be considered for establishing the normative data of subcortical brain volumes in normal aging*, Alzheimer's Dement., 16: e041919. 2020.

[28] Kühn S., Lorenz R., Banaschewski T., Barker G.J., Büchel C., Conrod P.J., et al., *Positive Association of Video Game Playing with Left Frontal Cortical Thickness in Adolescents.* (2014) PLoS ONE 9(3): e91506.

[29] Dabbs K., Jones J.E., Jackson D.C., Seidenberg M., Hermann B.P., *Patterns of cortical thickness and the Child Behavior Checklist in childhood epilepsy.* Epilepsy Behav. 2013 Oct;29(1):198-204.