STATS 557 FINAL PROJECT

Analysis of factors that affect early childhood education

Group Members: Geoffrey Li, Win Nawat Suvansinpan, Roshan Ramkeesoon

ABSTRACT

Investigation of factors associated with 3rd grade student grades on math, science, and reading assessment administered in 2002 U.S. nationwide Early Childhood Longitudinal study. Controlling for parent income, increased weekly television watching and increased after school television watching were both found to have negative associations ($p<0.01$) with all subject exams scores. Unit hour after school television watching associated with greater average negative change in exam scores than unit hour of weekly television watched. Consistent bedtime were found to be associated with average increase in all subject exam scores ($p<0.001$). Earlier bedtimes were found to be associated with increased reading and science exam scores ($p<0.001$), but no significant relationship was found with math scores. Greater student age at enrollment were found to be associated with average increase in all subject exam scores ($p<0.001$).

Introduction

The Early Childhood Longitudinal Study[5] was a longitudinal study conducted by the Institute for Education Sciences, part of U.S. federal government, in an effort to understand the factors that affect early childhood education. The conductors collected data on factors that affect educational outcomes by interviewing students, their teachers, and their parents/guardians. The study tracked students from kindergarten through 8th grade, conducting interviews approximately once every other year.

While many factors play a role in childhood education outcomes, in this paper we will (1) build on prior studies which claim a statistically significant negative association between television watching and education outcomes[3][4], (2) examine the association between child bedtime hour and education outcomes, and (3) examine the association between child age at enrollment on education outcomes.

Dataset Description

The ECLS study collected data on factors that affect educational outcomes by interviewing students, teachers, and parents of participants. Parents and teachers were interviewed over phone, while students were interviewed in-person. All interviews conducted by trained staff. Tracking 3,341 variables ranging from student bedtime to the teacher's level of interest in teaching, the study is one of the largest datasets of its kind. We will only use the 2001 dataset, which pertains to students in 3rd grade from the Kindergarten class of 1998-1999. There are 15,000 students in the dataset for the 3rd grade interviews. Upon inspecting the dataset, many fields report high levels (>90%) of data completeness.

The conductors used multi-stage stratified cluster sampling. Schools were randomly selected across the U.S.A, and of the selected schools, kindergartners were randomly selected. On average, 23 kindergartners were selected from each selected school. The conductors ensured the sample of students is diverse (socioeconomic status, race, private/public schools, etc). In addition, the students are from a one-year cohort, which reduces the likelihood that siblings are included in the dataset (except for the rare case that twins are both randomly selected). *Because of the sampling scheme, large sample size, and narrow age band of students, we will assume the responses as independent.*

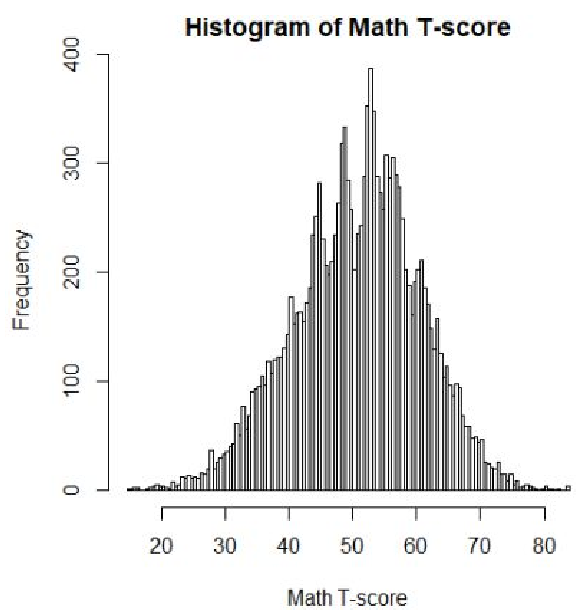
Common Response Variable

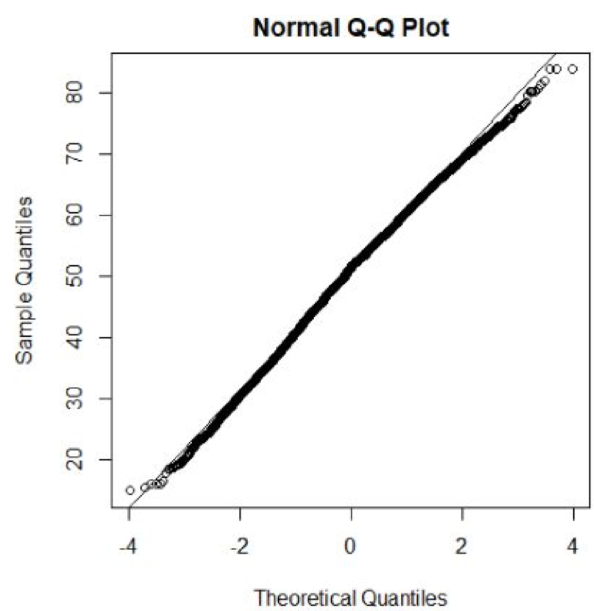
Math, reading, and science exams were administered during student surveys. Raw scores in each subject were also provided as T-scores, normalized to N(μ = 50, sd = 10). To measure students' performance in our questions, we will be using math, reading and science T-scores which reflect the students' performance relative to national average. Normalization is useful so we can make comparisons in response variables across the three different subjects

Response variables		
C5R2MTSC	Third grade math T-score	Math score relative to national average
C5R2RTSC	Third grade reading T-score	Reading score relative to national average
C5STSCOR	Third grade science T-score	Science score relative to national average

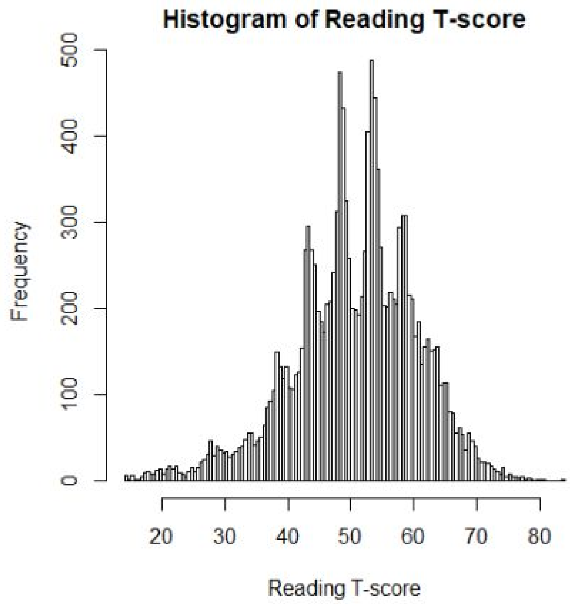
After removing missing and not-applicable data, the distributions of both math and reading T-scores are approximately normal.

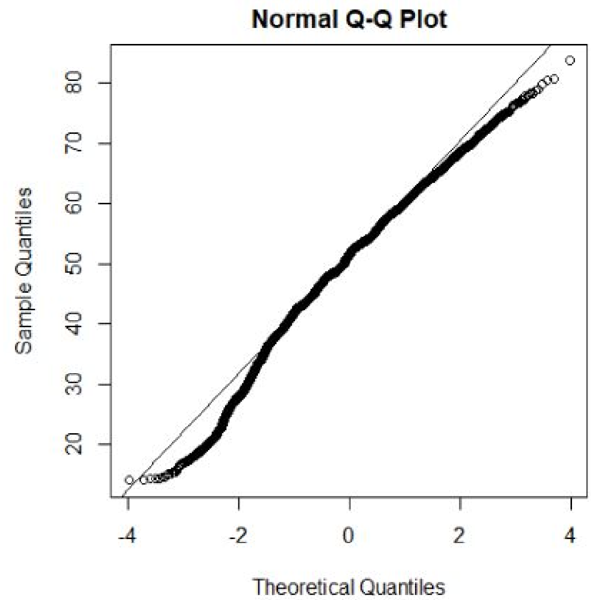
- **Distribution of math T-scores:**



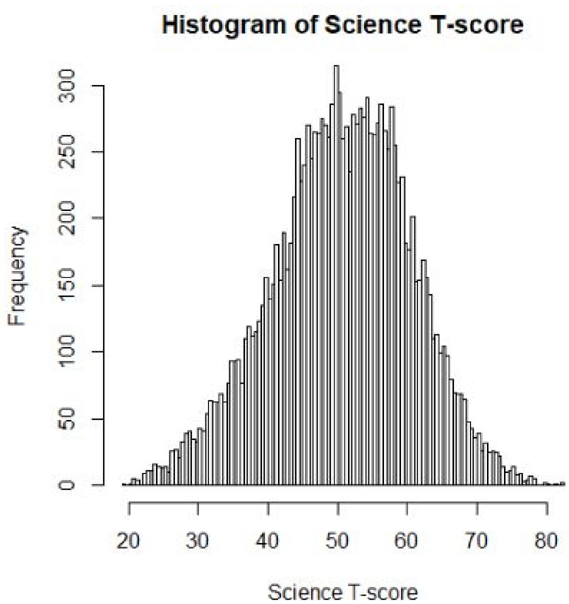


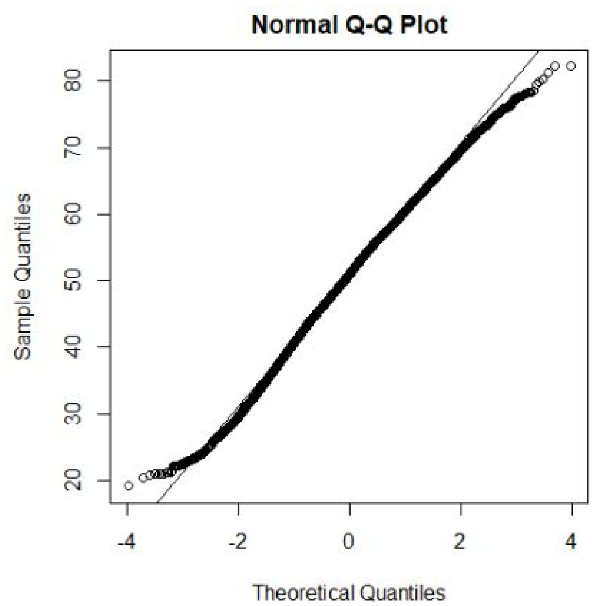
- **Distribution of reading T-scores**





- **Distribution of science T-scores**





Question 1: **What is the association of tv watching habits on child education outcomes?**

We will first see if we can use our data to match prior results that reject the hypothesis that there is no association between weekly television watching hours and child education outcomes.

Prior work has also indicated that there is a statistically significant positive association between hours of after school activities and child education outcomes[1]. While it is not an exact comparison, another way to explore that result is to see if there exists an association between after school television watching and child education outcomes. The null hypothesis of this second part of the analyses is that there is no association between after school television watching hours and child education outcomes.

We theorize that socioeconomic status may be associated with both behavior at home, including television watching, and educational outcomes. For that reason, we will control for parent income in both analyses. The null hypotheses are summarized:

H_{A0}: there is no association between hours of television watching **total per week** on normalized student grades in any subject controlling for parent income
H_{B0}: there is no association between hours of television watching **after school** on normalized student grades in any subject controlling for parent income

The tv watching variables are numerical hours and the parent income is numerical dollars. We will apply multilinear regression. By observing the estimate for the linear regression coefficients and their respective t-score, we can test the hypothesis of the effect of tv watching at different times on student educational outcomes.

Predictor Variables
We will use the sum of the television watching predictor variables for total weekly television watching hours. We will use P5TVW3DH TV BETWEEN 3PM AND DINNER as after school television watching hours.

848 P5TVBF8H	P5 HEQ065A TV WATCHING BEFORE 8 AM - HR
850 P5TVW83H	P5 HEQ065B TV BETWEEN 8AM AND 3PM - HR
852 P5TVW3DH	P5 HEQ065C TV BETWEEN 3PM AND DINNER-HR
854 P5TVAFDH	P5 HEQ065D TV WATCHING AFTER DINNER - HR
856 P5TVSATH	P5 HEQ070A TV WATCHING ON SATURDAY - HR
858 P5TVSUNH	P5 HEQ070B TV WATCHING ON SUNDAY - HR

Confounding variables
Prior research indicates that parent income has associations with family behavior which indirectly changes child education outcomes[2]. As television watching is a result of home behavior, we believe parent income might be a confounder and we will control for parent

income. While there may exist other confounders, at this time we have not gathered evidence to suggest we should control for other confounders.

We will assess the assumptions of the model by looking at the residuals plots. We should see if the distribution of residuals is constant, normal, and linear over fitted values. We know that the dataset is very large (>10^4 participants), so it will likely be enough data that we do not have to be concerned about the normality assumption.

We will report the results as the linear regression coefficients between weekly television watching and after school television watching with the math, science, and reading test T-scores. These coefficients can be thought of as the average change in points on math, reading, and science T-score associated with each hour tv watching in the interval, given the same parent income of the student.

Result

We found a statistically significant negative relationship between weekly television watching hours and student T-scores in all three subject exams. Each additional hour of tv watched per week is associated with an average 0.05 to 0.07 drop in mean test scores for a given income. Income has a statistically significant positive relationship with T-scores.

Math exam T-scores

	Estimate	Std Error	t value	p-value	significant
weekly tv watching hours	-5.35E-02	1.93E-02	-2.8	0.005	*
income	7.42E-05	2.06E-06	36.1	<0.0001	*

Reading exam T-scores

	Estimate	Std Error	t value	p-value	significant
weekly tv watching hours	-6.60E-02	2.17E-02	-3.1	0.002	*
income	8.20E-05	2.30E-06	35.6	<0.0001	*

Science exam T-scores

	Estimate	Std Error	t value	p-value	significant
weekly tv watching hours	-5.81E-02	1.98E-02	-2.9	0.003	*
income	7.95E-05	2.10E-06	37.8	<0.0001	*

Figure 1. Linear regression coefficients and t-test significance for total weekly tv watching hours on (a) math exam T-scores (b) reading exam T-scores and (c) science exam t-scores, controlling for parent income.

Assessing the linear regression assumptions (Appendix B), the residuals to be linear, normal, and constant variance in relationship to fitted values. Banded patterns in the residual plots is likely do to the large jumps in acceptable responses for reported parent income.

We found a statistically significant relationship between after school television watching hours and student T-scores in all three subject exams. Each additional hour of tv watched after school is associated with an average 0.23 to 0.37 drop in mean test scores for a given income. We reject the null hypothesis that after school television watching does not have an association with exam T-score. Income is significantly significant and positive association with exam T-score.

Math exam T-scores

	Estimate	**Std Error**	**t value**	**p-value**	**significant**
after school tv watching hours	-2.34E-01	9.59E-02	-2.4	0.015	*
income	7.43E-05	2.06E-06	36.1	<0.0001	*

Reading exam T-scores

	Estimate	**Std Error**	**t value**	**p-value**	**significant**
after school tv watching hours	-3.76E-01	1.07E-01	-3.5	0.0004	*
income	8.20E-05	2.30E-06	35.6	<0.0001	*

Science exam T-scores

	Estimate	**Std Error**	**t value**	**p-value**	**significant**
after school tv watching hours	-2.75E-01	9.80E-02	-2.8	0.005	*
income	7.95E-05	2.10E-06	37.8	<0.0001	*

Figure 2. Linear regression coefficients and t-test significance for after school tv watching hours on (a) math exam T-scores (b) reading exam T-scores and (c) science exam t-scores, controlling for parent income.

Assessing the linear regression assumptions, the residuals (Appendix C) to be linear, normal, and constant variance in relationship to fitted values. Banded patterns in the residual plots is likely do to the large jumps in acceptable responses for reported parent income.

Discussion

Using the ECLS data, we reproduced the result that there is a statistically significant relationship between weekly television watching hours on student grades. In addition, we reproduced the result that there is a statistically significant relationship between after school television watching hours on student grades.

Comparing the linear regression coefficients for the hours of television watched in the two periods examined against each subject exam T-score, we see that the coefficients are all higher for after school television watching. This indicates that each unit hour of after school television watching has a larger average isolated association on grades than each unit hour of weekly television watching. While our study is limited to only examining associations in television watching hours, in conjunction with results from prior work on the positive association between after school activities and student grades, evidence suggests that how student after school time is budgeted has a significant association with education outcomes.

Question 2: Is bedtime associated with changes in child academic performance on Math and Reading comprehension exams? Likewise, is having a consistent bedtime associated with child academic performance?

Studies[6] have shown that both consistent bedtimes and earlier bedtimes are strongly associated with academic performance. We'll look to evaluate these claims using our dataset. We'll evaluate two hypotheses:

A) Is consistency of bedtime associated with stronger academic performance? If so, on average how different are the students' performance on each of the Math, Reading, and Science exams?
   - **H_{A0}**: there is no association between consistency of bedtimes and academic performance (students with consistent bedtimes and inconsistent bedtimes show equivalent academic performance).

B) Is there evidence of an association of bedtimes with student academic performance on Math, Reading Comprehension, and Science exams? If so, what is the nature of the relationship (is earlier or later correlated with higher performance)?
   - **H_{B0}**: there is no association between bedtimes and academic performance of a student (the time the student goes to bed does not have an association with academic performance).

Variables

We use the following predictor variables for our two hypotheses:

A. P5GOTOBD: GO TO BED SAME TIME EACH NIGHT
   - Data type: Binary TRUE/FALSE
   - Description: Parents are asked if student goes to bed at a consistent time on school nights.

B. P5TIMEHR: TIME TO GO TO BED-HOUR
   - Data type: integer value in [6, 12] indicating bedtime
   - Description: Assuming student have a consistent bedtime, parents are asked what time that is. Only includes students with consistent bedtimes.

We use the same aforementioned response variables: T-scores for the math, reading, and science proficiency exams.

We did not find any compelling confounding variables around parent, school, or student factors that would consistently affect certain students' bedtimes. We considered controlling for official school start times, but after a descriptive look at this variable we found the majority of schools start within the same hour (7:45 - 8:45 am).

Method

Hypothesis A: Bedtime Consistency

We use a linear regression to investigate this hypothesis. We could have used a two-sample t-test since we're testing the means of two groups. However, we're also interested in the coefficient estimate for the bedtime consistency variable, since we'd like to report how different student performance is between the two groups - not just that they are different.
A statistically significant coefficient will highlight the direction and magnitude of the students' academic performance for students who have a consistent bedtime compared against students who do not.

Hypothesis B: Bedtime

We use a linear regression and we assume that the relationship between bedtime (by hour) and academic performance is linear. This seems to be a reasonable assumption as bedtimes are very closely correlated with the amount of sleep a student is able to get, and it wouldn't be unreasonable to think academic performance may change at a linear rate with the number of hours of sleep.
A statistically significant coefficient will highlight the direction and magnitude of the mean change in students' performance associated with an hour change in bedtime.

Results

The results of our analysis are summarized below. For each hypothesis in this question, we performed three tests: one for each type of exam (Math, Reading, Science).

A) Results: Bedtime Consistency

	Estimate	Std Error	p-value	significant
Math	2.61	0.315	<0.0001	*
Reading	2.65	0.314	<0.0001	*
Science	3.10	0.317	<0.0001	*

Figure 3. Linear regression coefficient estimates and significance for association of bedtime consistency with (a) math exam T-scores (b) reading exam T-scores and (c) science exam t-scores.

The results show that students who have a consistent bedtime perform significantly better than students who do not, across all proficiency exams. The coefficients in the table above show the difference in mean T-score points by exam.

Now, for those students who did have a consistent bedtime, we'll take a closer look at the relationship between academic performance and earlier vs. later bedtimes…

B) Results: Bedtime

	Estimate	**Std Error**	**p-value**	**significant**
Math	0.250	0.148	0.0927	
Reading	-0.548	0.148	0.0002	*
Science	-0.783	0.149	<0.0001	*

Figure 4. Linear regression coefficient estimates and significance for association of bedtime with (a) math exam T-scores (b) reading exam T-scores and (c) science exam t-scores.

There does not appear to be a significant relationship between performance on math exams and bedtime. Interestingly, this contrasts with Reading and Science.

For reading comprehension exams, there is a significant negative mean association of -0.55 T-score for each hour later the student's bedtime is. Meaning on average, students who went to bed earlier performed better on reading exams. A scatter-plot of the data for Reading T-scores by Bedtime are shown below. The red line shows the fitted linear regression line with negative slope, indicating worse performance for later bedtimes.

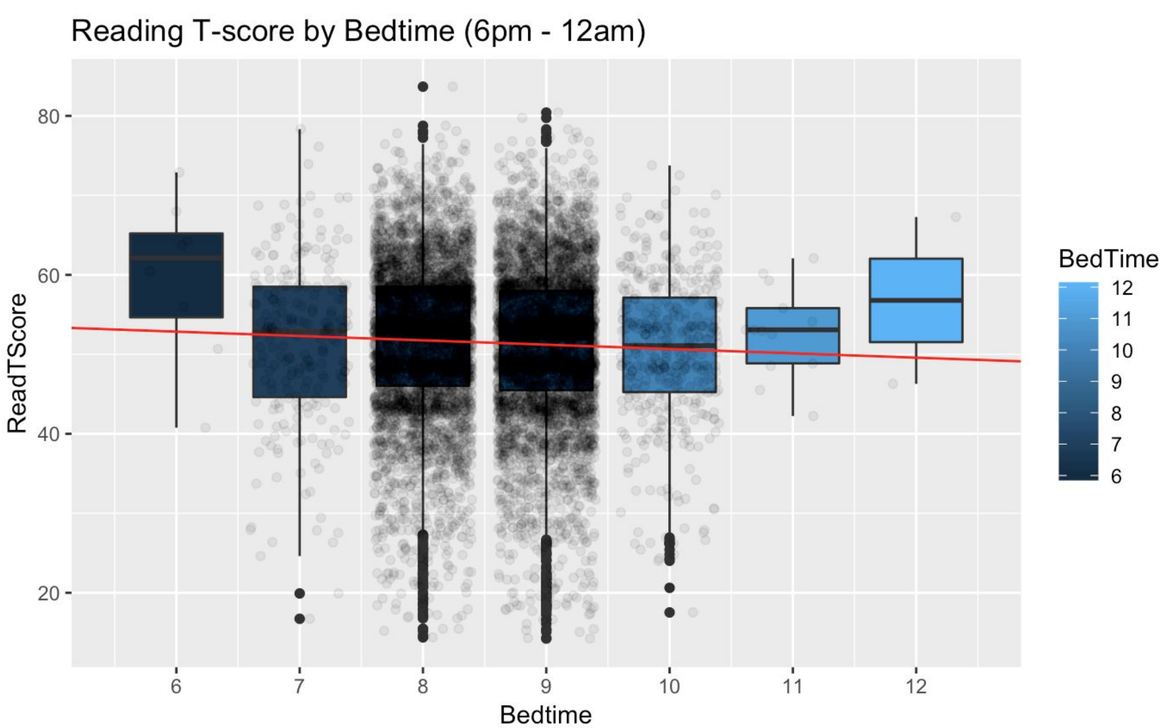


Figure 5. Fitted line regression line over jittered scatterplot data for association between bedtime and reading T-score.

For science exams, there is a significant negative mean association of -0.78 T-score for each hour later the student's bedtime is. Meaning on average, students who went to bed earlier performed better on science exams. A scatter-plot of the data for Science T-scores by Bedtime are shown below. The red line shows the fitted linear regression line with negative slope, indicating worse performance for later bedtimes.

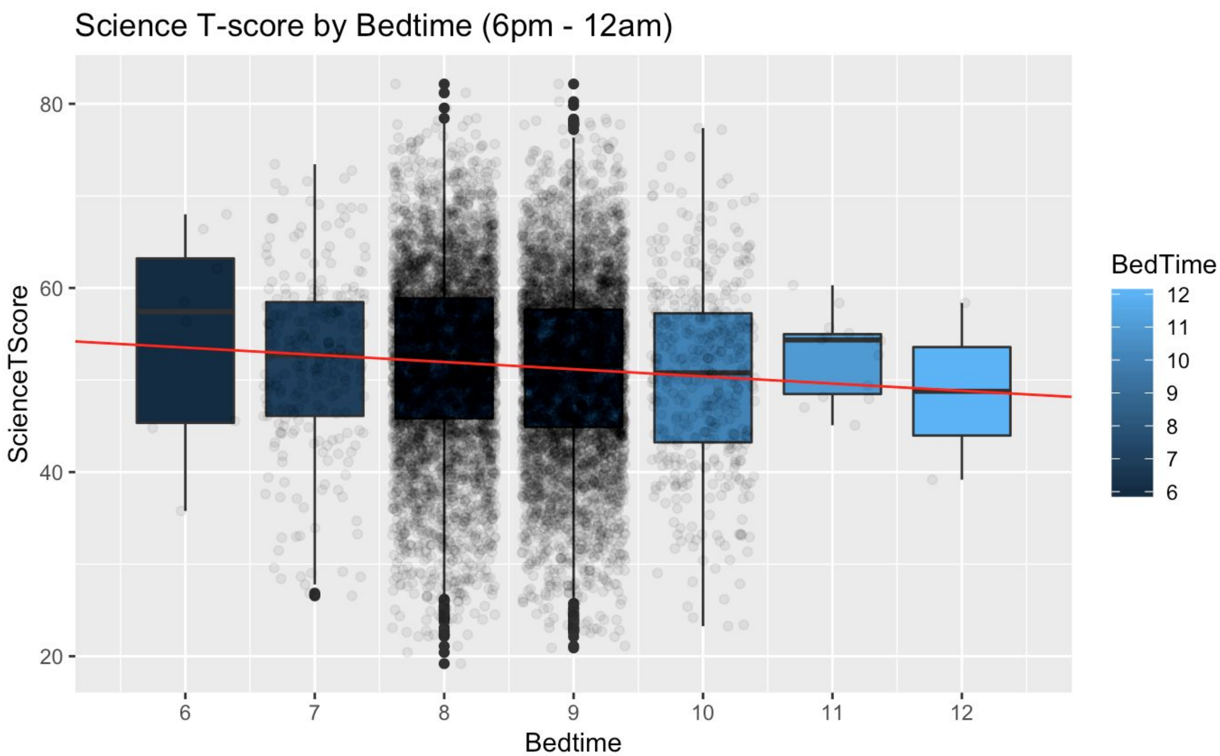


Figure 6. Fitted line regression line over jittered scatterplot data for association between bedtime and science T-score.

Verifying Assumptions (Appendix D,E,F,G,H,I)

- Independence
    - We assumed independence as these survey questions were asked of students and parents on a private basis. Additionally, we have no reason to believe that bedtimes are affected by other students, except in the case where multiple students in the same study live in the same household (e.g. siblings). However, since our dataset is limited to only one grade level, it's unlikely that there's many siblings in the dataset in the same grade and also in the study.
- Normality
    - As shown in the QQ plots of each of our regression residual diagnostic plots, our data very closely approximate the normal distribution. Additionally, we have quite a large sample size (order of magnitude of 10^4).
- Linearity
    - As shown in each of the residual vs. fitted residual diagnostic plots for these hypotheses, the flat line verifies linearity in each of our regressions.

- Constant Variance
    - As shown in each of the scale-location residual diagnostic plots for these hypotheses, the flat line verifies constant variance in each of our regressions.

Discussion

Using the ECLS data, we were able to confirm that having a consistent bedtime is strongly associated with increased academic performance. Interestingly, our results could only partially verify the claims that earlier bedtimes are associated with better academic performance: on Reading and Science exams, but not Math.

A potential area for a follow-up analysis would be to attempt to evaluate what makes Math performance different from Reading and Science. Perhaps the required cognitive skills in each exam are affected differently by sleep times, or some other explanation; we could only find out through more research.

Question 3: Is there a relationship between age(in months) and students' performance?

Hypothesis

The null hypothesis is that there is no difference in academic performance among students of different age. That is, students in the same grade have similar academic performance regardless of how old they are.

The alternative hypothesis then states that there is a difference in performance across students from different age groups.

Variables

Predictor variable		
R5AGE	Child's age at assessment	Measured in months and aggregated into bins.

The predictor variable is the age of students (in months) at the time the survey was conducted. The data is available in different buckets.

- 1: Less than 105 months
- 2: 105 - 108 months
- 3: 108 - 111 months
- 4: 111 - 114 months
- 5: 114 - 117 months
- 6: 117 months or more

The data is largely complete. Minimal cleaning is required.
The response variables are the third grade math, reading and science T-scores.

Method

To test the null hypothesis, linear regression is used to investigate the relationship between math and reading t score and the child's age at assessment. Linear regression is used over ANOVA analysis because, should there be any relationship between the two variables, given the nature of age, the relationship is likely to be linear, that is, either strictly increasing or strictly decreasing.

The regression study will not control for any other variables as it is highly unlikely that the month in which the students were born would have confounding effects with other variables in this study and should be randomly distributed.

We will use the residual plots to assess the assumptions of the model. A flat residuals vs fitted values plot implies that the model's assumptions of linearity is met. The QQ plot will help us assess the normality of the sample and the scale-location plot will help us assess the assumption of constant variance.

We will report the results through the coefficient of age of child and its statistical significance. This can be interpreted as the difference in math, reading and science T-scores associated with a unit difference in age bucket which corresponds to approximately 3 months of age difference.

Results

We found that there are statistically significant regression coefficients between age and all 3 subjects T-score performance.

Therefore, we have sufficient evidence to reject the null hypothesis and conclude that there is association between age and T-score performance.

ProficiencyExam	Estimate	Std. Error	p-value	Significance (0.05)
Math	0.7568247	0.0576887	0	***
Read	0.5747990	0.0579081	0	***
Science	1.0293823	0.0580991	0	***

Figure 7. Linear regression coefficient estimates and significance for association of age at enrollment with (a) math exam T-scores (b) reading exam T-scores and (c) science exam t-scores.

Discussion

The coefficients can be interpreted as follows:

- A unit change in age group (+3 months) is associated with ~0.76 movement math T-score.
- A unit change in age group (+3 months) is associated with ~0.57 movement in reading T-score.
- A unit change in age group (+3 months) is associated with ~1.03 movement in science T-score.

Give that T-scores are normalized to have standard deviation of 10, a unit increase in T-score translates to an improvement of 0.1 standard deviation compared to national average.

Checking Assumptions

The residual plots are used to assess the following assumptions (Appendix J)

- Independence
  - The survey was a stratified cluster sampling on students of the same academic year. Therefore, we have strong reasons to believe that this assumption is sufficiently satisfied.
- Linearity
  - The residuals vs fitted values plot is flat. This implies that the model's assumptions of linearity is met
- Constant variance
  - The scale-location plot is flat. This implies that the variance is constant for all fitted values.
- Normality or large sample size
  - The QQ plot assures that the data is sufficiently normal.

Visualization

- The association between age and math T-scores

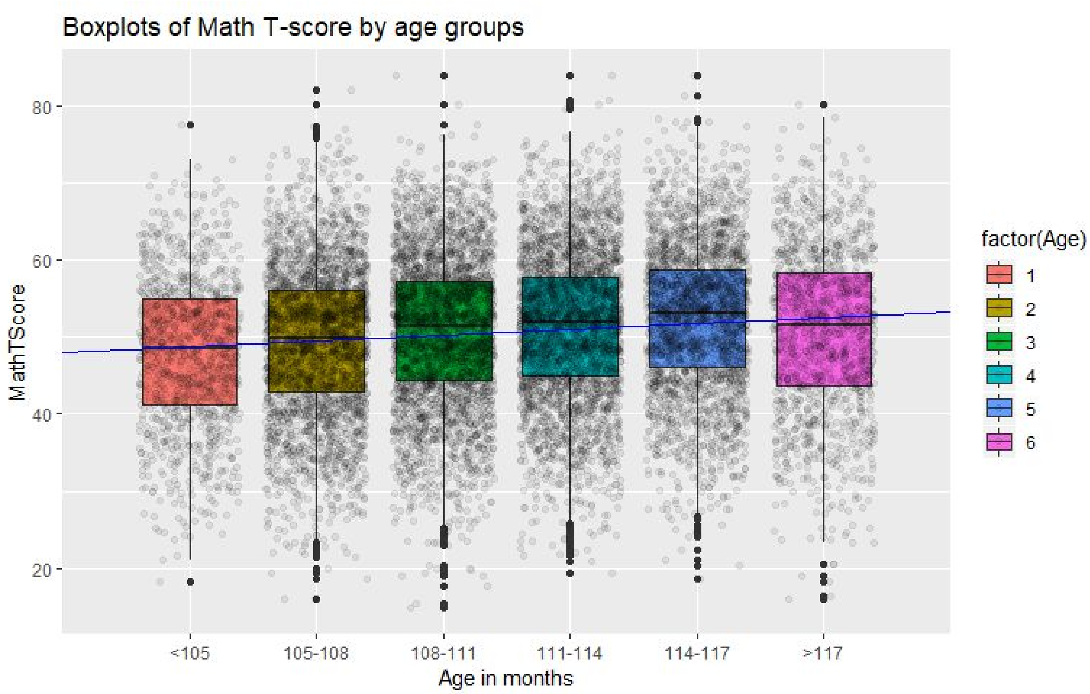


Figure 8. Fitted line regression line over jittered scatterplot data for association between age at enrollment (months) and Math T-score.

- The association between age and reading T-scores

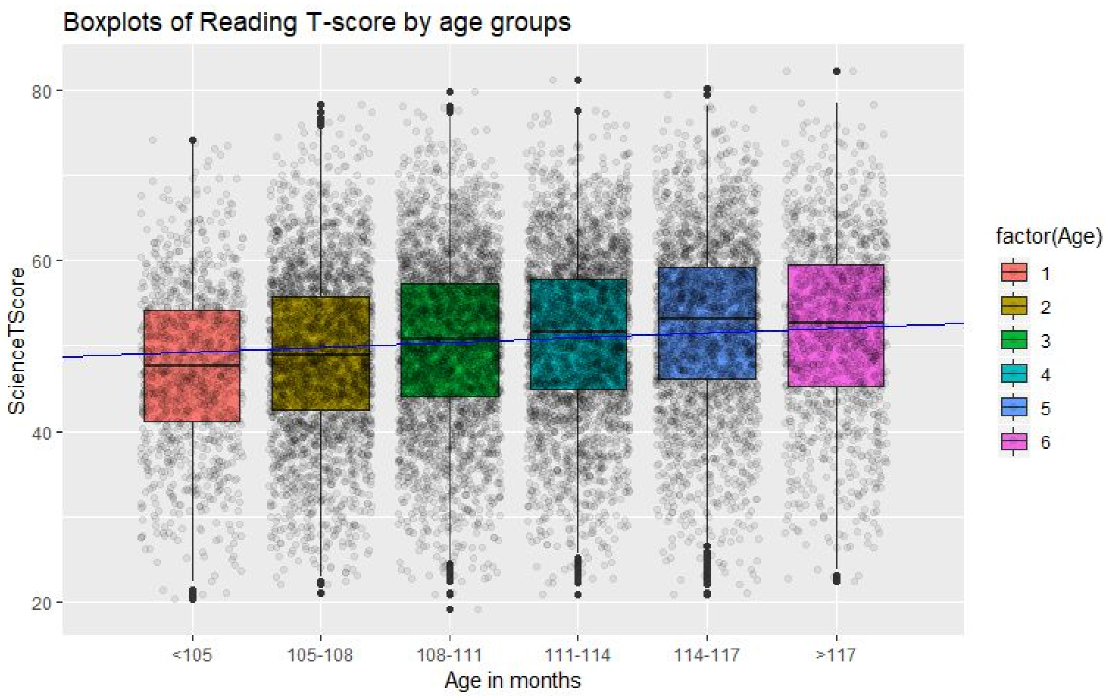


Figure 9. Fitted line regression line over jittered scatterplot data for association between age at enrollment (months) and Reading T-score.

- The association between age and science T-scores

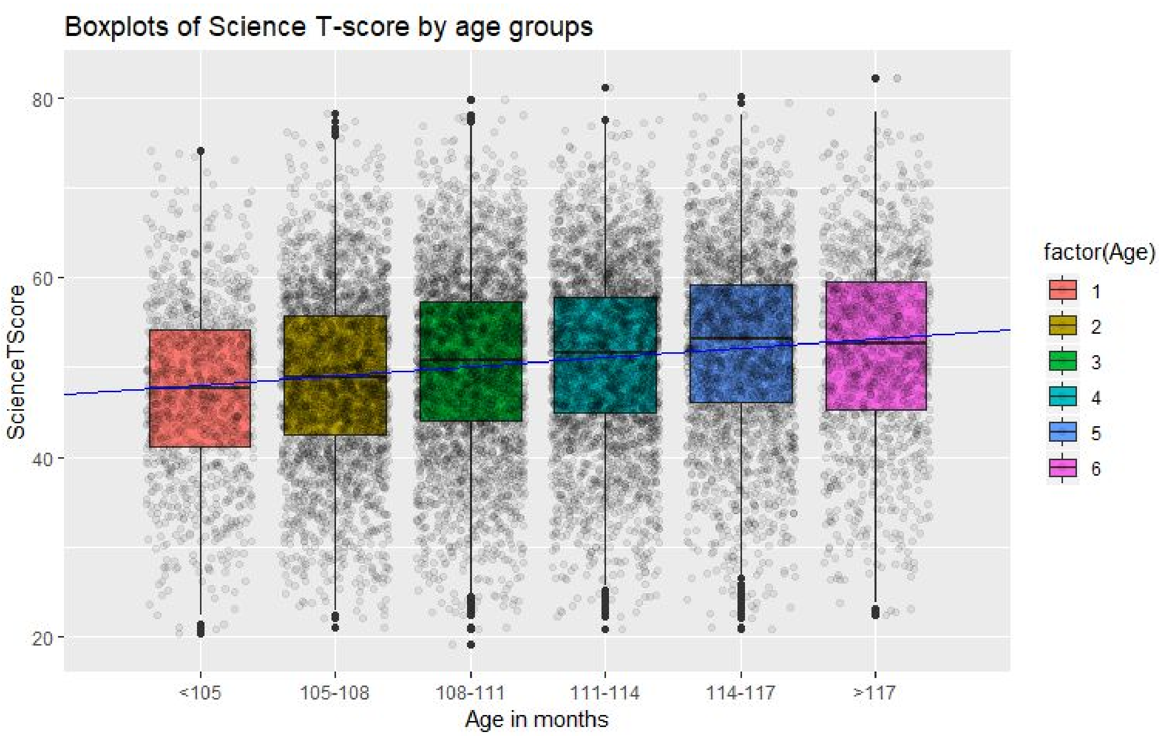


Figure 10. Fitted line regression line over jittered scatterplot data for association between age at enrollment (months) and Reading T-score.

REFERENCES

[1] Barnett, W. Steven. “Long-Term Effects of Early Childhood Programs on Cognitive and School Outcomes.” *The Future of Children*, vol. 5, no. 3, 1995, pp. 25–50. *JSTOR*, www.jstor.org/stable/1602366.

[2] Davis-Kean, P. “*The Influence of Parent Education and Family Income on Child Achievement: The Indirect Role of Parental Expectations and the Home Environment*.” American Psychological Association, American Psychological Association, psycnet.apa.org/buy/2005-06518-016.

[3] Neuman, et al. “Television Viewing and Reading Achievement.” *ERIC*, 31 Mar. 1981, eric.ed.gov/?id=ED200937.

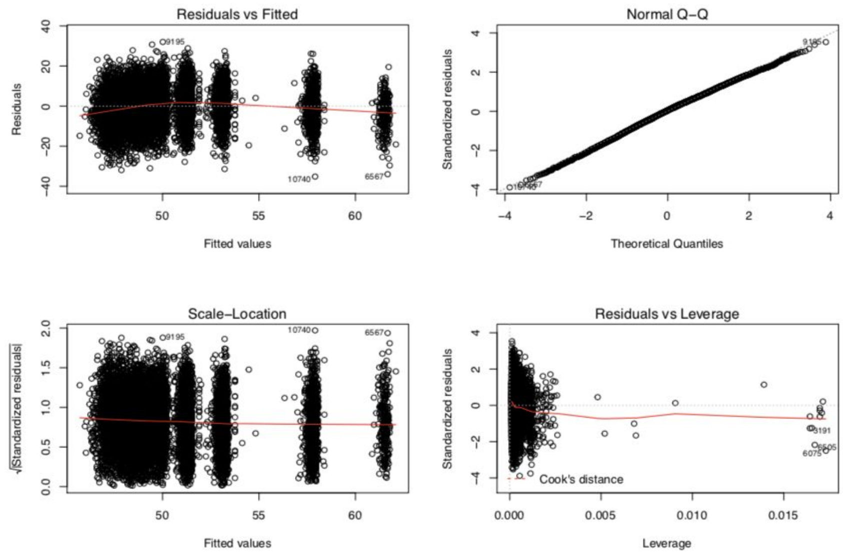
[4] Ridley-Johnson, et al. “The Relation of Children's Television Viewing to School Achievement and I.Q.” *Taylor & Francis*, www.tandfonline.com/doi/abs/10.1080/00220671.1983.10885469.

[5] U.S. Dept. of Education, National Center for Education Statistics. EARLY CHILDHOOD LONGITUDINAL STUDY [UNITED STATES]: KINDERGARTEN CLASS OF 1998-1999, THIRD GRADE [Computer file]. ICPSR version. Washington, DC: U.S. Dept of Education, Institute of Education Sciences [producer], 2004. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2004.

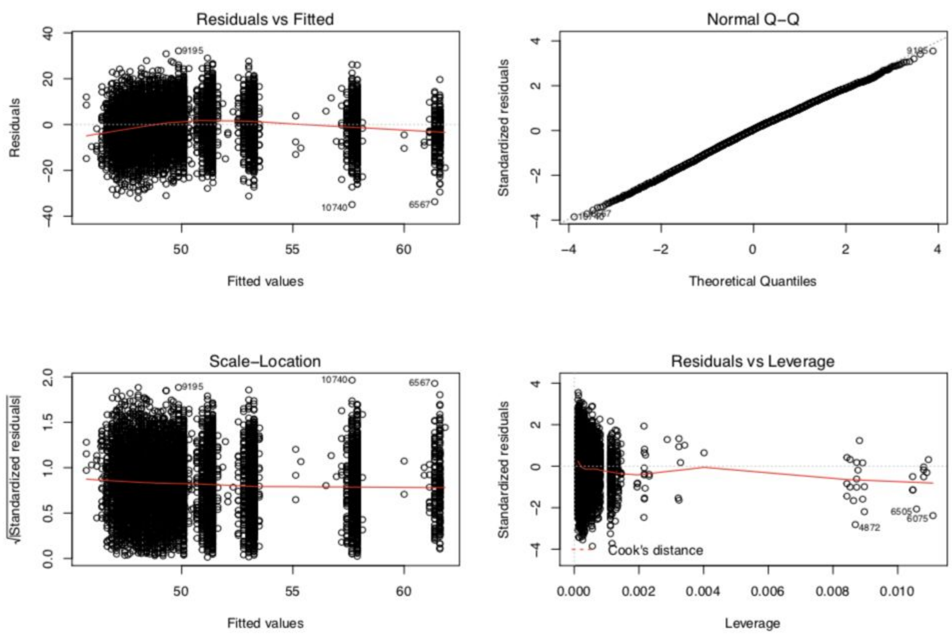
[6] Hysing, M. , Harvey, A. G., Linton, S. J., Askeland, K. G. and Sivertsen, B. (2016), Sleep and academic performance in later adolescence: results from a large population-based study. J Sleep Res, 25: 318-324. doi:10.1111/jsr.12373

TABLES & FIGURES - APPENDICES

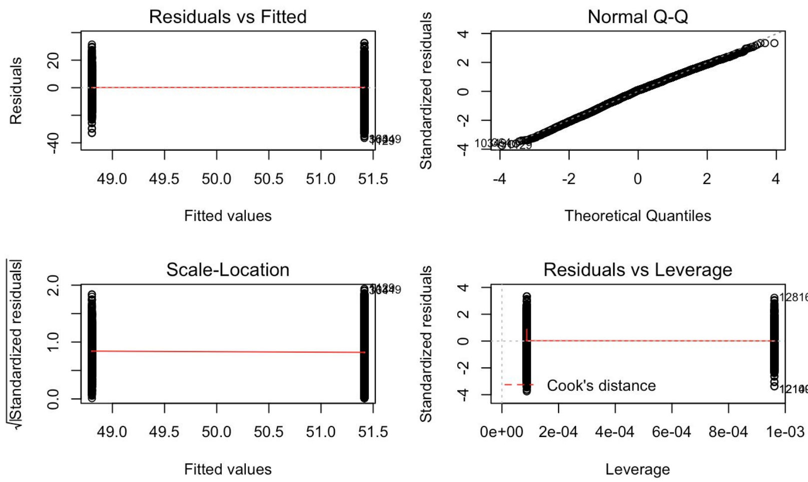
A. **Data Source -** The ECLS dataset is available for download on the ECLS website. https://nces.ed.gov/ecls/dataproducts.asp#K-8
B. Weekly total hours television watching linear regression residuals plots



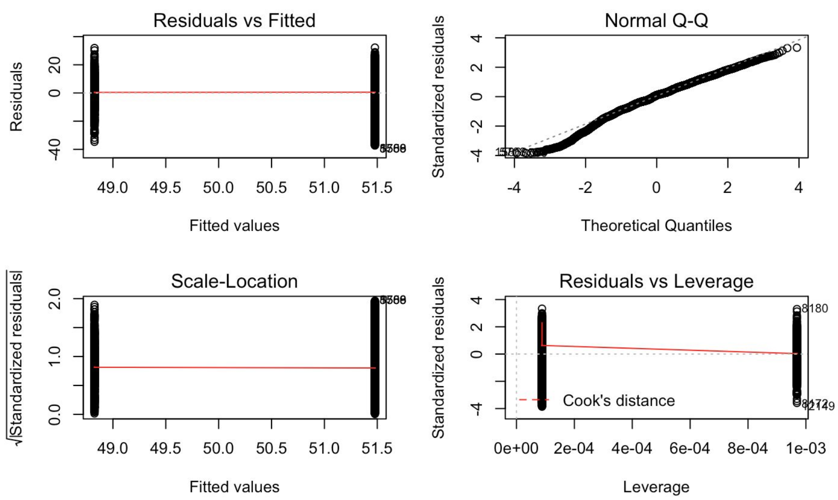
C. After school television watching hours linear regression residual plots



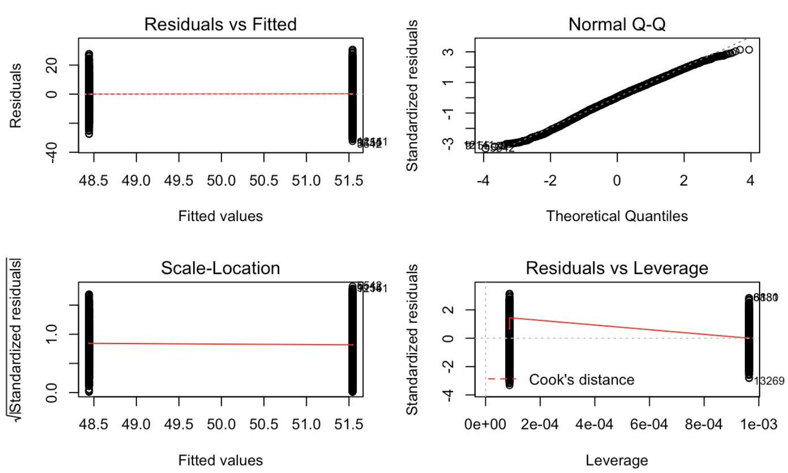
D. Bedtime Consistency: Math (Residual Diagnostic Plots)



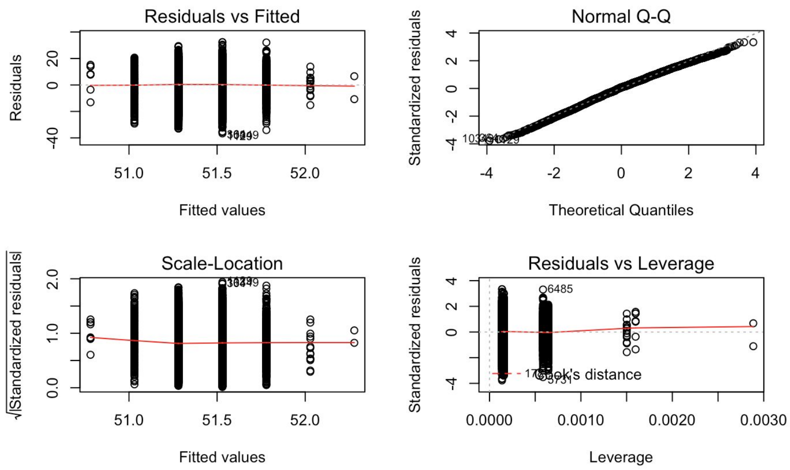
E. Bedtime Consistency: Reading (Residual Diagnostic Plots)



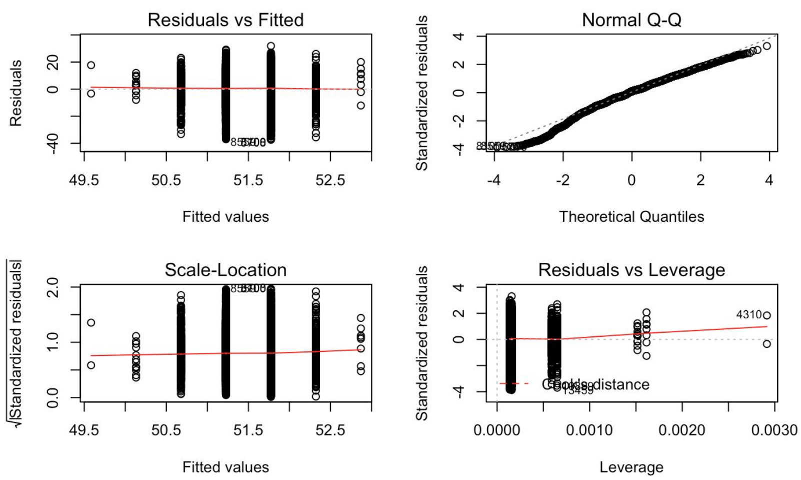
F. Bedtime Consistency: Science (Residual Diagnostic Plots)



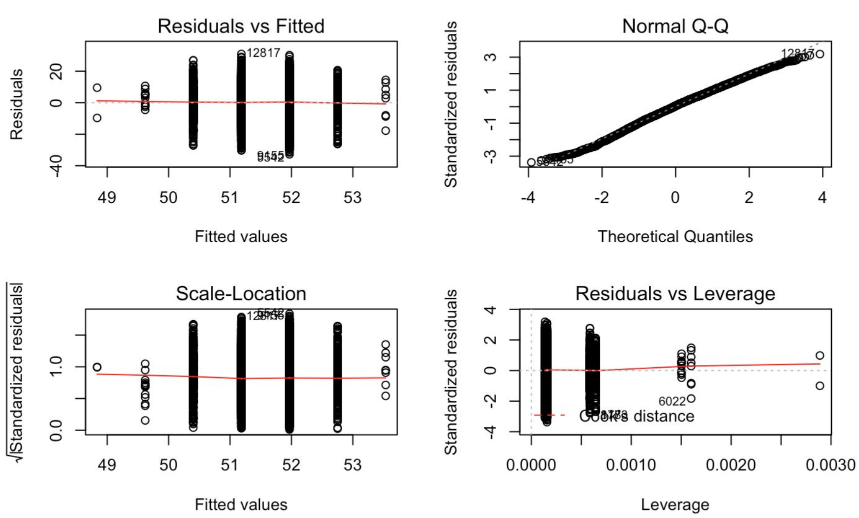
G. Bedtime: Math (Residual Diagnostic Plots)



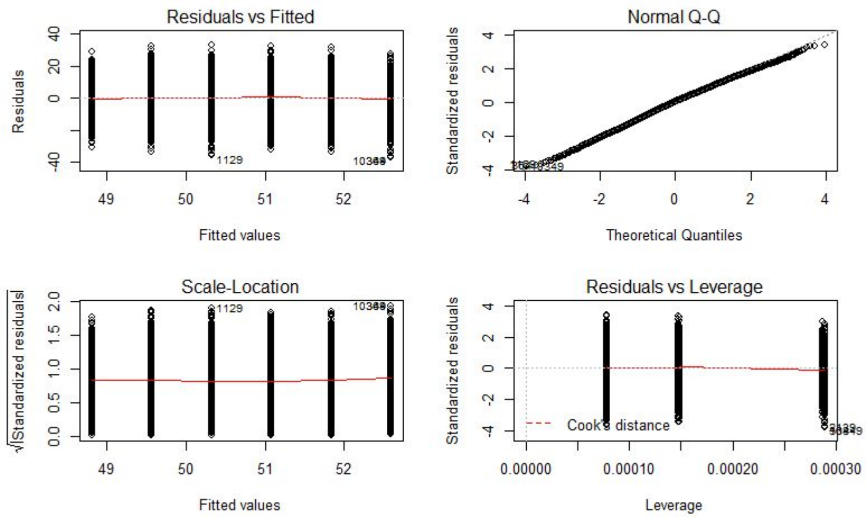
H. Bedtime: Reading (Residual Diagnostic Plots)



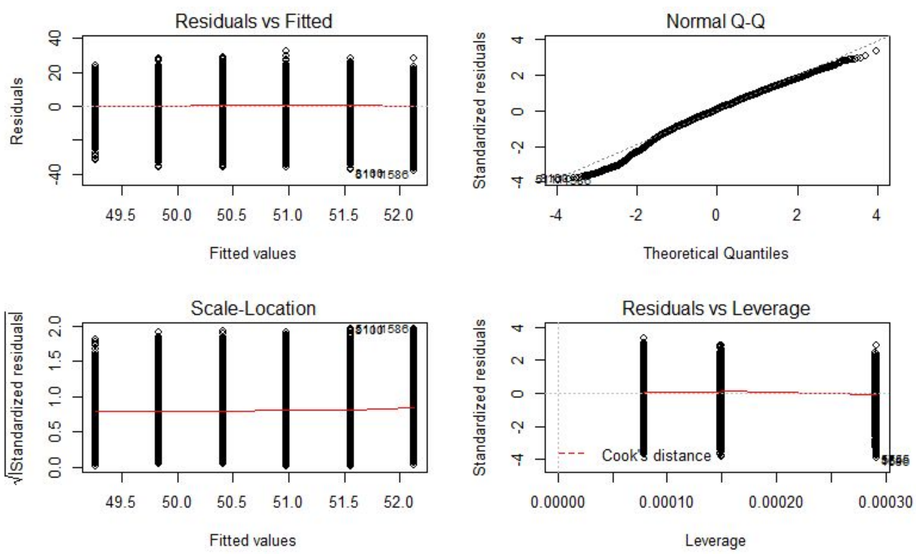
I. Bedtime: Science (Residual Diagnostic Plots)



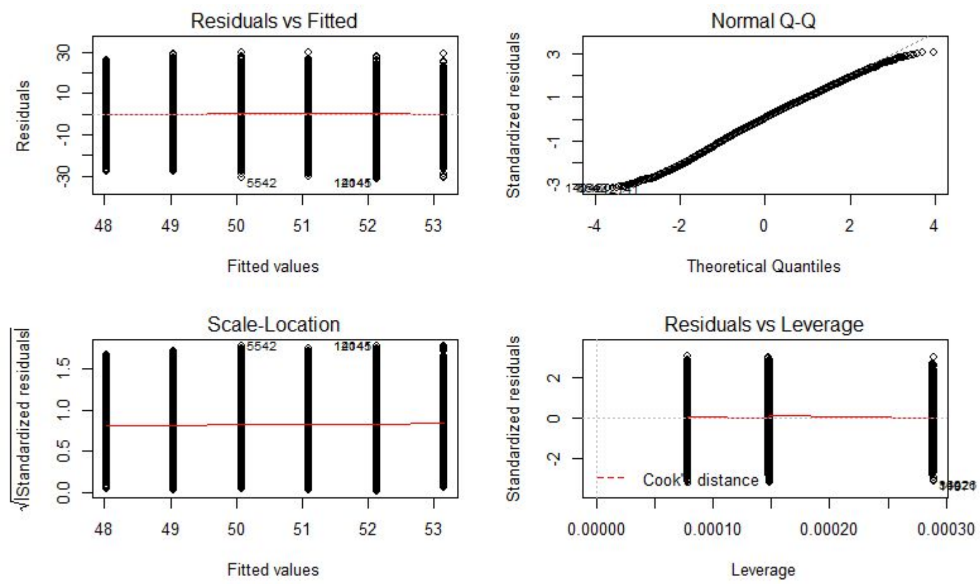
J. (Age) Math T-score residual plots



(Age) Reading T-score residual plots



(Age) Science T-score residual plots



Residuals vs Fitted
Residuals
Fitted values
5542
12045
Normal Q-Q
Standardized residuals
Theoretical Quantiles
Scale-Location
√|Standardized residuals|
Fitted values
Residuals vs Leverage
Standardized residuals
Cook's distance
Leverage