**Documents B: Research Q3 part 2**

**Pre- Regression Analysis**

My supervisors’ comments:

Since the results of the part did not show significance.

Create dummy variables for each of those variables -re-run the pre-regression stats with these new variables.  Calculate the cohen's d effect size for each one.

Run regression with anything either significant or with an effect size of >d=0.5 (medium effect size). Done

 Start explaining with a mind map to conceptualise the results for your discussion- as we talked about a) the language measure not being validated or sensitive; b) the use of secondary data making the all the categories that showed not significance categories and difficult to interpret; c) think about the bigger picture - maybe just living deprived conditions and environments puts children at risk AND/OR maybe parents who can't cope have genetic predisposition for language difficulty.  Think about how your data relates to the literature, where it matches and where it doesn't (I am looking into these stills)

**RQ3 part two: which of the demographic factors predict language difficulties in vulnerable children**

**1.Data Analysis**

The data analysis in part two involved in three phases. Phase one involved running pre-regression analysis via using T-tests and Pearson's correlation via using the language scores established in part one of the RQ3 above and potential predicters.

The first step involved in creating dummy variables for each of the factors (e.g., poor living conditions, low SES conditions, deprived environments, chaotic environments, white ethnicity, Black, mixed ethnicities, female, male, accessing school meals, accessing counsellings, speak second languages, age and time spent in school). Following, pre-regression analysis runs using T-tests and Pearson's correlations to determine whether a difference exists between the means of independent (demographic factors) and dependent variables (total language scores). As with RQ3 part 1, some skewness is expected in data and the Central Limit Theorem is followed when analysing the data. The normality of the data was assessed using Kolmogorov Smirnov’s test. The assumption of homogeneity of variance was evaluated through Levene’s test. Following, two determine which of the factors predicts language difficulties in this cohort.

The second phase of the RQ3 part two involved in running regression analysis between individual language development areas involving syntactic, semantic and social language scores and demographic factors. The researcher followed the same principles in RQ3 step 1 to establish the total language scores for these language areas. This model was established based on the RQ2 ANOVAs and post hoc results for total language scores. The process involved analysing the three language groups (red, amber and green) means/SD and post hoc scores for each question. The results of this process allowed the researcher to obtain the areas of difficulties these children presented and allowed the researcher to create a language developmental map/areas by dividing these 19 questioners into three areas. Based on that 7 of those questions were targeting children’s syntactic language skills, 8 were targeting children’s language and communication, and 4 were targeting children’s social skills. This model help researcher to establish children’s total syntactic language scores, semantic scores and social language scores. When establishing the totals scores for these areas, the researcher followed the same process as in RQ2 and coded the responses as: ‘0 ‘1 and 2, the researcher added these given responses together to establish the total scores for these variables. Based on that, zeros and ones add up as ‘1’ and twos were added up as ‘2’ which help the researcher to establish total syntactic language scores as 14, semantic skills 16 and social language scores as 8.

Following this, pre-regression analyses were run using T-tests and Pearson’s correlations and in doing so, the researcher followed the same process in part 1 when conduction analysis.

According to that, the researcher checked that there not any no outliers in the data through all variables. The data for all variables were not normally distributed, as assessed via histogram and Kolmogorov-Smirnov tests. Where data not distributed normally, Central Limit Theorem followed. Then independent t-test and person correlation were run. Effect sizes were calculated using appropriate guidance, which as mentioned previously. The results of the analysis are provided in table X below, which needs to be interpreted and written.

**2. RQ3 part two phase 1: Process of creating dummy variables and checking normality and skewness of the data**

A mentioned earlier; the first step was to create a dummy variable for each factor via using SPSS. This process involved recoding the original variables into new variables (0/1). The results of this process resulted in 11 new variables (e.g. poor living conditions

\_v2, low SES conditions \_v2, deprived environments -v2, chaotic environments

\_v2, white ethnicity\_v2, Black \_v2, mix ethnicities\_v2, gender (male\_v2 and females\_v2), accessing school meals\_v2, speak the second languages\_v2 and accessing counsellings \_v2, age and time spent in school, which will be used in part two of the RQ3. As with the part of the RQ3, some skewness was expected. Following that, normality tests were run to check whether all the data distributed normal which Kolmogorov- Smirnov normality test was used during this process. The test results revealed that none of the variables showed normal distributions. This is because the study sample was over 50 there; the Kolmogorov- Smirnov test was the best fitting test. The findings of the normality tests and skewness are presented in Table 1.2 below. As the sample size of the current study was large, the central limit theorem guidance followed, and the parametric tests were considered sufficiently robust tests to apply. In accordance with that, the T-tests were conducted to explore if any of the factors and total language scores show significant differences between them, and appropriate guidance used for calculation effect sizes for each of the factors. The P-value is used in statistical procedures, from t-tests to regression analysis which an alpha of 0.05 is used as the cut off for significance and Levene's test for equality of variances was used.

|  |  |
| --- | --- |
| **RQ3, part 2: Pre- Regression normality test results- Language scores and demographic variables (19Q)** |  |
| **Total language scores and demographic variables** | **Groups** | **P (df)** | **Test statistic**  | **p-value** | **Shape of distribution** |
| Language scores | Deprived environments\_v2 | 49 | 0.092 | 0.2 | Normal |
|  | Not deprived environments\_v2 | 29 | 0.192 | 0.008 | Not normal |
| Language scores | Chaotic envirimnets\_v2 | 21 | 0.22 | 0.009 | Not normal |
|  | Not chaotic environments\_v2 | 57 | 0.1 | 0.2 | Normal |
| Language scores | Poor living conditions\_v2 | 36 | 0.147 | 0.048 | Not normal |
|  | Not poor living conditions\_v2 | 42 | 0.155 | 0.012 | Not normal |
| Language scores | Low SES conditions\_v2 | 37 | 0.16 | 0.18 | Not normal |
|  | Not low SES conditions\_v2 | 41 | 0.132 | 0.07 | Not normal |
| Language scores | White ethnicity | 24 | 0.197 | 0.017 | Not normal |
|  | No White ethnicity | 54 | 0.124 | 0.038 | Not normal |
| Language scores | Black\_v2 | 27 | 0.154 | 0.099 | Not normal |
|  | Not Black \_v2 | 51 | 0.154 | 0.004 | Not normal |
|  | Other ethnicities  | 18 | 0.167 | 0.198 | Not normal |
|  | Not other ethnicities | 60 | 0.135 | 0.008 | Not normal |
| Language scores | Male\_v2  | 43 | 0.124 | 0.092 | Not normal |
|  | Not male | 35 | 0.148 | 0.05 | Not normal |
| Language scores | Female\_v2 | 35 | 0.148 | 0.05 | Not normal |
|  | Not female | 43 | 0.124 | 0.092 | Not normal |
| Language scores | Receiving school meals | 18 | 0.106 | 0.002 | Not normal |
|  | Not receiving school meals | 60 | 0.258 | 0.089 | Not normal |
| Language scores | Children second languages | 2 | 0.26 | . | Not normal |
|  | Not second languages | 72 | 0.125 | 0.127 | Not normal |
| Language scores | Accessing counselling | 6 | 0.195 | 0.2 | Normal |
|  | Not accessing counselling | 72 | 0.146 | 0.001 | Not normal |
| Language scores | Time spent in schools | 78 | 0.153 | 0.001 | Not normal |
| Language scores | Age | 78 | 0.076 | 0.2 | Normal |

From the above table, it can be channelized that only environments, counselling, and time spent in school were found to reflect a normal shape of distribution when analyzed language scores and others all were not normal. However, all these variables were found to depict insignificant normality as had acquired a p-value of 0.2 > 0.05.

**2.1 Pre regression t- tests and correlation analysis results for total language scores and demographic factors**

This process involved using numerous parametric tests to determine whether the difference between these IVs and DV was statistically significant. As seen in Table 3 below, the data shows where there are no significant differences between demographic factors variables as related participation’s language scores. The results of the statistical data analyses are presented in table 2 below after which the results are interpreted and discussed extensively.

|  |  |
| --- | --- |
| **Table X. RQ3 part 2 phase 1- Pre regression T- tests for and correlation analysis results for total language scores and demographic factors (19Q)** |  |
| **Groups** | **Df** | **Mean** | **SD** | **t** | ***p*** | ***Effect Sizes***  | **Skewness** | Levene's p |
| Deprived environments\_v2 | 49 | 18.78 | 3.508 | t (76) = 2.867 | 0.005 | d = 0.701718 | 0.322 | 0.038 |
| Not deprived environments\_v2 | 29 | 16.66 | 2.439 |  |  |  | 0.496 |  |
| Chaotic envirimnets\_v2 | 21 | 16.62 | 2.418 | t (76) = -2.282 | 0.025 |  0.627234. | 0.331 | 0.065 |
| Not chaotic environments\_v2 | 27 | 18.49 | 3.454 |  |  |  | 0.424 |  |
| Poor living conditions\_v2 | 36 | 18.14 | 3.523 | t (76) = 0.374 | 0.71 | d = 0.083955 | 0.555 | 0.442 |
| Not poor living conditions\_v2 | 42 | 17.86 | 3.136 |  |  |  | 0.557 |  |
| Low SES conditions\_v2 | 37 | 17.76 | 3.235 | t (76) = -.583 | 0.562 | d = 0.132896 | 0.616 | 0.773 |
| Not low SES conditions\_v2 | 41 | 18.2 | 3.385 |  |  |  | 0.527 |  |
| White ethnicity | 24 | 17.33 | 2.582 | t (76) = -1.169 | 0.246 | d = 0.305609 | 0.143 | 0.174 |
| No White ethnicity | 54 | 18.28 | 3.558 |  |  |  | 0.515 |  |
| Black  | 27 | 18.22 | 3.609 | t (76) = 0.455 | 0.65 | d = 0.106192 | 0.468 | 0.449 |
| Not Black  | 51 | 17.86 | 3.156 |  |  |  | 0.616 |  |
| Mix ethnicities  | 18 | 18.67 | 3.97 | t (76) = 0.996 | 0.323 |  | 512 | 0.173 |
| Not mix ethnicities | 60 | 17.78 | 3.081 |  |  |  | 419 |  |
| Male\_v2  | 43 | 17.7 | 3.136 | t (76) = -.857 | 0.394 | d= 0.192415 | 0.386 | 0.741 |
| Not male | 35 | 18.34 | 3.506 |  |  |  | 0.687 |  |
| Female\_v2 | 35 | 18.34 | 3.506 | t (76) = .857 | 0.394 | d= 0.192415. | 0.687 | 0.741 |
| Not female | 43 | 17.7 | 3.136 |  |  |  | 0.386 |  |
| Receiving school meals | 18 | 17.89 | 3.546 | *t* (74) = -.240 | 0.811 | d = 0.056106 | 1.118 | 0.053 |
| Not receiving school meals | 60 | 18.1 | 3.93 |  |  |  | 0.393 |  |
| Children with second languages | 2 | 19.5 | 2.121 | *t* (76) = 0.354 | 0.515 | d = 0.555686 | . | 1.075 |
| Not with second languages | 76 | 17.95 | 3.326 |  |  |  | 0.59 |  |
| Accessing counselling | 6 | 20.83 | 4.761 | *t* (71) = 2.116 | 0.038 | d = 0.361688. | -0.557 | 0.241 |
| Not accessing counselling | 72 | 17.76 | 3.092 |  |  |  | 0.56 |  |

Regarding poor living conditions, there were n= 49 participants living in poor living conditions and not poor living conditions (n=29), independent t-test run to see if there differences in total language scores to poor living conditions and not poor living conditions. Data are mean ± standard deviation unless otherwise stated. There were no outliers in the data, as inspected through boxplots and data. The data were not normally distributed, as assessed via histogram and Kolmogorov-Smirnov test (p < .05). As assessed by Levene's test for equality of variances, there was homogeneity of variances (p = .065). The mean differences between participant in poor living conditions (M=18.78, SD = 3.508) and low SES conditions (M = 16.66, SD = 2.439 were not too big. The independent sample t-tests (two-tailed) results showed that there no statistically significant difference between total language scores and poor living conditions, t (76) = 2.867, p = 0.005). The effect sizes were calculated, and they were found to be as d 0.701718., which is a small effect size based on the Cohen’s d guidelines.

In regard to deprived environments, an independent-samples t-test was run to determine if there were differences in **total language scores to deprived environments and not deprived environments v**ariables. There were 36 living deprived environments participants and 42, not deprived environments participants. The visual inspection of the boxplots and Q-Q Plots showed that were no outliers in the data. Total language scores for each variable were not normally distributed, as assessed by the Kolmogorov-Smirnov test (p > .05), and there was homogeneity of variances, as assessed by Levene's test for equality of variances (p = 0.442). The mean total language score for to deprived environments participants (M = 18.14, SD = 3.523) than not deprived environments participants (M = 17.86, SD = 3.136), and no statistically significant difference were found, t (76) = 0.374, p = 0.710. The calculated and found to be as d = 0.083955., which is a small effect size based on Cohen’s d guidelines

To determine whether the **total language scores of children who had accessed** school meals (n=18) differed from children who did not access, an independent sample t-test was performed (n=58). A boxplot analysis revealed that there were no outliers in the data. Kolmogorov- Smirnov test (p < .05) results indicated that the total language scores and accessing accessed school meals were not normally distributed. The Levene's test of homogeneity for equality of variances found that variances were homogeneous (p = 0.053). T-test results indicated that the mean total language score for accessing accessed school meals

 (M = 17.89, SD = 3.546 was not statistically different from that of children who did not (M = 18.10, SD = 3.930), with a none statistically significant effect, t (74) = -.240p = 0.811. The calculated and found **to be as d = 0.056106., which is a small effect size based on the Cohen’s d guidelines.**

In order to assess if there were differences in the total language scores of children living in chaotic (n = 21) and non-chaotic environments (n = 27), an independent t-test was conducted.  Levene's test for equality of variances suggested that there was homogeneity of variances (p = .038). The mean difference between the total language scores of children living in chaotic (M = 16.62; SD = 2.418) and non-chaotic environments were revealed (M = 18.49; SD = 3.454) to be not so big. The independent sample t-tests (two-tailed) results showed that there is no statistically significant difference between total language scores for children living in the chaotic and non-chaotic environments, t (76) = -2.282; p = 0.025. The effect sizes were calculated, and they were found to be as d = 0.627234, which is a medium effect size based on the Cohen’s d guidelines. Hence, it can be asserted that the total language scores of children living in chaotic environments are moderately affected.

An attempt was made to analyze if there were differences in the total language scores of children living in white ethnicity (n = 24) and no white ethnicity (n = 54), an independent t-test was conducted. Levene's test for equality of variances suggested that there was homogeneity of variances (p = .174). The mean difference between the total language scores of children living in white ethnicity (M = 17.33; SD = 2.582) and no white ethnicity was revealed (M = 18.28; SD = 3.558) to be not so big. The independent sample t-tests (two-tailed) results showed that there is no statistically significant difference between total language scores for children living in white ethnicity and no white ethnicity, t (76) = -1.169; p = 0.246. The effect sizes were calculated, and they were found to be as d = 0.305609, which is a small effect size based on Cohen’s d guidelines. Thus, it can be asserted that the total language scores of children living in white ethnicity affected at a low level.

**Correlations: Total language scores, time spent in school and age**

|  |  |  |
| --- | --- | --- |
| **Correlations Language scores (19Q) pre-reg analysis** |  |  |
|   |   | Total Language scores  | Time spent in schools | Age |
|  Total language scores | Pearson Correlation | 1 | -0.185 | -0.029 |
|  | Sig. (2-tailed) |  | 0.106 | 0.799 |
|  | N | 78 | 78 | 78 |
| Time spent in schools | Pearson Correlation | -0.185 | 1 | .306\*\* |
|  | Sig. (2-tailed) | 0.106 |  | 0.006 |
|  | N | 78 | 78 | 78 |
| Age | Pearson Correlation | -0.029 | .306\*\* | 1 |
|  | Sig. (2-tailed) | 0.799 | 0.006 |  |
|   | N | 78 | 78 | 78 |
| \*\* Correlation is significant at the 0.01 level (2-tailed). |  |  |

From the above table, it can be highlighted that the relationship between total language scores and time-spent in school was measured. Here r = -0.185; p-value = 0.106 < 0.05. Hence, it can be outlined that the correlation between total language scores and time spent in schools is negative which is statistically insignificant. This also indicates that an increase in total language scores will result in a decrease in time spent in schools and vice-versa if it would have been statistically significant. Thus, it can be highlighted that the relationship between total language scores and time spent in school is statistically insignificant.

Furthermore, it can be delineated that the Pearson correlation value acquired for gauging the relationship between total language scores and age is r = -0.029; p – value = .799 > 0.05. As the p-value is more than 0.05 level of significance, the relationship is insignificant. Moreover, a negative correlation coefficient depicts that an increase in total language scores may be observed due to a decrease in age had the relationship been significant.

Finally, the relationship between age and time-spent in school is measured. Here r = .306\*\*; p-value = 0.006 < 0.05 and 0.01. Thus, it can be delineated that the correlation between age and time spent in schools is positive which is statistically significant. This also illustrates that an increase in age will result in an increase in time spent in schools and vice-versa. Moreover, it can be concluded that the relationship between age and time spent in school is statistically significant

**RQ3 part two of the phase 2: How do demographic factors related model of language development involving syntactic, semantic and social language scores and demographic factors**

As mentioned earlier, the model of language development groups (syntax, semantic and social language) was established based on the finding of the RQ2. These processes permitted the researcher to establish the total language scores for syntactic and semantic, and social language scores. Following that, the researcher run pre-regression analysis, which the process involved in running numerous parametric tests to check whether there are any statistically significant differences between the means of children’s total language scores of syntactic language skills, semantic and social language skills (DVs) and demographic factors (e.g poor living conditions \_v2, low SES conditions \_v2, deprived environments -v2, chaotic environments \_v2, white ethnicity\_v2, Black \_v2, mix ethnicities\_v2, gender (male\_v2 and females\_v2), accessing school meals\_v2, speak second languages\_v2 and accessing counsellings \_v2, age and time spent in school).

In order to determine whether there were significant differences between the demographic factors, which were established in part 1 of the RQ3 (see the section called Pre-regression analysis) and these three developmental areas, the following steps were undertaken. Same procedures were applied and followed as with part 1. First, skewness and outlier check and no outliers were observed in the data. The normality of the data was checked via Kolmogorov Smirnov’s, and the majority of the data were not normally distributed. As with the RQ3 part 1 and RQ3 part 2 phase 1, the Central Limit Theorem principle is followed as guidance, and parametric tests of analyses were used through the analysis. The homogeneity of variances was assessed by Levene's test for each test to see if the assumption of homogeneity of variances is met (p > .05). Following, the parametric test of T-tests was run to determine if a difference exists between the means of 11 demographic factors (IVs) and dependent variables, namely total language scores of syntactic language skills, semantic and social l language skills (DVs). The results of the statistical data analyses are presentedin table X1 below, after which the results are interpreted and discussed extensively.

**1. Pre-Regression analysis involving total syntactic language scores and demographic variables**

In order to find out if there is any difference exists between the total syntactic language scores and all of the demographic variables, independent T-tests run. Priory to the T-tests, normality tests were to see if the variables distributed normally (p > .05). As mentioned earlier mentioned some skewness previously were expected, the recommended sample sizes of the current study were larger than the threshold (n=78) for the normality tests. Therefore, the Central Limit Theorem was applied when conducting analysis. Even though the data were not distributed normally, parametric tests were applied when exploring the difference between total syntactic language scores and the demographic variables. Aside from this, the homogeneity of the variable was checked via Levene's test for equality of Variances (p > .05). The results of the statistical data analyses are presented in table 3 below, after which the results are interpreted and discussed extensively.

|  |
| --- |
| **Table X1. RQ3, part 2: Pre- Regression normality test results- syntactic language scores and demographic variables**  |
| **Total language scores and demographic variables** | **Groups** | **df** | **Test statistic**  | **p-value** | **Shape of distribution** |
| Syntax language scores | Deprived environments\_v2 | 49 | 0.178 | 0.001 | Not normal |
|  | Not deprived environments\_v2 | 29 | 0.186 | 0.012 | Not normal |
| Syntax language scores | Chaotic envirimnets\_v2 | 21 | 0.169 | 0.119 | Not normal |
|  | Not chaotic environments\_v2 | 57 | 0.186 | 0.001 | Not normal |
| Syntax language scores | Poor living conditions\_v2 | 36 | 0.212 | 0.001 | Not normal |
|  | Not poor living conditions\_v2 | 42 | 0.195 | 0.001 | Not normal |
| Syntax language scores | Low SES conditions\_v2 | 37 | 0.197 | 0.001 | Not normal |
|  | Not low SES conditions\_v2 | 41 | 0.183 | 0.001 | Not normal |
| Syntax language scores | White ethnicity | 24 | 0.243 | 0.001 | Not normal |
|  | No White ethnicity | 54 | 0.191 | 0.001 | Not normal |
| Syntax language scores | Black \_v2 | 27 | 0.206 | 0.005 | Not normal |
|  | Not Black \_v2 | 51 | 0.182 | 0.001 | Not normal |
| Syntax language scores | Mix ethnicities  | 18 | 0.148 | 0.2 | Normal |
|  | Not mix ethnicities | 60 | 0.2 | 0.001 | Not normal |
| Syntax language scores | Male\_v2  | 43 | 0.222 | 0.001 | Not normal |
|  | Not male | 35 | 0.146 | 0.057 | Not normal |
| Syntax language scores | Female\_v2 | 35 | 0.146 | 0.057 | Not normal |
|  | Not female | 43 | 0.222 | 0.001 | Not normal |
| Syntax language scores | Receiving school meals | 60 | 0.202 | 0.5 | Not normal |
|  | Not receiving school meals | 18 | 0.192 | 0.001 | Not normal |
| Syntax language scores | Children second languages | 2 | . | . | Not normal |
|  | Not second languages | 76 | 0.198 | 0001 | Not normal |
| Syntax language scores | Accessing counselling | 6 | 0.24 | 0.2 | Not normal |
|  | Not accessing counselling | 72 | 0.183 | 0.001 | Not normal |
| Syntax language scores | Time spent in schools | 78 | 0.153 | 0.001 | Not normal |
| Syntax language scores | Age | 78 | 0.76 | 0.2 | Normal |

From the above table, it can be channelized that only mix ethnicities and age groups were found to depict a normal shape of distribution when analyzed syntax language scores and others all were not normal. Moreover, both these variables of age distribution and mix ethnicities were insignificant as had acquired a p-value of 0.2 > 0.05.

|  |  |
| --- | --- |
| **Table X. RQ3 part 2 phase 2- Descriptive statistic for potential predictors variables of syntactic language scores**  |  |
| **Groups** | **df** | **Mean** | **SD** | **F** | ***t*** | ***Effect Sizes***  | **Skewness** | Levene's p |
| Deprived environments\_v2 | 49 | 7.69 | 2.143 | t (76) = 1.696 | 0.094 | d= 0.130834. | 0.827 | 0.03 |
| Not deprived environments\_v2 | 29 | 6.93 | 1.462 |  |  |  | 0.422 |  |
| Chaotic envirimnets\_v2 | 21 | 7 | 1.549 | t (76) = -1.134 | 0.261 | d = 0.341286. | 357 | 0.137 |
| Not chaotic environments\_v2 | 57 | 7.56 | 2.062 |  |  |  | 923 |  |
| Poor living conditions\_v2 | 36 | 7.81 | 1.997 | t (76) = 1.683 | 0.097 | d = 0.254368 | 0.671 | 0.331 |
| Not poor living conditions\_v2 | 42 | 7.07 | 1.853 |  |  |  | 1.243 |  |
| Low SES conditions\_v2 | 37 | 7.05 | 1.9 | t (76) = -1.552 | 0.125 | d = 0.353217. | 1.358 | 0.467 |
| Not low SES conditions\_v2 | 41 | 7.73 | 1.95 |  |  |  | 0.653 |  |
| White ethnicity | 24 | 7.13 | 1.454 | t (76) = -.863 | 0.391 | d = 0.225191.,  | 0.505 | 0.02 |
| No White ethnicity | 54 | 7.54 | 2.125 |  |  |  | 0.871 |  |
| Black  | 27 | 7.78 | 2.501 | t (76) = 1.219 | 0.227 | d = 0.268386 | 0.719 | 0.001 |
| Not Black  | 51 | 7.22 | 1.566 |  |  |  | 0.7 |  |
| Mix ethnicities  | 18 | 7.39 | 1.852 | t (76) = -.053 | 0.958 | d = 0.015628 | 0.858 | 0.552 |
| Not mix ethnicities | 60 | 7.42 | 1.985 |  |  |  | 1.023 |  |
| Male\_v2  | 43 | 7.44 | 1.856 | t (76) = 0.158 | 0.875 | d= 0.035578 | 0.858 | 0.72 |
| Not male | 35 | 7.37 | 2.073 |  |  |  | 1.023 |  |
| Female\_v2 | 35 | 7.37 | 2.073 | t (76) = -.158 | 0.875 | d= 0.035578 | 1.023 | 0.72 |
| Not female | 43 | 7.44 | 1.856 |  |  |  | 0.858 |  |
| Receiving school meals | 18 | 7.56 | 2.406 | t (76) = 360 | 720 | d =0.089352 | . | 0.159 |
| Not receiving school meals | 58 | 7.33 | 1.8 |  |  |  | 393 |  |
| Children with second languages | 2 | 8 | 0.001 | t (76) = 0.432 | 0.667 | d = 0.816616. | . | 0.053 |
| Not with second languages | 39 | 7.92 | 2.095 |  |  |  | 5.9 |  |
| Accessing counselling | 6 | 9.83 | 3.371 | t (76) = -3.388 | 0.001 | d = 0.816616. | -0.557 | 0.001 |
| Not accessing counselling | 72 | 7.21 | 1.661 |  |  |  | 0.56 |  |

T-test was run to see if there were differences in total syntactic language scores to poor living conditions (N=49) and not poor livin conditions ( n=29). There were no outliers in the data, as inspected through boxplots and data. The data were not normally distributed, as assessed via histogram and Kolmogorov-Smirnov test (p < .05). There was homogeneity of variances, as assessed by Levene's test for equality of variances (p = 0.030). The mean differences between participant in poor living conditions (M=7.69, SD = 2.143) and low SES conditions (M = 6.93, SD = 1.462) were not too big. The independent sample t-tests (two-tailed) results showed that there was no statistically significant difference between total language scores and poor living conditions, t (76) = 1.696, p = .094)—the calculated and found to be as d= 0.130834,which was a small effect size based on the Cohen's d guidelines.

**Correlations**

The relationship between total syntactic language scores, time spent in school and age were investigated via using correlational analysis. Before calculating the correlations, it is necessary to explore if all variables included in the correlation analysis are normally distributed and which was done using a Kolmogorov Smirnov test. Before calculating the correlations, it is necessary to explore if all variables included in the correlation analysis are normally distributed and which was done using a Kolmogorov Smirnov test. Initial analysis revealed that the time spent in school and syntactic language scores showed not a linear relationship and normal distributions, as shown by the Kolmogorov Smirnov test (p <. 05). On the other hand, age and syntactic language scores were showed moderate linearity and distributions between the two variables were normal. Although the population is not normally distributed between the time spent school and t syntactic language scores, Pearson's correlation was chosen as an appropriate test in the current study; the decision was made based on the Central Limit Theorem

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Correlations** |  |  |  |  |
|  |  |  | Syntactic language scores | Time spent in schools | Age |
|  | Syntactic language scores | Pearson Correlation | 1 | -0.116 | -0.03 |
|  |  | Sig. (2-tailed) | 0.313 | 0.797 |
|  |  | N | 78 | 78 | 78 |
|  | Time spent in schools | Pearson Correlation | -0.116 | 1 | .306\*\* |
|  |  | Sig. (2-tailed) | 0.313 |  | 0.006 |
|  |  | N | 78 | 78 | 78 |
|  | Age | Pearson Correlation | -0.03 | .306\*\* | 1 |
|  |  | Sig. (2-tailed) | 0.797 | 0.006 |  |
|  |  | N | 78 | 78 | 78 |
|  | \*\* Correlation is significant at the 0.01 level (2-tailed). |  |

From the above table, it can be highlighted that the relationship between syntactic language scores and time-spent in school was measured. Here r = -0.116; p-value = 0.313 < 0.05. Thus, it can be delineated that the correlation between syntactic language scores and time spent in schools is negative which is statistically significant. This also illustrates that an increase in syntactic language scores will result in an decrease in time spent in schools and vice-versa. Hence, it can be concluded that the relationship between age and time spent in school is statistically significant.

Further, it can be outlined that the Pearson correlation value acquired for estimating the relationship between syntactic language scores and age is r = -0.03; p – value = .797 > 0.05. As the p-value is more than 0.05 level of significance, the relationship is insignificant. Moreover, a negative correlation coefficient depicts that an increase in syntactic language scores may be observed due to a decrease in age had the relationship been significant.

Lastly, the relationship between age and time-spent in school is measured. Here r = .306\*\*; p-value = 0.006 < 0.05 and 0.01. Thus, it can be delineated that the correlation between age and time spent in schools is positive which is statistically significant. This also illustrates that an increase in age will result in an increase in time spent in schools and vice-versa. Moreover, it can be concluded that the relationship between age and time spent in school is statistically significant.

**2.Pre regression analysis involving semantic language scores and demographic variables**

In order to find out if there is any difference exists between the total semantic language cores and all of the demographic variables, independent T-tests run. Prior to the T-tests, normality tests were to see if the variables distributed normally (p > .05). As mentioned earlier, the same process was applied. The results of the statistical data analyses are presented in table 4 below, after which the results are interpreted and discussed extensively

.

|  |  |
| --- | --- |
| **Table X2. RQ3, part 2: Pre- Regression normality test results- semantic scores and demographic variables**  |  |
| **Total language scores and demographic variables** | **Groups** | **df** | **Test statistic**  | **p-value** | **Shape of distribution** |
| Semantic language scores | Deprived environments\_v2 | 49 | 0.173 | 0.001 | Not normal |
|  | Not deprived environments\_v2 | 29 | 0.18 | 0.017 | Not normal |
| Semantic language scores | Chaotic envirimnets\_v2 | 21 | 0.17 | 0.115 | Not normal |
|  | Not chaotic environments\_v2 | 57 | 0.179 | 0.001 | Not normal |
| Semantic language scores | Poor living conditions\_v2 | 36 | 0.227 | 0.001 | Not normal |
|  | Not poor living conditions\_v2 | 42 | 0.157 | 0.011 | Not normal |
| Semantic language scores | Low SES conditions\_v2 | 37 | 0.18 | 0.004 | Not normal |
|  | Not low SES conditions\_v2 | 41 | 0.214 | 0.001 | Not normal |
| Semantic language scores | White ethnicity | 24 | 0.19 | 0.025 | Not normal |
|  | No White ethnicity | 54 | 0.176 | 0.001 | Not normal |
| Semantic language scores | Black African/Caribbean/ Black British\_v2 | 27 | 0.152 | 0.111 | Not normal |
|  | Not Black African/Caribbean/ Black British\_v2 | 51 | 0.182 | 0.001 | Not normal |
|  | Mix ethnicities  | 18 | 0.207 | 0.04 | Not normal |
|  | Not mix ethnicities | 60 | 0.158 | 0.001 | Not normal |
| Semantic language scores | Male\_v2  | 43 | 0.222 | 0.001 | Not normal |
|  | Not male | 35 | 0.17 | 0.012 | Not normal |
| Semantic language scores | Female\_v2 | 35 | 0.17 | 0.012 | Not normal |
|  | Not female | 43 | 0.222 | 0.001 | Not normal |
| Semantic language scores | Receiving school meals | 18 | 0.207 | 0.04 | Not normal |
|  | Not receiving school meals | 58 | 0.186 | 0.001 | Not normal |
| Semantic language scores | Children second languages | 2 | 0.26 |  . | Not normal |
|  | Not second languages | 76 | 0.166 | 0.001 | Not normal |
| Semantic language scores | Accessing counselling | 6 | 0.286 | 0.136 | Not normal |
|  | Not accessing counselling | 72 | 0.17 | 0.001 | Not normal |
| Semantic language scores | Time spent in schools | 78 | 0.153 | 0.001 | Not normal |
| Semantic language scores | Age | 78 | 0.076 | 0.2 | Normal |

From the above table, it can be highlighted that only age was found to reflect a normal shape of distribution when analyzed with semantic language scores and others all were not normal. However, this variable was found to depict insignificant normality as it had acquired a p-value of 0.2 > 0.05.

|  |  |
| --- | --- |
| **Table X. RQ3 part two phase 2- Descriptive statistic for potential predictors variables of semantic language difficulties**  |  |
| **Groups** | **df** | **Mean** | **SD** | **F** | ***t*** | ***Effect Sizes***  | **Skewness** | Levene's p |
| Deprived environments\_v2 | 49 | 7.24 | 1.995 | t (76) = 2.384 | 0.02 | d=0.526732. | 0.549 | 0.011 |
| Not deprived environments\_v2 | 29 | 6.28 | 1.162 |  |  |  | -0.251 |  |
| Chaotic envirimnets\_v2 | 21 | 6.29 | 1.231 | t (76) = -1.823 | 0.072 | d =0.509396. | -0.251 | 0.09 |
| Not chaotic environments\_v2 | 57 | 7.11 | 1.915 |  |  |  | 0.667 |  |
| Poor living conditions\_v2 | 36 | 7.33 | 1.942 | t (76) = 2.098 | 0.039 | d = 0.470509. | 0.748 | 0.18 |
| Not poor living conditions\_v2 | 42 | 6.5 | 1.566 |  |  |  | 0.191 |  |
| Low SES conditions\_v2 | 37 | 6.43 | 1.608 | t (76) = -2.173 | 0.033 | d = 0.494507. | 0.324 | 0.386 |
| Not low SES conditions\_v2 | 41 | 7.29 | 1.861 |  |  |  | 0.741 |  |
| White ethnicity | 24 | 6.92 | 1.792 | t (76) = 0.105 | 0.917 | d = 0.027824. | 0.829 | 0.955 |
| No White ethnicity | 54 | 6.87 | 1.802 |  |  |  | 0.628 |  |
| Black  | 27 | 6.7 | 1.706 | t (76) = -.648 | 0.519 | d = 0.157904. | 0.602 | 0.972 |
| Not Black  | 51 | 6.98 | 1.838 |  |  |  | 0.772 |  |
| Mix ethnicities  | 18 | 7.39 | 2.118 | t (76) = 1.373 | 0.174 | d = 0.346375.  | 0.668 | 0.275 |
| Not mix ethnicities | 60 | 6.73 | 1.666 |  |  |  | 0.741 |  |
| Male\_v2  | 43 | 6.88 | 1.721 | t (76) = -.005 | 0.996 | d= 0.005531 | 0.942 | 0.587 |
| Not male | 35 | 6.89 | 1.891 |  |  |  | 0.35 |  |
| Female\_v2 | 35 | 6.89 | 1.891 | t (76) = 0.005 | 0.996 | d= 0.005531 | 0.35 | 0.587 |
| Not female | 43 | 6.88 | 1.721 |  |  |  | 0.942 |  |
| Receiving school meals | 18 | 6.94 | 1.474 | t (74) = -.106 | 0.916 | d =0.024856. | 0.231 | 0.328 |
| Not receiving school meals | 58 | 6.9 | 1.734 |  |  |  | 0.832 |  |
| Children with second languages | 2 | 8 | 2.828 | t (39) = -.730 | 0.47 | d = 0.393878.  | . | 0.378 |
| Not with second languages | 39 | 7.08 | 1.707 |  |  |  | 0.704 |  |
| Accessing counselling | 6 | 7.5 | 1.643 | t (76) = -.986 | 0.327 | d = 0.425353. | 0.811 | 0.985 |
| Not accessing counselling | 72 | 6.79 | 1.695 |  |  |  | 0.728 |  |

There were no outliers in the data, as inspected through boxplots and data. The data were not normally distributed, as assessed via histogram and Kolmogorov-Smirnov test (p < .05). There was homogeneity of variances, as assessed by Levene's test for equality of variances (p = 0.011). The mean differences between participant in poor living conditions (M=7.24, SD = 1.995) and low SES conditions (M = 6.28, SD = 1.632) were not too big. The independent sample t-tests (two-tailed) results showed that there no statistically significant difference between semantic language and poor living conditions, t (76) = 2.384, p = 0.020). The effect size was calculated, which was found to be as d 0.526732, which is a small effect size based on the Cohen’s d guidelines.

**Correlations**

In regard to poor living conditions, there were n= 49 participants living in poor living conditions and not poor living conditions (n=29), independent t-test was run to see if there were differences in semantic language to poor living conditions and not poor living conditions. Data are mean ± standard deviation unless otherwise stated. The relationship between semantic language, time spent in school and age were investigated via using correlational analysis. Before calculating the correlations, it is necessary to explore if all variables included in the correlation analysis are normally distributed and which was done using a Kolmogrov Smirnov test. Initial analysis revealed that the time spent in school and semantic language scores showed not a linear relationship and normal distributions, as shown by the Kolmogorov Smirnov test (p <. 05). On the other hand, age and social language scores were showed moderate linearity and distributions between the two variables were normal. Although the population is not normally distributed between the time spent school and total semantic language, Pearson's correlation was chosen as an appropriate test in the current study; the decision was made based on the Central Limit Theorem. -

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Correlations |  |  |  |  |
|   |   | Semantic language scores | Time spent in schools | Age |
| Semantic language scores | Pearson Correlation | 1 | -0.078 | 0.121 |
|  | Sig. (2-tailed) |  | 0.498 | 0.292 |
|  | N | 78 | 78 | 78 |
| Time spent in schools | Pearson Correlation | -0.078 | 1 | .306\*\* |
|  | Sig. (2-tailed) | 0.498 |  | 0.006 |
|  | N | 78 | 78 | 78 |
| Age | Pearson Correlation | 0.121 | .306\*\* | 1 |
|  | Sig. (2-tailed) | 0.292 | 0.006 |  |
|   | N | 78 | 78 | 78 |
| \*\* Correlation is significant at the 0.01 level (2-tailed). |  |  |

From the above table, it can be highlighted that the relationship between semantic language scores and time-spent in school was measured. Here r = -0.078; p-value = 0.498 < 0.05. Thus, it can be delineated that the correlation between semantic language scores and time spent in schools is negative which is statistically significant. This also illustrates that an increase in semantic language scores will result in an decrease in time spent in schools and vice-versa. Hence, it can be concluded that the relationship between semantic language scores and time spent in school is statistically significant.

Further, it can be outlined that the Pearson correlation value acquired for estimating the relationship between semantic language scores and age is r = 0.121; p – value = .292 > 0.05. As the p-value is more than 0.05 level of significance, the relationship is insignificant. Moreover, a negative correlation coefficient depicts that an increase in semantic language scores may be observed due to a decrease in age had the relationship been significant.

Finally, the relationship between age and time-spent in school is measured. Here r = .306\*\*; p-value = 0.006 < 0.05 and 0.01. Thus, it can be delineated that the correlation between age and time spent in schools is positive which is statistically significant. This also illustrates that an increase in age will result in an increase in time spent in schools and vice-versa. Moreover, it can be concluded that the relationship between age and time spent in school is statistically significant.

3.**Pre regression analysis involving total social language scores and demographic variables**

In order to find out if there are any differences exists between the social language cores and all of the demographic variables, independent T-tests run. Prior to the T-tests, normality tests were to see if the variables distributed normally (p > .05). As mentioned earlier, the same process was applied. The results of the statistical data analyses are presented in table X3 below, after which the results are interpreted and discussed extensively.

|  |  |
| --- | --- |
| **RQ3, part 2: Pre- Regression normality test results- social language scores and demographic variables**  |  |
| **Total language scores and demographic variables** | **Groups** | **P (df)** | **Test statistic**  | **p-value** | **Shape of distribution** |
| Social language scores | Deprived environments\_v2 | 49 | 0.163 | 0.002 | Not normal |
|  | Not deprived environments\_v2 | 29 | 0.226 | 0.001 | Not normal |
| Social language scores | Chaotic envirimnets\_v2 | 21 | 0.274 | 0.001 | Not normal |
|  | Not chaotic environments\_v2 | 57 | 0.149 | 0.003 | Not normal |
| Social language scores | Poor living conditions\_v2 | 36 | 0.233 | 0.001 | Not normal |
|  | Not poor living conditions\_v2 | 42 | 0.132 | 0.062 | Not normal |
| Social language scores | Low SES conditions\_v2 | 37 | 0.165 | 0.012 | Not normal |
|  | Not low SES conditions\_v2 | 41 | 0.238 | 0.001 | Not normal |
| Social language scores | White ethnicity | 24 | 0.166 | 0.086 | Not normal |
|  | No White ethnicity | 54 | 0.184 | 0.001 | Not normal |
| Social language scores | Black African/Caribbean/ Black British\_v2 | 27 | 0.256 | 0.001 | Not normal |
|  | Not Black African/Caribbean/ Black British\_v2 | 51 | 0.166 | 0.001 | Not normal |
| Social language scores | Other ethnicities  | 18 | 0.245 | 0.006 | Not normal |
|  | Not other ethnicities | 60 | 0.197 | 0.001 | Not normal |
| Social language scores | Male\_v2  | 43 | 0.172 | 0.003 | Not normal |
|  | Not male | 35 | 0.211 | 0.001 | Not normal |
| Social language scores | Female\_v2 | 35 | 0.211 | 0.001 | Not normal |
|  | Not female | 43 | 0.172 | 0.003 | Not normal |
| Social language scores | Receiving school meals | 18 | 0.269 | 0.001 | Not normal |
|  | Not receiving school meals | 60 | 0.149 | 0.002 | Not normal |
| Social language scores | Children speak second languages | 2 | 0.26 |  . | Not normal |
|  | Not second languages | 76 | 0.178 | 0.001 | Not normal |
| Social language scores | Accessing counselling | 18 | 0.269 | 0.001 | Not normal |
|  | Not accessing counselling | 60 | 0.149 | 0.002 | Not normal |
| Social language scores | Time spent in schools | 78 | 0.153 | 0.001 | Not normal |
| Social language scores | Age | 78 | 0.76 | 0.2 | Normal |

From the above table, it can be channelized that only the variable of age found to reflect a normal shape of distribution when analyzed with social language scores and others all were not normal. However, it was found to depict insignificant normality as it had acquired a p-value of 0.2 > 0.05.

**Table X. RQ3 part two phase 2- Descriptive statistic for potential predictors variables of social language difficulties**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Groups** | **df** | **Mean** | **SD** | **F** | ***t*** | ***Effect Sizes***  | **Skewness** | Levene's p |
| Deprived environments\_v2 | 49 | 3.84 | 1.214 | t (76) = 0.772 | 0.442 | d=0.173478. | 0.034 | 0.008 |
| Not deprived environments\_v2 | 29 | 3.59 | 1.637 |  |  |  | 0.57 |  |
| Chaotic envirimnets\_v2 | 21 | 3.43 | 1.568 | t (76) = -1.227 | 0.224 | d =0.296354 | 0.835 | 0.176 |
| Not chaotic environments\_v2 | 57 | 3.86 | 1.302 |  |  |  | 0.068 |  |
| Poor living conditions\_v2 | 36 | 3.39 | 1.248 | t (76) = -2.149 | 0.035 | d = 0.491576 | 0.408 | 0.682 |
| Not poor living conditions\_v2 | 42 | 4.05 | 1.431 |  |  |  | 0.07 |  |
| Low SES conditions\_v2 | 37 | 3.95 | 1.311 | t (76) = 1.233 | 0.221 | d = 0.283976. | -0.052 | 0.233 |
| Not low SES conditions\_v2 | 41 | 3.56 | 1.433 |  |  |  | 0.568 |  |
| White ethnicity | 24 | 3.58 | 1.381 | t (76) = -.681 | 0.498 | d = 0.166124 | 0.406 | 0.848 |
| No White ethnicity | 54 | 3.81 | 1.388 |  |  |  | 0.213 |  |
| Black  | 27 | 3.67 | 1.271 | t (76) = -.356 | 0.723 | d = 0.080773 | 0.324 | 0.49 |
| Not Black  | 51 | 3.78 | 1.447 |  |  |  | 0.228 |  |
| Mix ethnicities  | 18 | 3.83 | 1.383 | t (76) = 0.312 | 0.756 | d = 0.079308.  | -0.267 | 0.913 |
| Not mix ethnicities | 60 | 3.72 | 1.391 |  |  |  | 0.413 |  |
| Male\_v2  | 43 | 3.53 | 1.297 | t (76) = -1.491 | 0.14 | d= 0.341008 | 0.4 | 0.328 |
| Not male | 35 | 4 | 1.455 |  |  |  | 0.061 |  |
| Female\_v2 | 35 | 4 | 1.455 | t (76) = 1.491 | 0.14 | d= 0.341008 | 0.061 | 0.328 |
| Not female | 43 | 3.53 | 1.297 |  |  |  | 0.4 |  |
| Receiving school meals | 18 | 3.33 | 1.283 | t (74) = -1.572 | 0.12 |  d =0.402777. | 0.417 | 0.822 |
| Not receiving school meals | 58 | 3.91 | 1.393 |  |  |  | 0.207 |  |
| Children with second languages | 2 | 3.5 | 0.707 | t (39) = -.247 | 0.806 | d = 0.220031. | . | 0.177 |
| Not with second languages | 39 | 3.74 | 1.371 |  |  |  | 0.426 |  |
| Accessing counselling | 6 | 3.83 | 1.329 | t (76) = 0.165 | 0.87 | d = 0.066085. | 0.207 | 0.922 |
| Not accessing counselling | 72 | 3.74 | 1.394 |   |   |   | 0.417 |   |

Concerning poor living conditions, there were n= 49 participants living in poor living conditions and not poor living (n=29); independent t-test was run to see if there were differences in social language scores to poor living conditions and not poor living conditions. There were no outliers in the data, as inspected through boxplots and data. The data were not normally distributed, as assessed via histogram and Kolmogorov-Smirnov test (p < .05). There was homogeneity of variances, as assessed by Levene's test for equality of variances (p = .0.008). The mean differences between participant in poor living conditions placements (M=3.84, SD = 1.214) and low SES conditions (M = 3.59, SD = 1.637 were not too big. The independent sample t-tests (two-tailed) results showed that there was no statistically significant difference between total social l language scores and poor living conditions t (76) = 0.772, p = 0.442). The effect size was calculated, which was found to be d 0.173478. This is a small effect size based on the Cohen’s d guidelines.

**Correlations**

The relationship between social language scores, time spent in school and age was investigated via using correlational analysis. Before calculating the correlations, it is necessary to explore if all variables included in the correlation analysis are normally distributed and which was done using a Kolmogrov Smirnov test. Initial analysis revealed that the time spent in school and social language scores showed not a linear relationship and normal distributions, as shown by the Kolmogorov Smirnov test (p <. 05). On the other hand, age and social language scores were showed moderate linearity and distributions between the two variables were normal. Although the population is not normally distributed between the time spent school and total social language scores, Pearson's correlation was chosen as an appropriate test in the current study; the decision was made based on the Central Limit Theorem.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Correlations |  |  |  |  |
|   |   | Social language Scores | Time spent in schools | Age |
| Social language scores | Pearson Correlation | 1 | -0.018 | -0.068 |
|  | Sig. (2-tailed) |  | 0.878 | 0.554 |
|  | N | 78 | 78 | 78 |
| Time spent in schools | Pearson Correlation | -0.018 | 1 | .306\*\* |
|  | Sig. (2-tailed) | 0.878 |  | 0.006 |
|  | N | 78 | 78 | 78 |
| Age | Pearson Correlation | -0.068 | .306\*\* | 1 |
|  | Sig. (2-tailed) | 0.554 | 0.006 |  |
|   | N | 78 | 78 | 78 |
| \*\* Correlation is significant at the 0.01 level (2-tailed). |  |  |

From the above table, it can be highlighted that the relationship between social language scores and time-spent in school was measured. Here r = -0.018; p-value = 0.878 < 0.05. Thus, it can be delineated that the correlation between social language scores and time spent in schools is negative which is statistically insignificant. This also illustrates that an increase in social language scores will result in an decrease in time spent in schools and vice-versa if it would have been statistically significant. Thus, it can be highlighted that the relationship between social language scores and time spent in school is statistically insignificant.

Moving on, it can be delineated that the Pearson correlation value acquired for measuring the relationship between social language scores and age is r = -0.068; p – value = .554 > 0.05. As the p-value is more than 0.05 level of significance, the relationship is insignificant. Moreover, a negative correlation coefficient depicts that an increase in social language scores may be observed due to a decrease in age had the relationship been significant.

Finally, the relationship between age and time-spent in school is measured. Here r = .306\*\*; p-value = 0.006 < 0.05 and 0.01. Thus, it can be delineated that the correlation between age and time spent in schools is positive which is statistically significant. This also illustrates that an increase in age will result in an increase in time spent in schools and vice-versa. Moreover, it can be concluded that the relationship between age and time spent in school is statistically significant