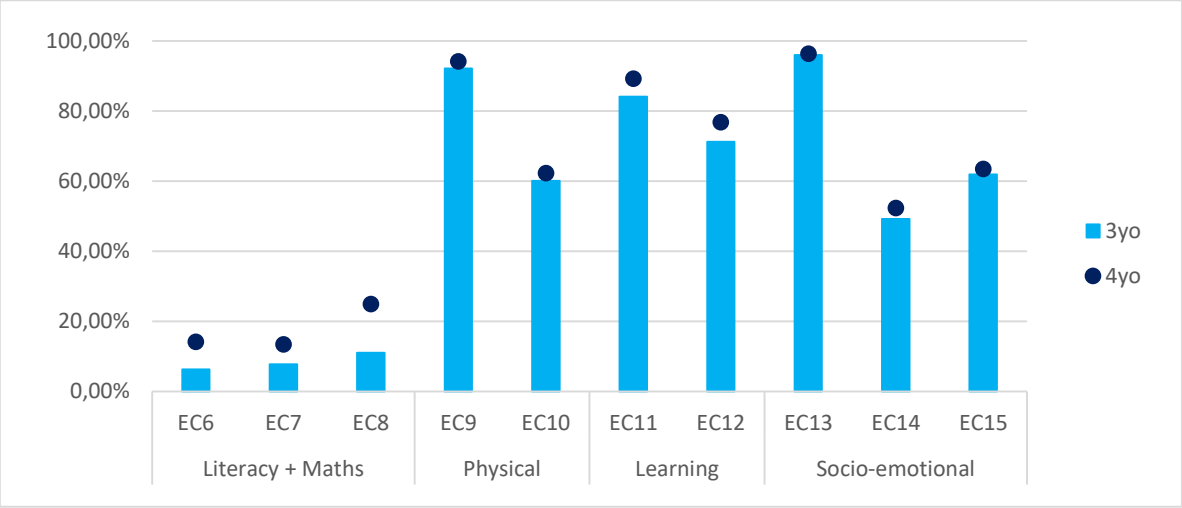
In order to parse how educational performance evolves at the critical ages (3 and 4 years old) considering both general education and specific subjects, an exploratory data analysis and a Chi-square test were performed using Excel and R software.

A simple exploratory analysis reveals that the educational performance has enhanced in all dimensions when comparing 3-year-old children (represented by blue columns) to 4-year-old- children (depicted as dark blue points) groups in the graph below. Significant improvements were observed in Literacy and Mathematics, with the greatest differences in the recognition of numbers from 1 to 10 (EC8). All these numbers were extracted from percentages of affirmative responses (category "yes") to the questions, except for the variables EC10, EC14, and EC15, as the right answer of the question that demonstrated enhancement was the category "no".



The Chi-square test performed was able to prove such results by indicating precisely if the differences were significant statistically. This test is typically applied on categorical variables and its objective is assess if the variables are independent or not. In this case, it evaluated whether 4-year-old children group performed better than the 3-year-old-children group across multiple attributes. Through the statistical test and the p-value below is possible to state that EC6, EC7, EC8, EC11, and EC12 were associated with the age groups (p-value < 0.05), which indicates that as the children grew up, they were able to learn and assimilate new letters, words, numbers, and instructions, as well as do tasks independently.

On the other hand, the test also suggests that EC9, EC10, EC13, EC14, and EC15 are independent of age, which means that 4 year-old-children did not show significant enhancement in physical and socio-emotional skills compared to the younger group (p-value > 0.05). Although there was an increase in percentages between the age groups, the difference was not statistically significant.

Dimension	EC	Chi-square test	p-value	Result
Literacy + Mathematics	6	40.154	**2.347e-10**	4yo group better performance
	7	20.136	**7.214e-06**	4yo group better performance
	8	79.15	**< 2.2e-16**	4yo group better performance
Physical	9	3.5661	0.05897	Independent variables
	10	1.0963	0.2951	Independent variables
Learning	11	13.996	**0.0001832**	4yo group better performance
	12	9.6454	**0.001898**	4yo group better performance
Socio-emotional	13	0.1423	0.706	Independent variables
	14	2.2684	0.132	Independent variables
	15	0.58368	0.4449	Independent variables

Code

```
# setting directory
getwd()
setwd("C:\\Users\\hcsbu\\Desktop\\UNICEF-P3-assessment-public\\01_rawdata")

# importing and reading data
zimb <- read.csv("Zimbabwe_children_under5_interview.csv", sep = ';', header = T)
str(zimb)
attach(zimb)

# transforming the variables types into categorical ones
zimb$child_age_years <- as.factor(zimb$child_age_years)
zimb$EC6 <- as.factor(zimb$EC6)
zimb$EC7 <- as.factor(zimb$EC7)
zimb$EC8 <- as.factor(zimb$EC8)
zimb$EC9 <- as.factor(zimb$EC9)
zimb$EC10 <- as.factor(zimb$EC10)
zimb$EC11 <- as.factor(zimb$EC11)
zimb$EC12 <- as.factor(zimb$EC12)
zimb$EC13 <- as.factor(zimb$EC13)
zimb$EC14 <- as.factor(zimb$EC14)
zimb$EC15 <- as.factor(zimb$EC15)

# creating tables
EC6 <- table(EC6, child_age_years)
EC7 <- table(EC7, child_age_years)
EC8 <- table(EC8, child_age_years)
EC9 <- table(EC9, child_age_years)
EC10 <- table(EC10, child_age_years)
EC11 <- table(EC11, child_age_years)
EC12 <- table(EC12, child_age_years)
EC13 <- table(EC13, child_age_years)
EC14 <- table(EC14, child_age_years)
EC15 <- table(EC15, child_age_years)

# performing chi-square test
cq_test_EC6 <- chisq.test(EC6)
cq_test_EC7 <- chisq.test(EC7)
cq_test_EC8 <- chisq.test(EC8)
cq_test_EC9 <- chisq.test(EC9)
cq_test_EC10 <- chisq.test(EC10)
cq_test_EC11 <- chisq.test(EC11)
cq_test_EC12 <- chisq.test(EC12)
cq_test_EC13 <- chisq.test(EC13)
cq_test_EC14 <- chisq.test(EC14)
cq_test_EC15 <- chisq.test(EC15)
```