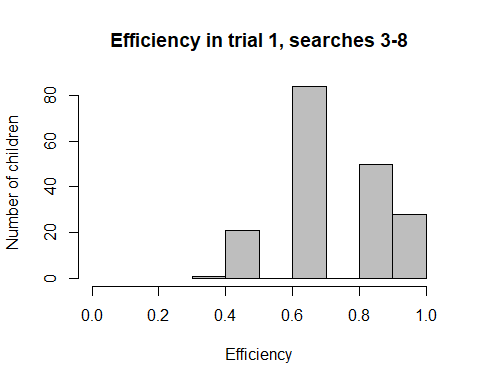
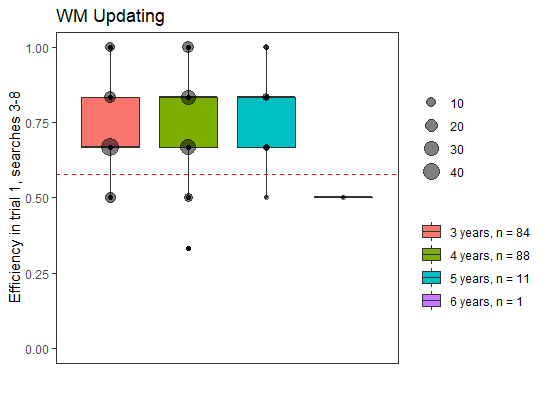
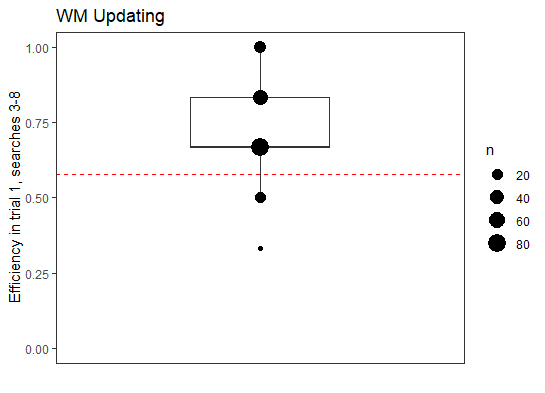
WM Interference IB Boxes Testbattery

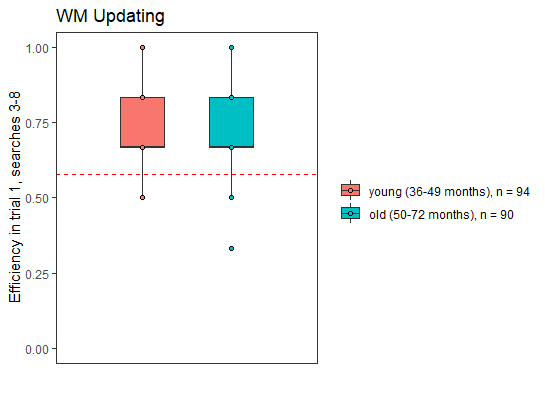
Eva Reindl

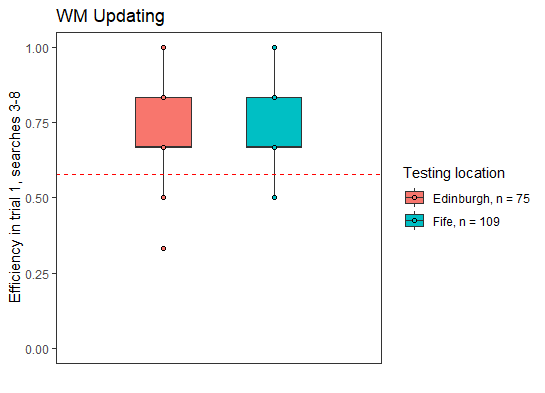
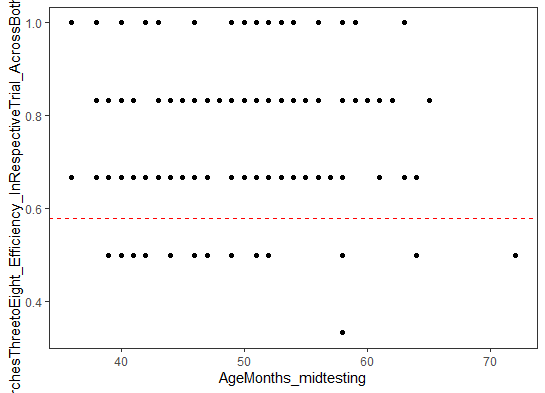
30 3 2020

Key highlights:

* 184 valid data points
* DV: Proportion of correct searches in trial 1 (counting only searches 3-8, so that “0” is a possible outcome): **.74 (SD = .15, range .33-1)**
  + **not normally distributed** W = .873, p < .001
  + Performance significantly **better than chance** (0.57965), V = 16683, p < .001
  + Marginally significant effect of age
  + When age and BPVS score included together, model not better than a null model







## [1] "R version 3.6.1 (2019-07-05)"

# Dropouts

There is **1 dropout** (ID 13) who did not reach 8 trials in either trial 1 nor 2, so this child will be excluded from any further data analysis.

# Valid data

There are **184 children** in the dataset.

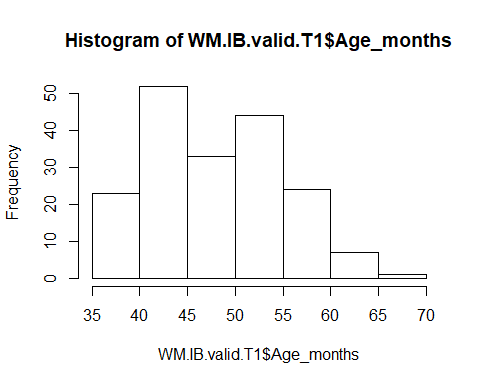
# Sample description

## Gender distribution

There are 100 females and 84 males.

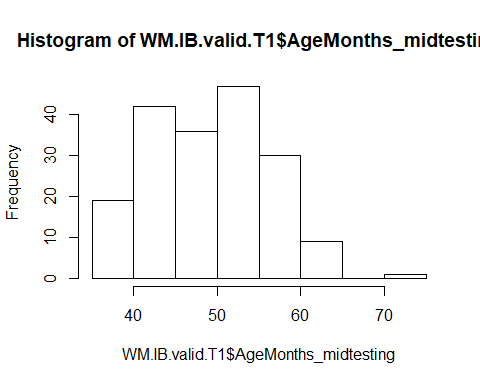
## Age

### Age at beginning of testing



Children who contributed to this task were, at the beginning of testing, on average 48.52 months old (SD = 6.97, range 35-70). There were 1 2-year-old, 86 3-year-olds, 89 4-year-olds, and 8 5-year-olds.

## Age in the middle of testing



Children in this task were on average **49.54 months old (SD = 6.99, range 36-72)** when they were in the middle of the test battery. There were

* 84 3-year-olds
* 88 4-year-olds
* 11 5-year-olds
* 1 6-year-old

## Age mediansplit

There are **94 young** and **90 old** children.

# Testing location

109 children were from Fife, 75 children were from Edinburgh.

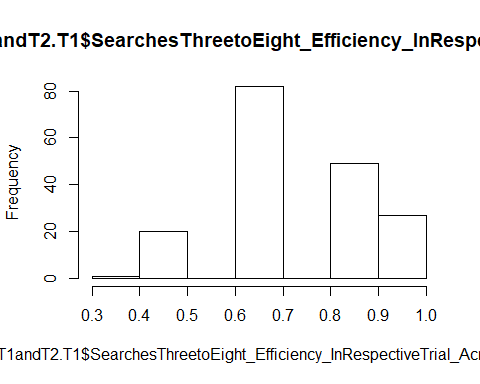
# Comparison of performance between trial 1 and 2

We first want to check whether the efficiencies in trial 1 and 2 are comparable. We hereby only focus on the efficiency in the searches 3 til 8 (we cut the search numbers at 8 because 1) this increases comparability with the ape sample and 2) because these seem the most meaningful trials - more searches just mean that children are guessing. We cut the first 2 searches because those are always correct and thus children could not have an efficiency of 0).

For this analysis, we will only select those children who have completed both trials (n = 179).

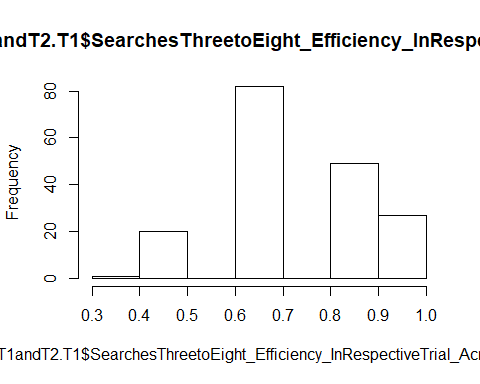
**179 children completed both trials**.

## Efficiency for trial 1



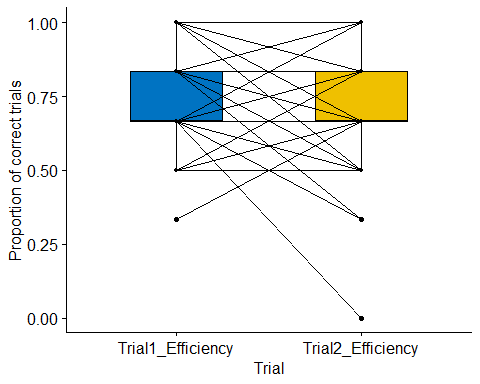
In trial 1, mean efficiency (searches 3-8) is **.74 (SD = 0.15, range 0.33-1)**.

## Efficiency for trial 2



In trial 2, mean efficiency (searches 3-8) is **.72 (SD = 0.17, range 0-1)**.

Efficiency in both trials is **not different from each other**, two-sided Wilcoxon test, V = 4195.5, p = .093.

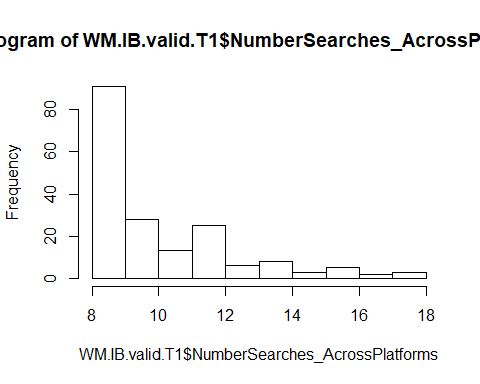


The **correlation** **between performance in both trials is** **r = .23, p = .001**.

The performance between trial 1 and 2 is not different from each other, but correlation is rather low, i.e., some children get better, some get worse. Therefore, we decide to only use the trial 1 data. Note that there are 5 children who did only do trial 1 anyway.

# Description of performance in trial 1

## Total number of searches



The mean number of searches until all stickers were found or the max number of searches was reached was 10.22 (SD = 2.47, range 8-18). 50% of the children needed 10 or fewer searches, and 75% of the children needed 12 or fewer searches.

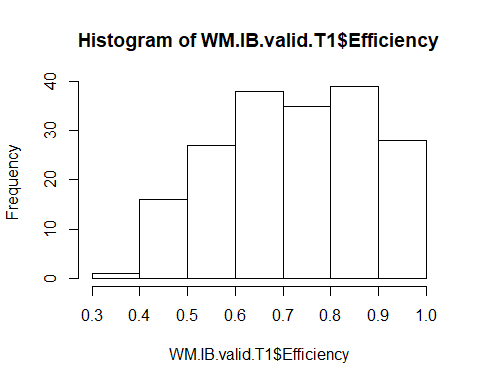
* 3y (n = 84): M = 10.89 (SD = 2.71, range 8-18)
* 4y (n = 88): M = 9.67 (SD = 2.14, range 8-18)
* 5y (n = 11): M = 9.54 (SD = 1.81, range 8-12)
* 6y (n =1): 10
* young (n = 94): M = 10.69 (SD = 2.66, range 8-18)
* old (n = 90): M = 9.73 (SD = 2.15, range 8-18)
* Edinburgh (n = 75): M = 10.07 (SD = 2.15, range 8-18)
* Fife (n = 109): M = 10.33 (SD = 2.67, range 8-18)

## Emptied all boxes?

Of the 184 children, 70 children (38%) emptied all boxes.

* 3y (n = 84): 28 children (33%) emptied all boxes
* 4y (n = 88): 36 children (41%) emptied all boxes
* 5y (n = 11): 6 children (54%) emptied all boxes
* 6y (n =1): 0
* young (n = 94): 31 children (33%) emptied all boxes
* old (n = 90): 39 children (43%) emptied all boxes
* Edinburgh (n = 75): 28 children (37%) emptied all boxes
* Fife (n = 109): 42 children (38%) emptied all boxes

## Efficiency



The mean efficiency in trial 1 was .75 (SD = .17, range .33-1). 50% of the children had an efficiency of 75% or less, and 75% of the children had an efficiency of 87.5% or less.

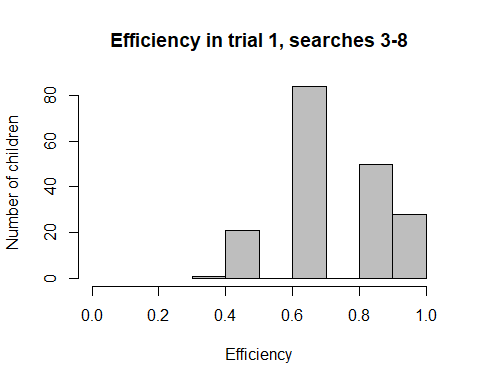
* 3y (n = 84): M = .70 (SD = .17, range .41-1)
* 4y (n = 88): M = .79 (SD = .15, range .33-1)
* 5y (n = 11): M = .81 (SD = .14, range .58-1)
* 6y (n =1): .70
* young (n = 94): M = .71 (SD = .17, range .41-1)
* old (n = 90): M = .79 (SD = .16, range .33-1)
* Edinburgh (n = 75): M = .75 (SD = .15, range .33-1)
* Fife (n = 109): M = .75 (SD = .18, range .41-1)

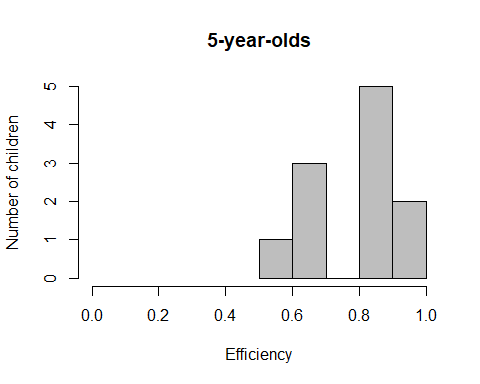
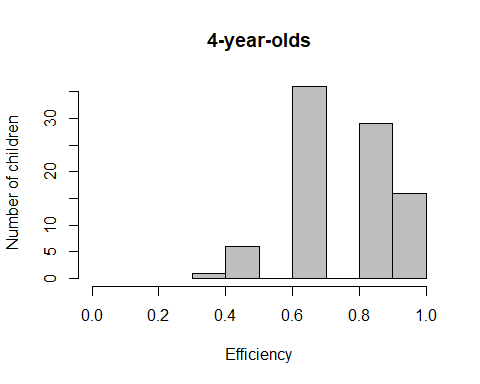
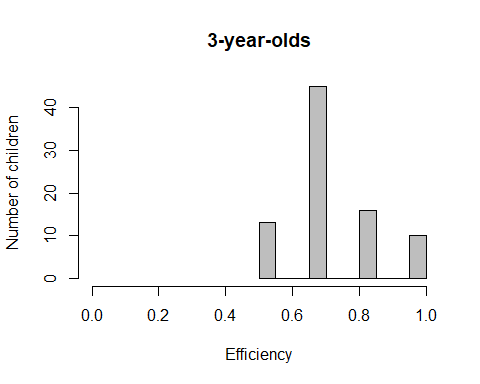
## Search strategy

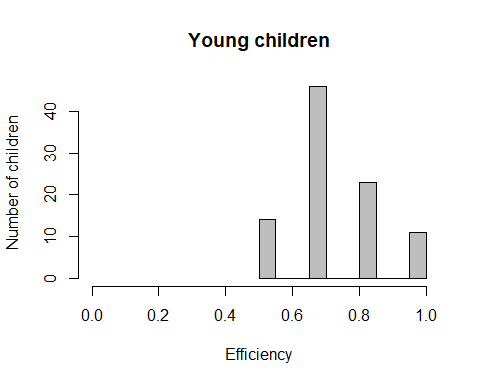
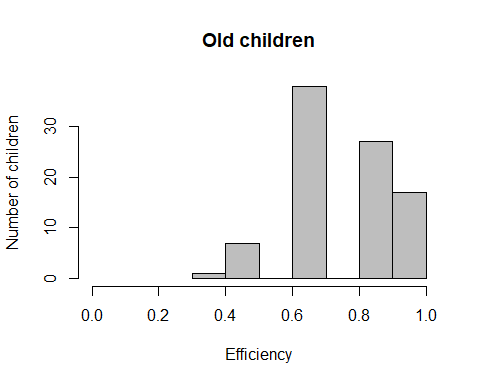
Of the 184 children, 5 children (3%) had a search strategy.

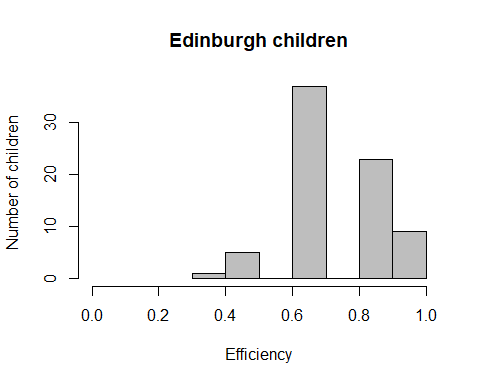
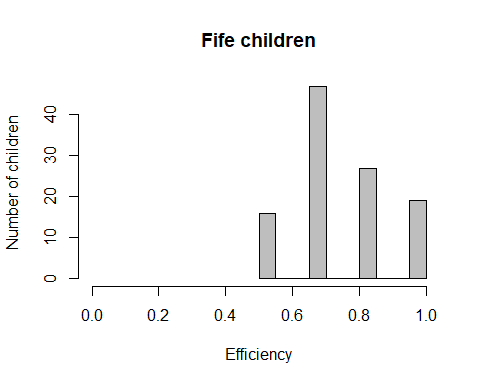
* 3y (n = 84): 1 child (1%) used a strategy
* 4y (n = 88): 3 children (3%) used a strategy
* 5y (n = 11): 1 child (9%) used a strategy
* 6y (n =1): 0
* young (n = 94): 2 children (2%) used a strategy
* old (n = 90): 3 children (3%) used a strategy
* Edinburgh (n = 75): 1 child (1%) used a strategy
* Fife (n = 109): 4 children (4%) used a strategy

## DV: Proportion of correct searches in trial 1 in searches 3-8





The mean efficiency in trial 1, searches 3-8, was **.74 (SD = .15, range .33-1)**. 50% of the children had an efficiency of 67% or less, and 75% of the children had an efficiency of 83% or less. The variable was **not normally distributed**, W = .873, p < .001. Performance was significantly better than chance (0.57965), two-sided Wilcoxon test, V = 16683, p < .001.

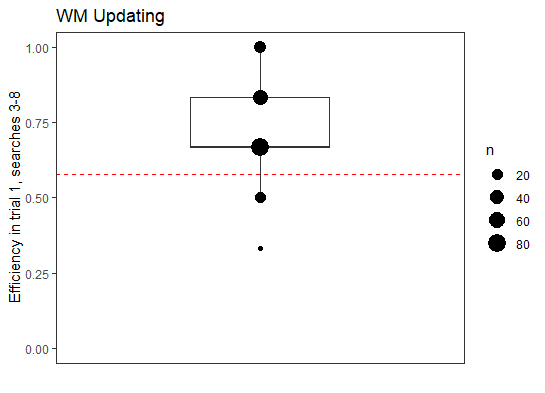
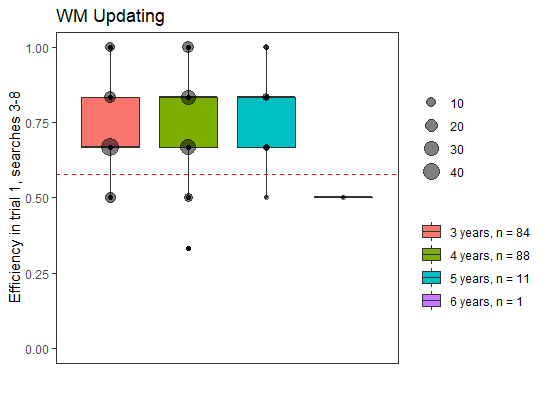
* 3y (n = 84): **M = .71 (SD = .14, range .50-1)**, not normally distributed, W = .835, p < .001, better than chance, V = 3479, p < .001
* 4y (n = 88): **M = .77 (SD = .15, range .33-1)**, not normally distributed, W = .876, p < .001, better than chance, V = 3852, p < .001
* 5y (n = 11): **M = .79 (SD = .15, range .50-1)**, normally distributed, W = .899, p = .181, better than chance, V = 65, p = .004
* 6y (n =1): .50

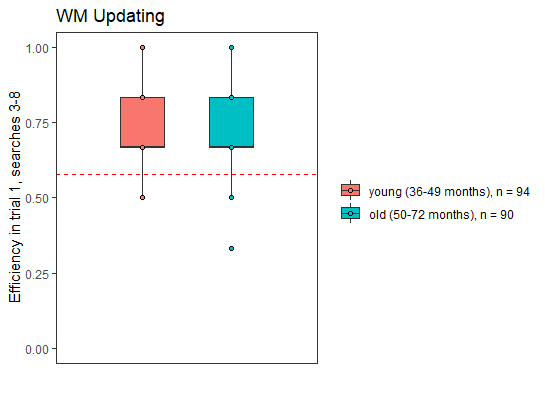
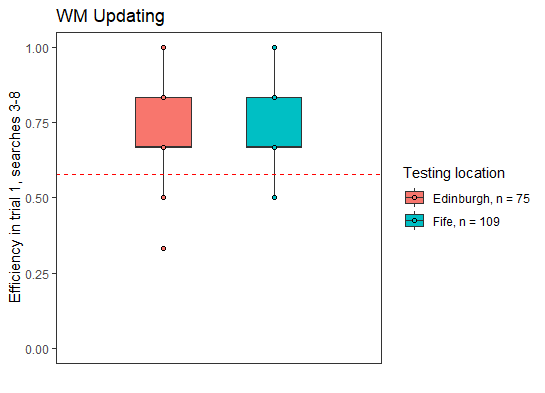
4-year-olds performed better than 3-year-olds, one-sided Wilcoxon test, W = 2897, p = .004.

* young (n = 94): **M = .72 (SD = .14, range .50-1)**, not normally distributed, W = .857, p < .001, better than chance, V = 4360, p < .001
* old (n = 90): **M = .76 (SD = .15, range .33-1)**, not normally distributed, W = .876, p < .001, better than chance, V = 4021, p < .001

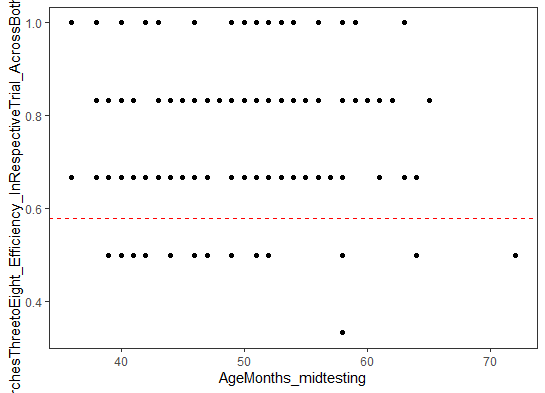
Older children performed better than younger children, one-sided Wilcoxon test, W = 4890, p = .026.

* Edinburgh (n = 75): M = .74 (SD = .14, range .33-1), not normally distributed, W = .859, p < .001
* Fife (n = 109): M = .74 (SD = .16, range .50-1), not normally distributed, W = .867, p < .001





## Plot age as continuous variable against proportion correct



`

## Can age predict efficiency?

#Full model

res<-lm(SearchesThreetoEight\_Efficiency\_InRespectiveTrial\_AcrossBothPlatforms ~ z.age.midtesting, data=WM.IB.valid.T1)

Age has only a marginally significant effect on proportion correct, X2(1) = 0.068, p = .081.

Note that the age effect is stronger if we choose efficiency across all administered trials as the DV (X2(1) = 0.315, p < .001.