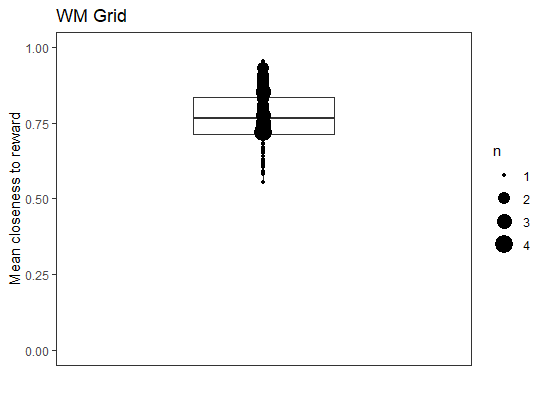
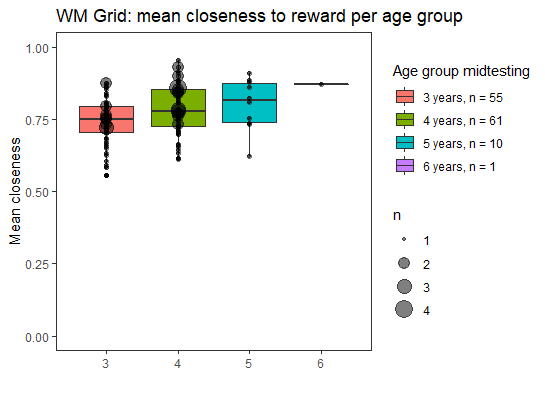
WM\_grid

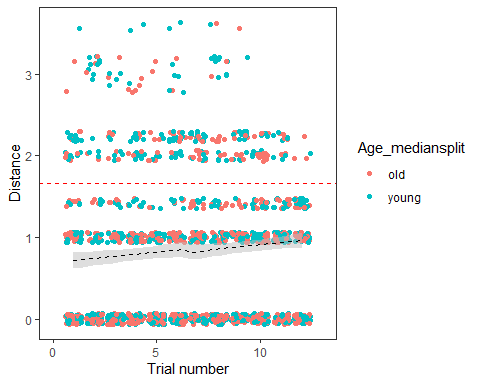
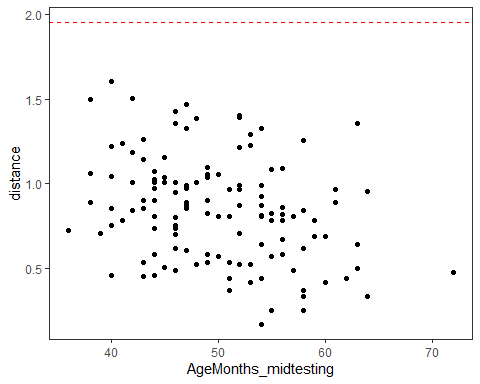
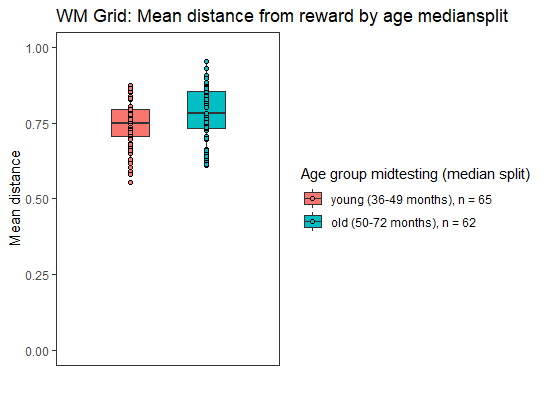
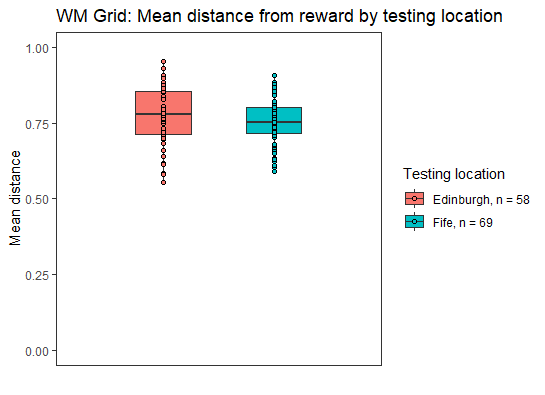
Eva Reindl

31 3 2020

Key findings:

* 127 valid datapoints
* DV: Closeness to reward (scale 0-1)
  + Normally distributed
  + Children (across sample all age groups, by mediansplit) perform above chance
  + Age, trial number, and testing location have an effect on the DV





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

# Dropouts

We have 3 dropouts (2 children stopped after 5 test trials, 1 in the middle of the 8th test trial.

# Select the test trials and remove dropouts

Final **sample size is 127**.

# Which children to include?

We eventually decided to include children who have completed 75% (i.e., 9) of the test trials. Before deciding this, we wanted to know whether performance in trials 1-4 is different from performance in trials 5-12, and thus we wanted to explore whether we could cut the test length (either here to include children who stopped earlier or with regard to future administrations of this test). For this analysis, we need to **remove the children who have fewer than 12 trials**.

## Is performance in trials 1-4 comparable to performance in trials 5-12?

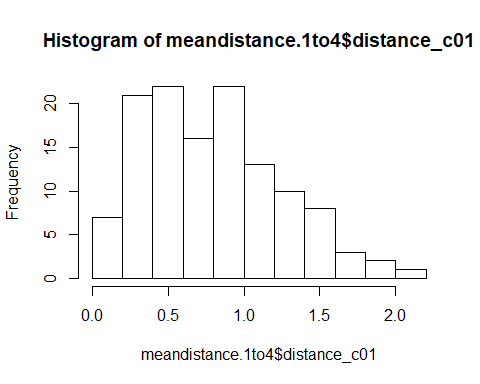
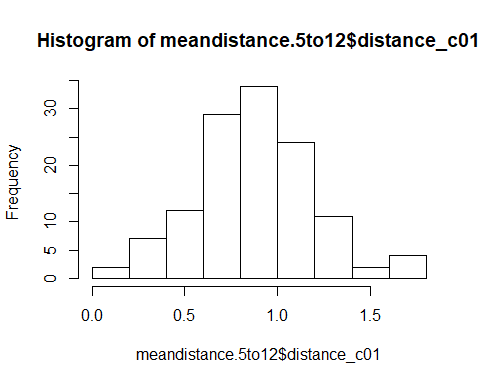
### Performance in trials 1-4

Aggregating over trials 1 to 4, the mean distance was 0.78 (SD = 0.46, range 0-2.04).

### Performance in trials 5-12

Aggregating over trials 5 to 12, the mean distance was 0.87 (SD = 0.32, range 0.12-1.68).

### Comparison performance in trials 1-4 and 5-12

Performance in Trials 1-4 and 5-12 is significantly different from each other (paired samples Wilcoxon test, V = 2908.5, p = .022.

## Is performance in trials 1-6 comparable to performance in trials 7-12?

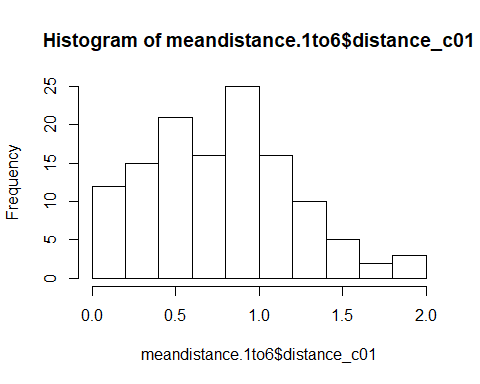
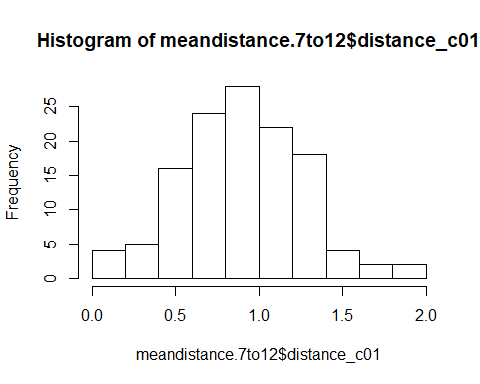
### Performance in trials 1-6

Aggregating over trials 1 to 7, the mean distance was 0.78 (SD = 0.42, range 0-1.90).

### Performance in trials 7-12

Aggregating over trials 1 to 7, the mean distance was 0.90 (SD = 0.35, range 0.17-1.97).

### Comparison performance in trials 1-6 and 7-12

Performance in Trials 1-6 and 7-12 is significantly different from each other (paired samples Wilcoxon test, V = 2771, p = .012).

## Is performance in trials 1-8 comparable to performance in trials 9-12?

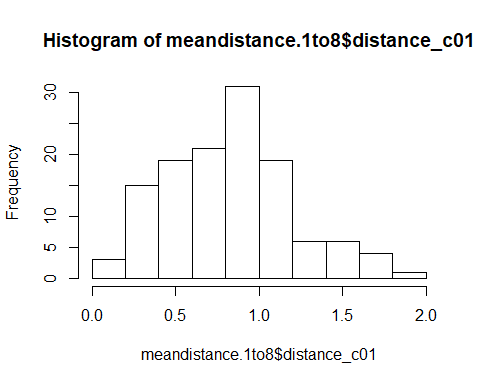
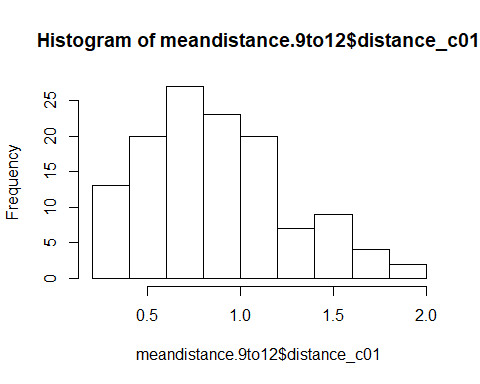
### Performance in trials 1-6

Aggregating over trials 1 to 8, the mean distance was 0.83 (SD = 0.38, range 0.12-1.98).

### Performance in trials 9-12

Aggregating over trials 1 to 8, the mean distance was 0.86 (SD = 0.38, range 0.25-1.82).

### Comparison performance in trials 1-8 and 9-12

Performance in Trials 1-8 and 9-12 is not different from each other (paired samples Wilcoxon test, V = 3463.5, p = .463). Therefore, **in the future, this task could potentially be shortened to 8 instead of 12 trials**.

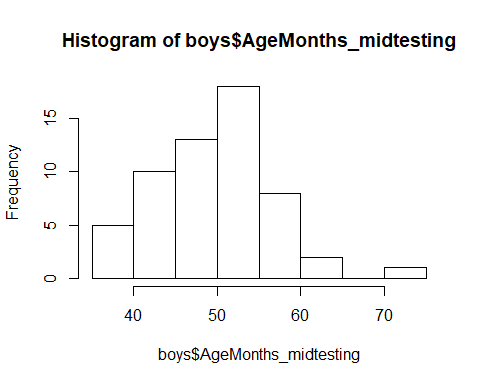
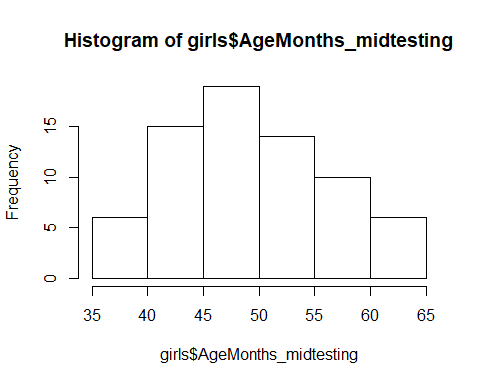
# Sample description

## Gender distribution

There are **70 girls** and **57 boys**.

* Girls: 34 3y, 29 4y, 7 5y
* Boys: 21 3y, 32 4y, 3 5y, 1 6y

There is no difference in the age distribution between boys and girls.

## Age

### Age at beginning of testing

At the beginning of testing, the children who had valid data on the WM grid task were on average 48.83 months (SD = 6.90, range 36-70) old. There were 58 3-year-olds, 62 4-year-olds, and 7 5-year-olds.

### Age in the middle of testing

In the middle of testing, the children who had valid data on the WM grid task were on average **49.90 months (SD = 6.86, range 36-72)** old. There were

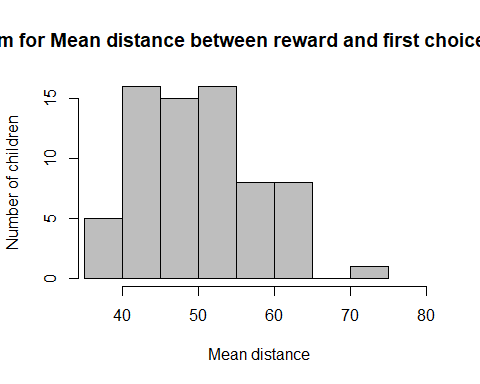
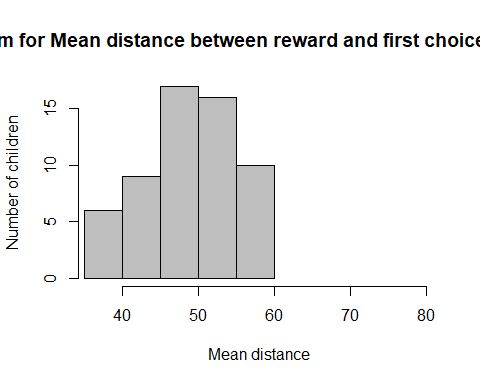
* 55 3-year-olds
* 61 4-year-olds
* 10 5-year-olds
* 1 6-year-old

### Age mediansplit by entire sample

Median is 49 months

There are **65 young** and **62 old** children.

## Testing Location

69 children were from the Fife area, 58 children were from Edinburgh. There is no difference in the age distribution between the two locations, two-sided Wilcoxon test, W = 1904, p = .640.

### Testing location and age

Fife:

* 3y: 31
* 4y: 27
* 5y: 10
* 6y: 1

Edinburgh:

* 3y: 24
* 4y: 34

# Warm-up

## Mean number of toys won over the two warm-up trials

In the two warm-up trials, children won on average 1.87 (SD = 0.36, range 0-2) toys.

## Mean number of searches in the two warm-up trials

Out of 254 warm-up trials, children were successful in 238 (94%). In those 238 warm-up trials where children were successful, children needed on average 1.13 (SD = 0.31, range 1-3) searches to find the toy.

### In how many warm-up trials did children not find the toy?

Out of 254 warm-up trials, there were only 16 (6%) in which the children did not get the toy.

# How many searches were done?

table(test.trials$NrSearchesUntilRewardFound)

##   
## 1 2 3 4   
## 669 210 98 541

## By age groups

* 3-year-olds: 656 searches in total, out of which 257 (39%) are 1 search, 90 (14%) are 2 searches, 48 (7%) are 3 searches, and in 261 (40%) searches the toy was not found.
* 4-year-olds: 730 searches in total, out of which 344 (47%) are 1 search, 103 (14%) are 2 searches, 44 (6%) are 3 searches, and in 239 (33%) searches the toy was not found.
* 5-year-olds: 120 searches in total, out of which 60 (50%) are 1 search, 15 (12%) are 2 searches, 6 (5%) are 3 searches, and in 39 (32%) searches the toy was not found.
* 6-year-old: 12 searches in total, out of which 8 (67%) are 1 search, 2 (17%) are 2 searches, and in 2 (17%) searches the toy was not found.

## By age mediansplit

* young: 774 searches in total, out of which 307 (40%) are 1 search, 105 (13%) are 2 searches, 59 (8%) are 3 searches, and in 303 (39%) searches the toy was not found.
* old: 744 searches in total, out of which 362 (49%) are 1 search, 105 (14%) are 2 searches, 39 (5%) are 3 searches, and in 238 (32%) searches the toy was not found.

# How did children perform in the distractor task?

Out of the 1518 test trials, in 1259 trials (83%) children found the sticker in their first try of solving the distractor task. In 258 trials (17%) children failed to find the sticker.

## By age groups

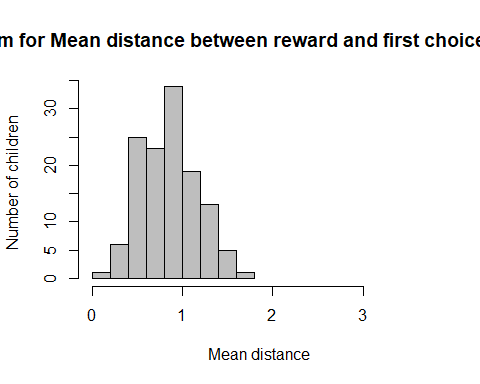
* 3y: out of the 656 test trials, in 520 (79%) children found the sticker, in 136 they did not.
* 4y: out of the 729 test trials, in 622 (85%) children found the sticker, in 107 they did not.
* 5y: out of the 120 test trials, in 107 (89%) children found the sticker, in 13 they did not.
* 6y: out of the 12 test trials, in 10 (83%) the child found the sticker, in 2 they did not.

## By age mediansplit

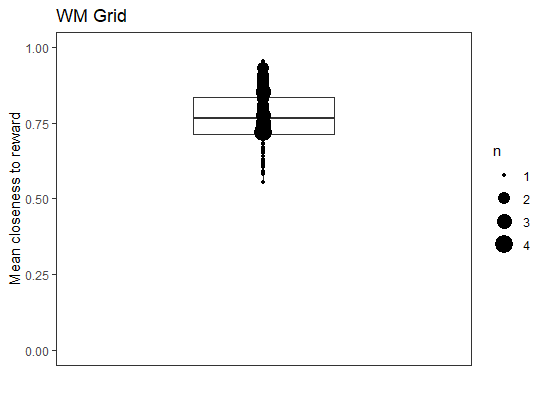
* young: out of the 774 test trials, in 621 (80%) children found the sticker, in 153 they did not.
* old: out of the 743 test trials, in 638 (86%) children found the sticker, in 105 they did not.

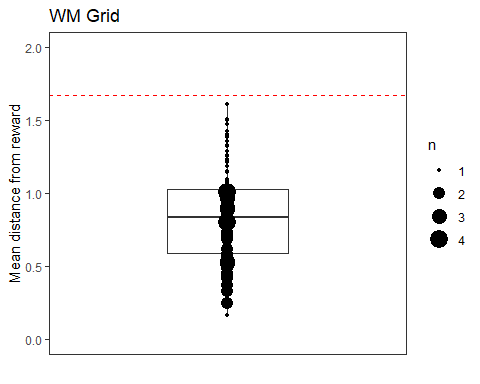
# DV: Average distance across test trials

## Across the sample



The mean distance from the reward was 0.84 (SD = 0.31, range 0.17-1.60). The DV is normally distributed, W = 0.987, p = .251.

Original scale



### Do children perform significantly below chance?

#### Test against all cells

Christoph’s simulation is based on 1000000 iterations of 12 random choices of a cell on the 4x4 grid. The distance between these random choices and the hiding location of the reward was then calculated. The mean distance of the these random choices from the reward is used as the hypothetical chance value.

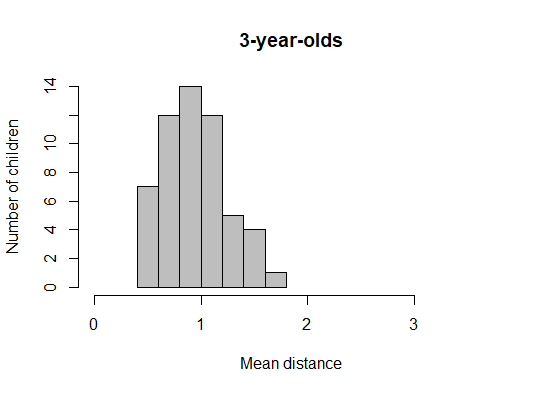
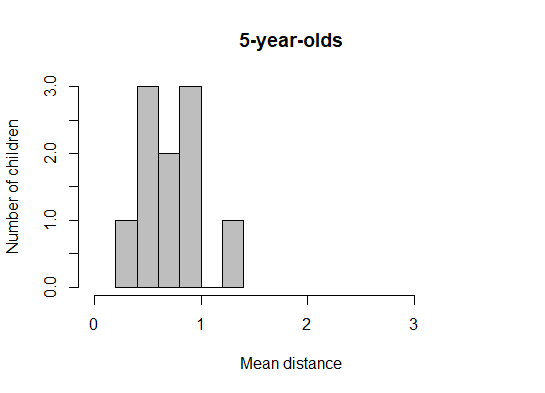
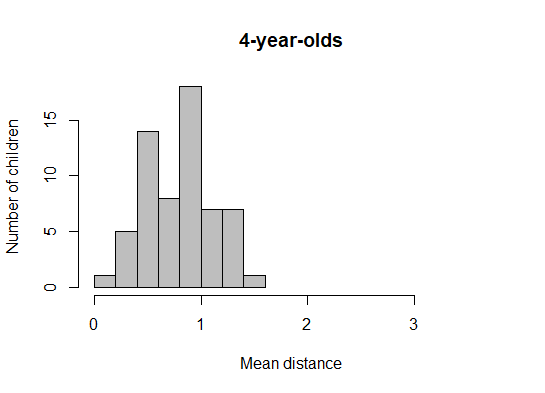
Mean distance vs chance value (1.954): M = 0.84 (SD = 0.31, range 0.17-1.60) is significantly better than chance, t(126) = -40.65, p < .001.

#### Test against inner cell preference

The simulation is identical to the aforementioned one but only the inner cells are sampled, resulting in a smaller average distance from the reward.

Mean distance between chosen box and baited cell vs chance value (inner cells only: 1.667): M = 0.84 (SD = 0.31, range 0.17-1.60) is significantly better than chance, t(126) = -30.14, p < .001.

## By age groups



* 3y (n = 55): M = 0.93 (SD = 0.29, range 0.45-1.60)
* 4y (n = 61): M = 0.79 (SD = 0.31, range 0.17-1.40)
* 5y (n = 10): M = 0.72 (SD = 0.32, range 0.33-1.34)
* 6y (n = 1): 0.47

### Do children perform significantly below chance?

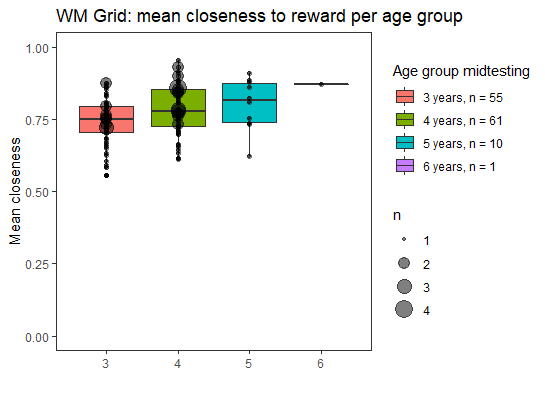
#### Test against all cells

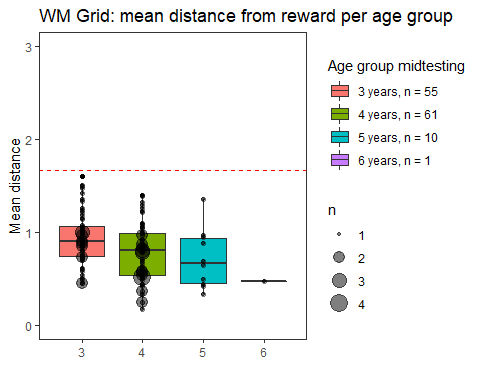
* 3y: Mean distance (0.93) vs chance value (1.954): significantly better than chance, t(54) = -26.30, p < .001.
* 4y: Mean distance (0.79) vs chance value (1.954): significantly better than chance, t(60) = -29.69, p < .001.
* 5y: Mean distance (0.72) vs chance value (1.954): significantly better than chance, t(9) = -12.19, p < .001.

#### Test against inner cell preference

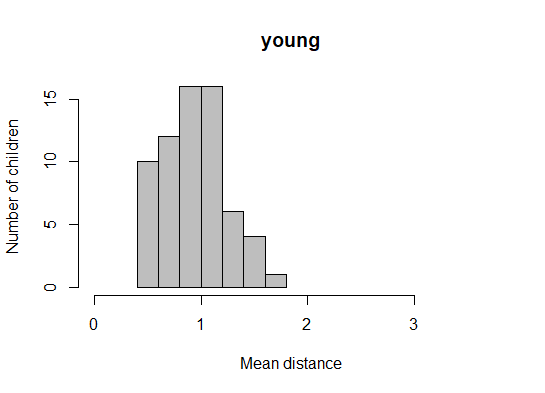
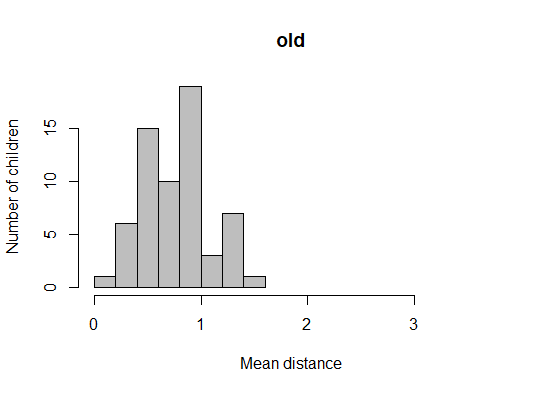
The simulation is identical to the aforementioned one but only the inner cells are sampled, resulting in a smaller average distance from the reward.

* 3y: Mean distance (0.93) vs chance value (1.954): significantly better than chance, t(54) = -18.91, p < .001.
* 4y: Mean distance (0.79) vs chance value (1.954): significantly better than chance, t(60) = -22.37, p < .001.
* 5y: Mean distance (0.72) vs chance value (1.954): significantly better than chance, t(9) = -9.36, p < .001.





## By mediansplit



* young (n = 65): M = 0.93 (SD = 0.28, range 0.45-1.60)
* old (n = 62): M = 0.76 (SD = 0.31, range 0.17-1.40)

Older children were significantly better than younger children, t(123) = - 3.22, p < .001.

### Do children perform significantly below chance?

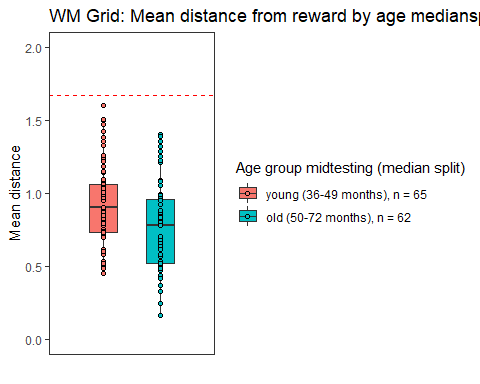
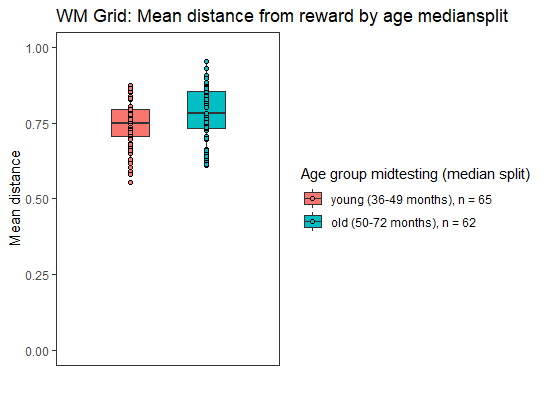
#### Test against all cells

* young: M = 0.93 vs chance value (1.954): significantly better than chance, t(64) = -29.05, p < .001
* old: M = 0.76 vs chance value (1.954): significantly better than chance, t(61) = -30.53, p < .001

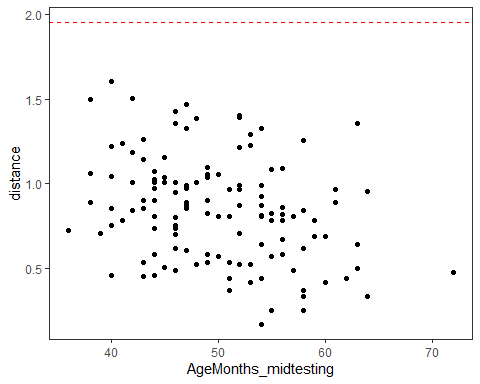
#### Test against inner cell preference

The simulation is identical to the aforementioned one but only the inner cells are sampled from resulting in a smaller average distance from the food reward.

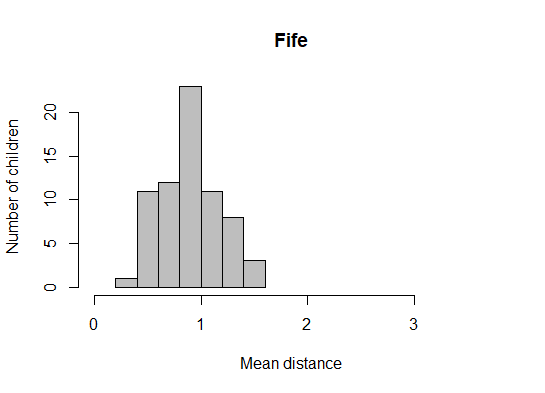
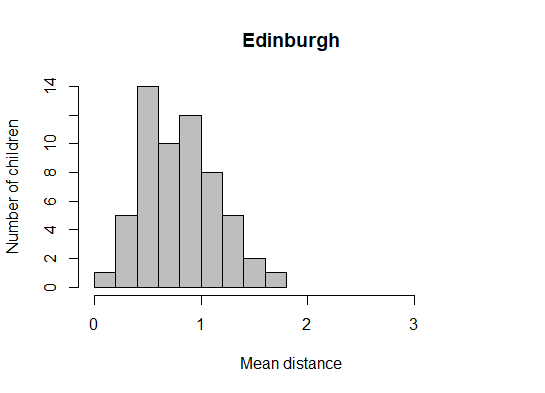
* young: M = 0.93 vs chance value (1.667): significantly better than chance, t(64) = -20.94, p < .001
* old: M = 0.76 vs chance value (1.667): significantly better than chance, t(61) = -23.21, p < .001



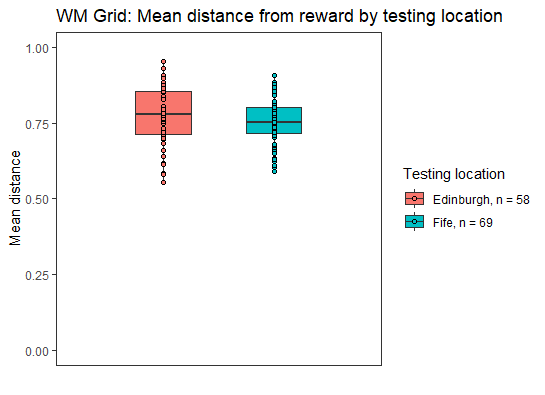
## Plot distance against age as continuous variable

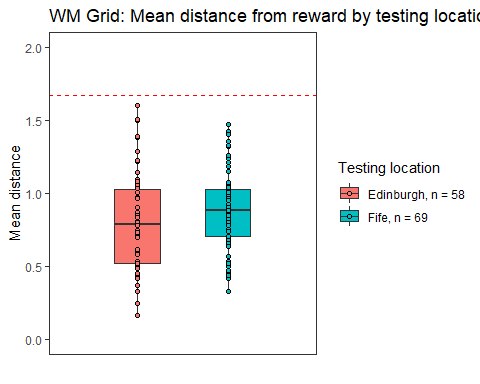


## By testing location

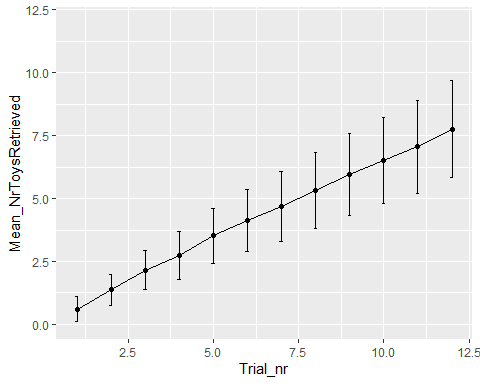


* Fife (n = 69): 0.88 (SD = 0.27, range 0.33-1.47)
* Edinburgh (n = 58): 0.79 (SD = 0.34, range 0.17-1.60)

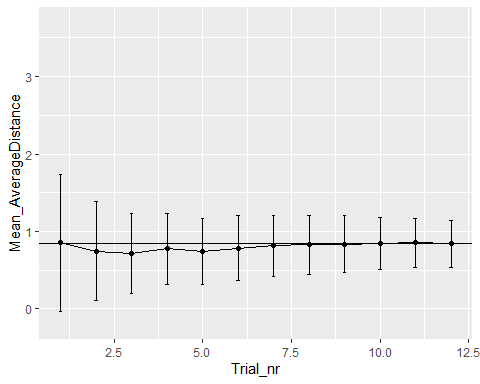
Children in Edinburgh are significantly better than children in Fife, one-sided t-test, t(109.5) = -1.66, p = .05.



# Cumulative number of toys retrieved



# Cumulative average distance



# Can the distance be predicted by age?

dist<-lmer(distance\_c01 ~ z.age + z.Trial + z.age:z.Trial + (1+z.Trial|ID) + (1 + z.Trial + z.age|NurseryCode), data=test.trials, REML=FALSE)#singular fit

## boundary (singular) fit: see ?isSingular

#remove correlation between random effects  
dist<-lmer(distance\_c01 ~ z.age + z.Trial + z.age:z.Trial + (1|ID) + (0+z.Trial|ID) + (1|NurseryCode) + (0 + z.Trial|NurseryCode) + (0 + z.age|NurseryCode), data=test.trials, REML=FALSE)#singular fit

The variance for the random slope of trial on ID and for the random slope of age on nursery code is effectively 0, so we remove these from the model.

dist<-lmer(distance\_c01 ~ z.age + z.Trial + z.age:z.Trial + (1|ID) + (1|NurseryCode) + (0 + z.Trial|NurseryCode), data=test.trials, REML=FALSE)#singular fit

Together, trial number, age, and the interaction between trial number and age can explain the data significantly better than a null model only containing the intercept, X2(3) = 15.94, p = . 001.

The interaction term did not significantly improve model fit, X2(1) = 0.432, p = .511, so we remove it from the model, and we enter trial (the non-z-trasnformed version, for better interpretability).

New full model:

dist<-lmer(distance\_c01 ~ z.age + Trial + (1|ID) + (1|NurseryCode) + (0 + Trial|NurseryCode), data=test.trials, REML=FALSE)#singular fit

The model including age and trial number can explain the data on distance significantly better than a model only containing the intercept, X2(2) = 16.67, p < .001.

We see that the variance for the random slope of trial on nursery code is effectively 0, so we exclude it from the model.

dist<-lmer(distance\_c01 ~ z.age + Trial + (1|ID) + (1|NurseryCode), data=test.trials, REML=FALSE)#singular fit

The model including age and trial number can explain the data on distance significantly better than a model only containing the intercept, X2(2) = 16.67, p < .001.

We see that the variance of nursery is very small. However, we know from above that there is an effect of testing location (Fife/Edinburgh), so we modify the model to look for that difference.

We aim to investigate the potential role of testing location.

dist<-lmer(distance\_c01 ~ z.age + Trial + TestingLocation + (1+Trial|ID), data=test.trials, REML=FALSE)#singular fit

## boundary (singular) fit: see ?isSingular

#remove correlation  
dist<-lmer(distance\_c01 ~ z.age + Trial + TestingLocation + (1|ID) + (0+Trial|ID), data=test.trials, REML=FALSE)#singular fit

## boundary (singular) fit: see ?isSingular

The variance for the random slope of trial on ID is effectively 0, so we exclude it from the model.

dist<-lmer(distance\_c01 ~ z.age + Trial + TestingLocation + (1|ID), data=test.trials, REML=FALSE)

Together, age, trial number, and testing location explain the data significantly better than the null model, X(3) = 21.05, p < .001.

**Effect of age**

Age has a significant effect on distance, X2(1) = 14.35, p < .001.

**Effect of trial number**

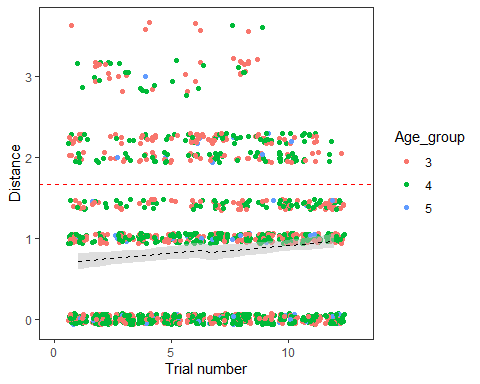
Trial number has a significant effect on distance, X2(1) = 3.96, p = .047.

**Effect of testing location**

Location has a significant effect on distance, X2(1) = 4.37, p = .036.

For each increase of 1 SD in age, the distance decreases (i.e., children get better) by 0.10. With each trial, distance gets longer (i.e., children get worse) by 0.01. Going from Edinburgh to Fife, the distance increases by 0.11.

Plot effect of trial number on distance:



Plot, by mediansplit age

