Bing machine analysis

Zeynep Civelek

August 11, 2020

# Dropouts

We have **5 dropouts** (1 child stopped after 4 trials, 1 child stopped after 8 trials, 1 child stopped after 9 trials, 1 child where RA gave away the game by saying "let's watch and listen", 1 child with an experimenter error in trial 11 (wrong noise) so we only had 10 trials).

# Which children to include?

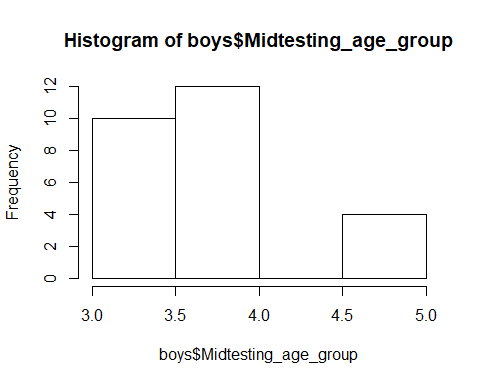
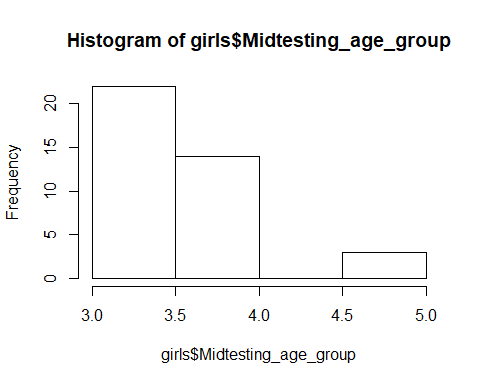
We decided to include children who have completed 75% (i.e., 15 trials) of the test trials. We included 4 children who did not have 100% of the datapoints: 2 children where we used the wrong coding sheet and accidentally only administered 16 trials; 1 child where there was an experimenter error (wrong noise in trial 18), so we could only use 17 trials, 1 child who stopped after 16 trials

# Removing the dropouts from data

Final **sample size is 65**.

# Sample description

## Gender distribution



There are **39 girls** and **26 boys**.

* Girls: 22 3y, 14 4y, 3 5y
* Boys: 10 3y, 12 4y, 4 5y

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 Bing Machine task were on average 47.57 months (SD = 6.93, range 36-64) old. There were 34 3-year-olds, 26 4-year-olds, and 5 5-year-olds.

### Age in the middle of testing

In the middle of testing, the children who had valid data on the Bing Machine task were on average **49.29 months (SD = 6.90, range 39-65)** old. There were

* 32 3-year-olds
* 26 4-year-olds
* 7 5-year-olds

### Age mediansplit by entire sample

Median is 49 months.

There are **34 young** and **31 old** children.

# Warm-up

All children (N=65) passed the transparent training within 3 trials (the minimum number of required trials to pass).

# Number of correct trials in Test

The average number of correct trials was **9.70** (SD = 2.57, range 5-17). 50% of the children had 9 or fewer trials correct. The distribution of correct trials is skewed.

Split by age:

- 3-year-olds: M = **9.53** (2.32, range 6-17)

- 4-year-olds: M = **9.76** (2.79, range 5-17)

- 5-year-olds: M = **10.29** (2.72, range 7-14)

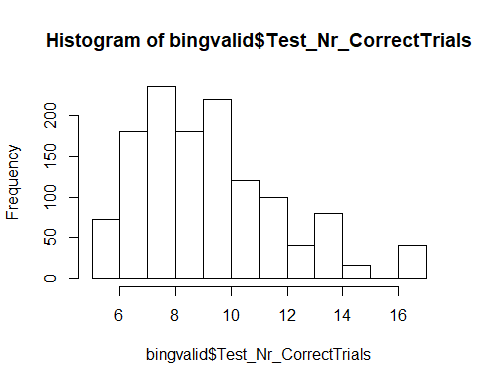
Thus, numerically, all age groups seem to be performing equally.

In terms of the median split of age, we found:

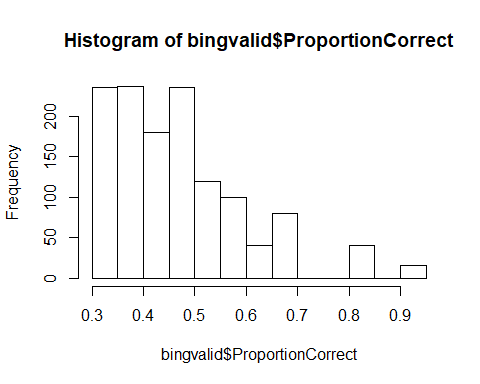
- young children: **9.71** (SD = 2.38, range 6-17)

- old children: **9.70** (SD = 2.76, range 5-17)

So using the median split, there was no difference between the groups.



# DV: Proportion correct



The average proportion of correct trials was **49.03%** (SD = 13.23, range 30-93.75%). 50% of the children had 45% or a smaller proportion of their trials correct. The DV is not normally distributed, W = 0.907, p < .001.

Split by age:

- 3-year-olds: M = 47.65% (11.60, range 30-85%)

- 4-year-olds: M = 50.09% (14.80, range 31.25-93.75%)

- 5-year-olds: M = 51.43 (13.60, range 35-70%)

Thus, numerically, all age groups seem to perform equally.

In terms of the median split of age, we found:

- Young children: **48.53** (SD = 11.92, range 30-85%)

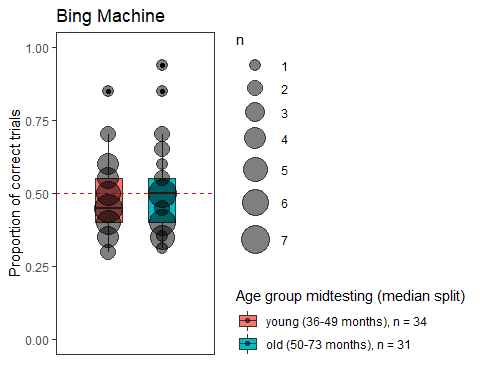
- Old children: **49.50** (SD = 14.55, range 31.25-93.75%)

So using the median split, there doesn't seem to be a difference between the groups.

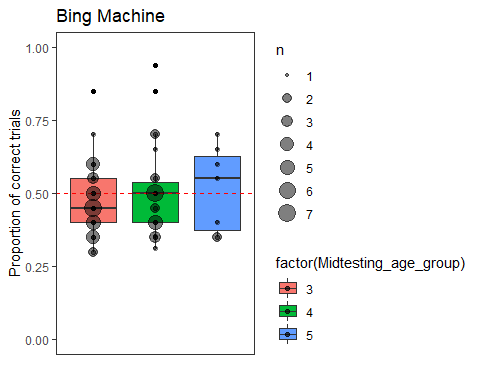
# Does performance deviate from chance?

Children's proportion correct **is not different from chance**, V = 610.5, p = .352. Neither young (V = 182, p = .447) nor old children (V = 137, p = .719) performed differently than chance.

# Boxplot split by median split\_ age:



# Boxplot split by the three age groups:



# Can children's performance be predicted by age?

full<-glmer(Success\_in0and1 ~ z.age.midtesting + z.trialno + z.age.midtesting:z.trialno + (1+z.trialno|id), data=bingvalid, family = binomial, control=glmerControl(optimizer="bobyqa",optCtrl=list(maxfun=100000)))

**Comparison against null model**

null<-glmer(Success\_in0and1 ~ 1 + (1|id) + (0+z.trialno|id), data=bingvalid, family = binomial, control=glmerControl(optimizer="bobyqa",optCtrl=list(maxfun=100000)))

Trial number, age, and the interaction between trial number and age together **does not explain the data better than a null model** only containing an intercept, X2(4) = 5.781 p = .216. But the interaction term is not significant x2(1)=2.291, p=0.13 so it is removed from the model to explore the main effect of trial number and age.

full2<-glmer(Success\_in0and1 ~ z.age.midtesting + z.trialno + (1|id) + (0+z.trialno|id), data=bingvalid, family = binomial, control=glmerControl(optimizer="bobyqa",optCtrl=list(maxfun=100000)))  
  
null2<-glmer(Success\_in0and1 ~ 1 + (1|id) + (0+z.trialno|id), data=bingvalid, family = binomial, control=glmerControl(optimizer="bobyqa",optCtrl=list(maxfun=100000)))

The model including trial number and age can explain the data significantly better than a null model only containing an intercept, X2(2) = 7.255 p < .05.

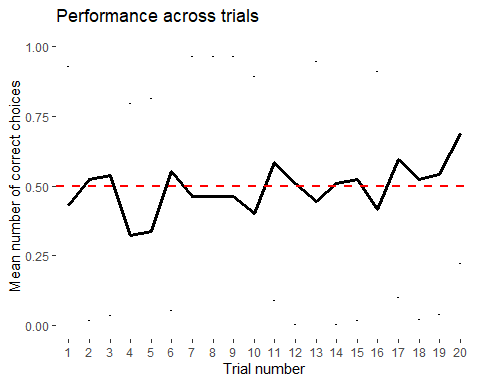
**Effect of age and trial number**

modeldrop2=drop1(full2, test="Chisq",control=contr)  
round(modeldrop2,3)

## Single term deletions  
##   
## Model:  
## Success\_in0and1 ~ z.age.midtesting + z.trialno + (1 | id) + (0 +   
## z.trialno | id)  
## npar AIC LRT Pr(Chi)   
## <none> 1776.6   
## z.age.midtesting 1 1774.9 0.272 0.602   
## z.trialno 1 1781.7 7.014 0.008 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

There is a significant effect of trial number, X2(1) = 7.014, p < .01, but no effect of age X2(1) = 0.272, p = .602.

**Plot**



Performance seems to get better over trials but this is expected in this task as children have no way of knowing where to locate the reward in the first trial.