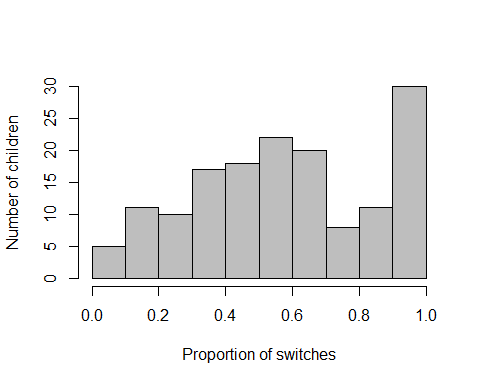
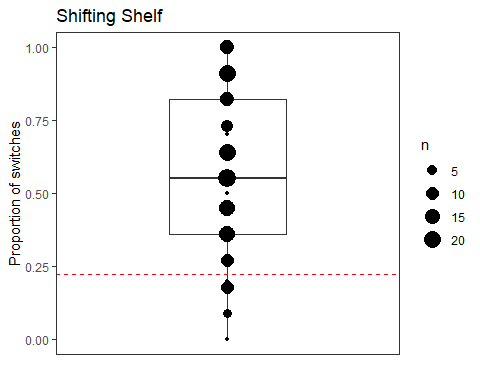
Shelf

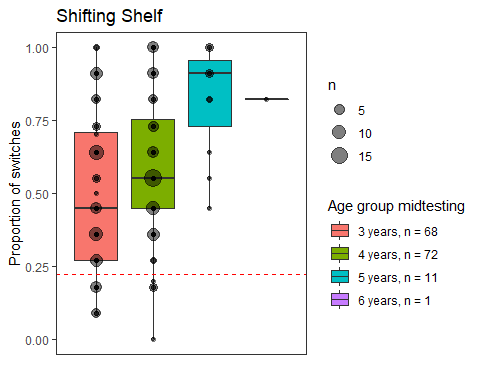
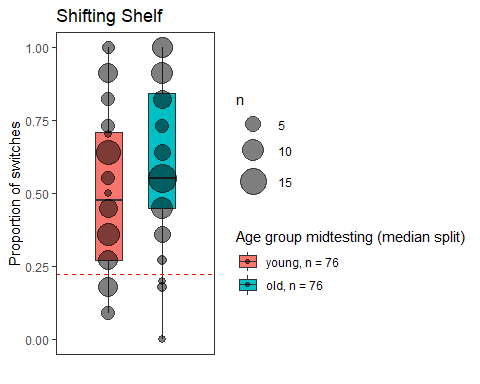
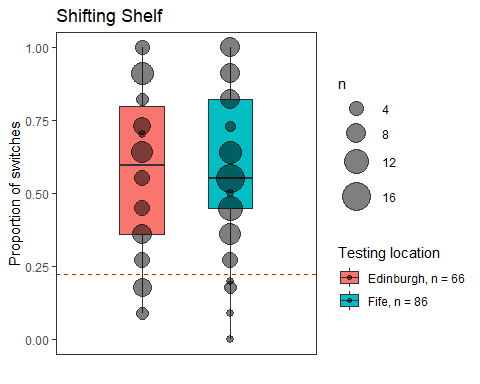
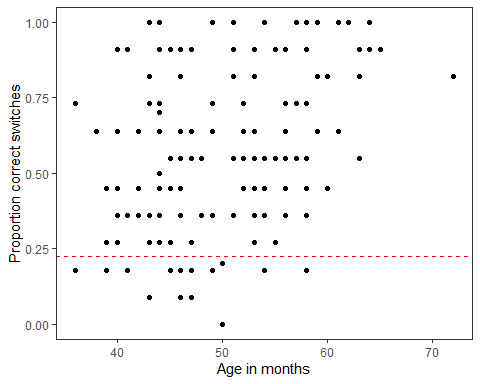
Eva Reindl

30 07 2020

Key highlights:

* **152 valid cases** in the test phase
* DV: **proportion of switches** out of all possible switches: 0.57 (SD = 0.26, range 0-1)
* **not normally distributed**
* Performance significantly **above chance level** (0.2242245), V = 11425, p < .001
* **3-year-olds performed fewer switches than 4-year-olds**, **younger children made fewer switches than older** children
* Significant **effect of age and trial number** on success (trial by trial data), no effect of testing location

*  
*  

# Description of sample (Age)

## At the beginning of testing

There are **178 children in the dataset**. There are 97 females and 81 males. At the beginning of testing, the children tested in the shelf task were 48.48 months (SD = 6.85, range 35-70). There were 1 2-year-old, 83 3-year-olds, 86 4-year-olds, and 8 5-year-olds.

## In the middle of testing

In the middle of testing, the children tested in the shelf task were **49.59 months (SD = 6.89, range 36-72)**. There were **81 3-year-olds, 85 4-year-olds, 11 5-year-olds, and 1 6-year-old**. There were **91 young** and **87 old** children (mediansplit based on entire sample).

# Testing location

72 children were from Edinburgh, 106 from Fife.

# Training 1

## How many children stopped Training 1?

From the 178 children who started the Shelf task, 2 children (ID 49 and 129) stopped Training 1.

## We select those children who completed Training 1.

**176 children have valid data** on Training 1.

## Reached criterion

Of the 176 children who completed Training 1, all but one child (ID 155) reached the criterion.

## Including this one child, how many trials were administered in Training 1?

Of all 176 children who attempted Training 1, the mean number of administered trials was 12.42 (SD = 4.26, range 7-36).

* 3 years (n = 80): 12.71 (SD = 3.89, range 7-29)
* 4 years (n = 84): 12.39 (SD = 4.80, range 7-36)
* 5 years (n = 11): 10.64 (SD = 1.36, range 8-12)
* 6 years (n = 1): 11

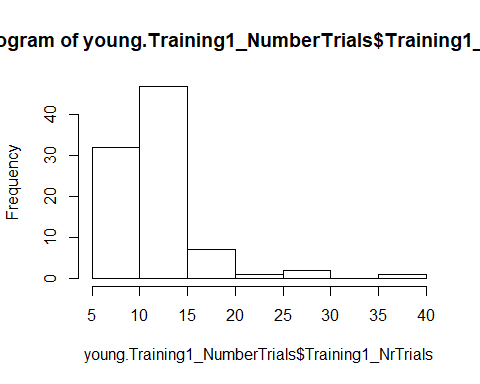
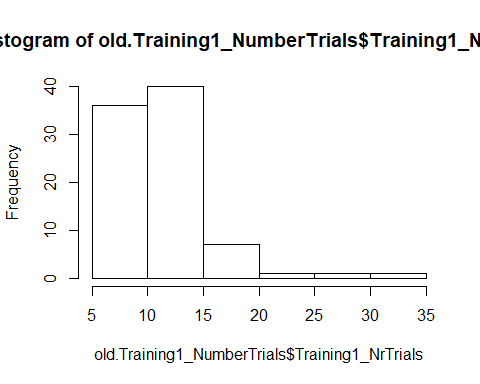
3- and 4-year-olds did not differ in the number of trials administered in trial 1.

* Young (n = 90): 12.83 (SD = 4.53, range 7-36)
* Old (n = 86): 11.99 (SD = 3.94, range 7-31)

Older children needed significantly fewer trials in Training 1 than younger children, W = 4453.5, p = .041.

* Fife (n = 104): 12.44 (SD = 4.00, range 8-29)
* Edinburgh (n = 72): 12.39 (SD = 4.63, range 7-36)

Children from Edinburgh and Fife did not differ in the number of administered trials.

## Including this one child who did not reach criterion, how many errors were made in Training 1?

Of all 176 children who attempted Training 1, the mean number of errors was 3.90 (SD = 2.39, range 1-18).

* 3 years (n = 80): 3.85 (SD = 2.03, range 1-12)
* 4 years (n = 84): 4.03 (SD = 2.83, range 1-18)
* 5 years (n = 11): 3.27 (SD = 0.90, range 2-5)
* 6 years (n = 1): 4

3- year-olds did not make more mistakes than 4-year-olds, W = 3414, p = .428.

* Young (n = 90): 3.93 (SD = 2.47, range 1-18)
* Old (n = 86): 3.87 (SD = 2.33, range 1-16)

There is no difference between young and old children in the number of errors.

* Fife (n = 104): 3.85 (SD = 2.20, range 1-12)
* Edinburgh (n = 72): 3.97 (SD = 2.66, range 1-18)

Children from Edinburgh and Fife did not differ in the number of errors.

## For those children who reached criterion, how many trials were administered?

Of the 175 children who reached the learning criterion in Training 1, the mean number of administered trials was 12.28 (SD = 3.87, range 7-31).

* 3 years (n = 80): 12.71 (SD = 3.89, range 7-29)
* 4 years (n = 83): 12.11 (SD = 4.05, range 7-31)
* 5 years (n = 11): 10.64 (SD = 1.36, range 8-12)
* 6 years (n = 1): 11

There was no difference between 3- and 4-year-olds.

* Young (n = 89): 12.57 (SD = 3.82, range 7-29)
* Old (n = 86): 11.99 (SD = 3.94, range 7-31)

Younger children were administered significantly more trials than older children, one-sided Wilcoxon test, W = 4453.5, p = .041.

* Fife (n = 104): 12.44 (SD = 4.00, range 8-29)
* Edinburgh (n = 71): 12.06 (SD = 3.70, range 7-31)

Children from Edinburgh and Fife did not differ in the number of administered trials.

## How many errors were made in Training 1 to reach the learning criterion?

Of the 175 children who reached the learning criterion in Training 1, the mean number of errors was 3.82 (SD = 2.15, range 1-16).

* 3 years (n = 80): 3.85 (SD = 2.03, range 1-12)
* 4 years (n = 83): 3.87 (SD = 2.38, range 1-16)
* 5 years (n = 11): 3.27 (SD = 0.90, range 2-5)
* 6 years (n = 1): 4

3- year-olds did not make more mistakes than 4-year-olds.

* Young (n = 89): 3.77 (SD = 1.97, range 1-12)
* Old (n = 86): 3.87 (SD = 2.33, range 1-16)

There is no difference between young and old children in the number of errors.

* Fife (n = 104): 3.85 (SD = 2.20, range 1-12)
* Edinburgh (n = 71): 3.77 (SD = 2.08, range 1-16)

Children from Edinburgh and Fife did not differ in the number of errors.

# Training 2a

## Of all 175 children who reached the learning criterion in Training 1, how many were successful in Training 2a?

Of all 175 children who reached the learning criterion in Training 1, 3 children (IDs 30, 34, and 105) did not reach the learning criterion in Training 2a.

## Including those children who did not reach the learning criterion, how many trials were administered in Training 2a?

Of all 175 children who attempted Training 2a, the mean number of administered trials was 11.49 (SD = 6.15, range 6-36).

* 3 years (n = 80): 11.79 (SD = 5.68, range 6-36)
* 4 years (n = 83): 11.67 (SD = 6.88, range 6-36)
* 5 years (n = 11): 8.27 (SD = 1.19, range 6-10)
* 6 years (n = 1): 8

3-year-olds were not administered more trials than 4-year-olds.

* Young (n = 89): 11.77 (SD = 5.97, range 6-36)
* Old (n = 86): 11.20 (SD = 6.35, range 6-36)

There is no difference between young and old children in the number of administered trials.

* Fife (n = 104): 11.51 (SD = 6.05, range 6-36)
* Edinburgh (n = 71): 11.46 (SD = 6.33, range 6-36)

Children from Edinburgh and Fife did not differ in the number of administered trials.

## Including the three children who did not reach criterion, how many errors were made in Training 2a?

Of all 175 children who attempted Training 2a, the mean number of errors was 3.48 (SD = 3.30, range 0-20).

* 3 years (n = 80): 3.66 (SD = 3.07, range 0-20)
* 4 years (n = 83): 3.54 (SD = 3.68, range 0-19)
* 5 years (n = 11): 1.91 (SD = 0.94, range 0-3)
* 6 years (n = 1): 2

3-year-olds did not make more errors than 4-year-olds.

* Young (n = 89): 3.68 (SD = 3.28, range 0-20)
* Old (n = 86): 3.28 (SD = 3.33, range 0-19)

There is no difference between young and old children in the number of errors.

* Fife (n = 104): 3.58 (SD = 3.35, range 0-20)
* Edinburgh (n = 71): 3.35 (SD = 3.26, range 0-19)

Children from Edinburgh and Fife did not differ in the number of errors.

## How many shifts were made in Training 2a?

Of all 175 children who attempted Training 2a, the mean number of shifts was 1.24 (SD = 0.50, range 1-3).

* 3 years (n = 80): 1.25 (SD = 0.49, range 1-3)
* 4 years (n = 83): 1.26 (SD = 0.54, range 1-3)
* 5 years (n = 11): 1 (SD = 0)
* 6 years (n = 1): 1

There was no differendce between 3- and 4-year-olds in terms of the number of shifts.

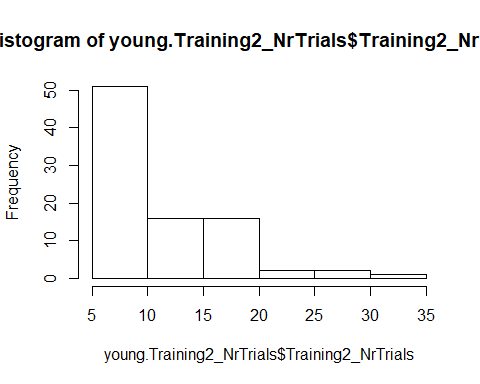
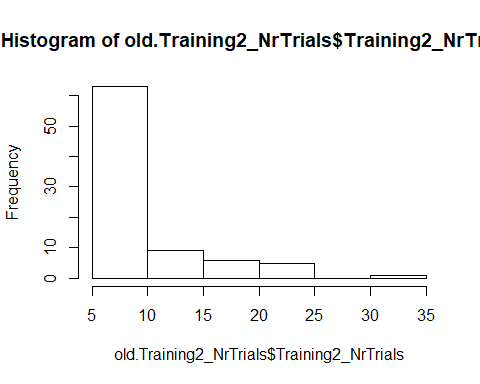
* Young (n = 89): 1.25 (SD = 0.48, range 1-3)
* Old (n = 86): 1.23 (SD = 0.52, range 1-3)

There is no difference between young and old children in the number of shifts.

* Fife (n = 104): 1.22 (SD = 0.50, range 1-3)
* Edinburgh (n = 71): 1.27 (SD = 0.50, range 1-3)

Children from Edinburgh and Fife did not differ in the number of shifts.

## When excluding the three children who did not reach criterion in Training 2a, how many trials were administered?

Of the 172 children who reached the learning criterion in Training 2a, the mean number of administered trials was 11.06 (SD = 5.27, range 6-34).

* 3 years (n = 79): 11.48 (SD = 5.01, range 6-28)
* 4 years (n = 81): 11.07 (SD = 5.77, range 6-34)
* 5 years (n = 11): 8.27 (SD = 1.19, range 6-10)
* 6 years (n = 1): 8

3-year-olds were not administered more trials than 4-year-olds.

* Young (n = 88): 11.50 (SD = 5.41, range 6-33)
* Old (n = 84): 10.61 (SD = 5.11, range 6-34)

There is no difference between young and old children in the number of trials, W = 3988, p = .183.

* Fife (n = 102): 11.03 (SD = 5.03, range 6-26)
* Edinburgh (n = 70): 11.11 (SD = 5.64, range 6-34)

Children from Edinburgh and Fife did not differ in the number of trials until criterion.

## How many errors were made in Training 2a to reach the learning criterion?

Of the 172 children who reached the learning criterion in Training 2a, the mean number of errors was 3.21 (SD = 2.57, range 0-17).

* 3 years (n =79): 3.45 (SD = 2.47, range 0-11)
* 4 years (n = 81): 3.16 (SD = 2.79, range 0-17)
* 5 years (n = 11): 1.91 (SD = 0.94, range 0-3)
* 6 years (n = 1): 2

3-year-olds did not make more errors than 4-year-olds, W = 3520, p = .133.

* Young (n = 88): 3.50 (SD = 2.79, range 0-17)
* Old (n = 84): 2.90 (SD = 2.29, range 0-12)

Young children do not make more errors than old children, W = 4130.5, p = .087.

* Fife (n = 102): 3.26 (SD = 2.51, range 0-11)
* Edinburgh (n = 70): 3.13 (SD = 2.67, range 0-17)

Children from Edinburgh and Fife did not differ in the number of errors.

## Number of shifts

Of the 172 children who reached the learning criterion in Training 2a, the mean number of shifts was 1.23 (SD = 0.48, range 1-3).

* 3 years (n = 79): 1.25 (SD = 0.49, range 1-3)
* 4 years (n = 81): 1.23 (SD = 0.51, range 1-3)
* 5 years (n = 11): 1.00 (SD = 0)
* 6 years (n = 1): 1

There was no difference between 3- and 4-year-olds in the number of shifts made in Training 2a.

* Young (n = 88): 1.25 (SD = 0.48, range 1-3)
* Old (n = 84): 1.20 (SD = 0.48, range 1-3)

Young and old children did not differ in the number of shifts.

* Fife (n = 102): 1.20 (SD = 0.47, range 1-3)
* Edinburgh (n = 70): 1.26 (SD = 0.50, range 1-3)

Children from Edinburgh and Fife did not differ in the number of shifts.

**172 children proceeded to Day 2.**

# Training 2b

**172 children qualified for Day 2, but due to Covid-19 we could not test 5 of them on the second session**. So **167 children started Day 2**.

## Dropouts

Out of the 167 children who started Training 2b, **1 child stopped the training** after 18 trials (ID 33).

## Excluding the child who stopped Training 2b, how many trials were administered?

**All of the 166 children who attempted Training 2b and did not drop out also reached the learning criterion**. Out of the 166 children who reached the learning criterion in Training 2b, the mean number of administered trials was 8.79 (SD = 4.12, range 6-30).

* 3 years (n = 77): 8.91 (SD = 4.55, range 6-30)
* 4 years (n = 77): 8.91 (SD = 3.91, range 6-26)
* 5 years (n = 11): 7.18 (SD = 1.89, range 6-12)
* 6 years (n = 1): 9

There was no difference between 3- and 4-year-olds in the number of trials administered.

* Young (n = 86): 8.80 (SD = 4.38, range 6-30)
* Old (n = 80): 8.79 (SD = 3.85, range 6-26)

Young and old children did not differ in the number of trials administered.

* Fife (n = 97): 9.25 (SD = 4.64, range 6-30)
* Edinburgh (n = 69): 8.16 (SD = 3.18, range 6-25)

Children from Edinburgh and Fife did not differ in the number of trials, W = 3743, p = .183.

## Excluding the child who stopped Training 2b, how many errors were made?

Out of the 166 children who reached the learning criterion in Training 2b, the mean number of errors was 1.64 (SD = 1.94, range 0-12).

* 3 years (n =77): 1.69 (SD = 2.18, range 0-12)
* 4 years (n =77): 1.69 (SD = 1.77, range 0-9)
* 5 years (n =11): 0.91 (SD = 1.30, range 0-4)
* 6 years (n =1): 2

There was no difference between 3- and 4-year-olds in the number of errors.

* Young (n = 86): 1.64 (SD = 2.10, range 0-12)
* Old (n = 80): 1.64 (SD = 1.77, range 0-9)

Young and old children did not differ in the number of errors.

* Fife (n = 97): 1.86 (SD = 2.18, range 0-12)
* Edinburgh (n = 69): 1.32 (SD = 1.52, range 0-9)

Children from Edinburgh and Fife did not differ in the number of errors, W = 3753, p = .170.

## Excluding the child who stopped Training 2b, how many shifts were made?

Out of the 166 children who reached the learning criterion in Training 2b, the mean number of shifts was 1.20 (SD = 0.43, range 1-3).

* 3 years (n =77): 1.18 (SD = 0.42, range 1-3)
* 4 years (n =77): 1.25 (SD = 0.46, range 1-3)
* 5 years (n =11): 1 (SD = 0)
* 6 years (n =1): 1

There was no difference between 3- and 4-year-olds in the number of shifts.

* Young (n = 86): 1.17 (SD = 0.41, range 1-3)
* Old (n = 80): 1.22 (SD = 0.45, range 1-3)

Young and old children did not differ in the number of shifts.

* Fife (n = 97): 1.24 (SD = 0.47, range 1-3)
* Edinburgh (n = 69): 1.14 (SD = 0.35, range 1-2)

Children from Edinburgh and Fife did not differ in the number of shifts.

# Test

## Dropouts

167 children reached the learning criterion in training 2b and thus proceeded to the test. How many dropouts where there in the test session?

There were **15 dropouts** from the test phase:

* 1 child stopped after 7 test trials (ID 13)
* 1 child stopped after 9 test trials (ID 10)
* 1 child stopped after 13 test trials (ID 39)
* 1 child stopped after 19 test trials (ID 43)
* 1 child stopped after 22 test trials (ID 17)
* 1 child: experimenter error in trial 23 (placed sticker under wrong cup; ID 22)
* 1 child: experimenter error trial 8 (ID 27)
* 7 children: experimenter error, E showed sticker after every unsuccessful trial
* 1 child: experimenter error: E delayed a switch and then missed a switch (ID 52)

We remove all dropouts, so **152 children remain**.

## Valid data

We include all those children who have 27 or more completed test trials. This includes:

* 1 child with 27 trials (because of experimenter error, ID46)
* 1 child with 32 trials (ID 143)
* 1 child with 33 trials (because of experimenter error, ID 55)
* 7 children with 35 trials (because of experimenter error, IDs 25, 52, 82, 127, 154, 165, 139)

There were 84 girls and 68 boys with valid data in the test phase.

## How many test trials were administered?

There are 152 valid cases in the Shelf test. Of these, children were administered on average 35.88 trials (SD = 0.60, range 31-36).

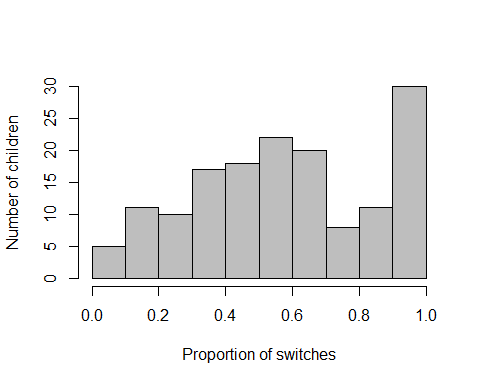
* 3 years (n = 68): 35.87 (SD = 0.62, range 32-36)
* 4 years (n = 77): 35.87 (SD = 0.63, range 31-36)
* 5 years (n = 11): 36 (SD = 0)
* 6 years (n = 1): 36
* Young (n = 76): 35.88 (SD = 0.59, range 32-36)
* Old (n = 76): 35.88 (SD = 0.61, range 31-36)

Young and old children did not differ in the number of administered test trials.

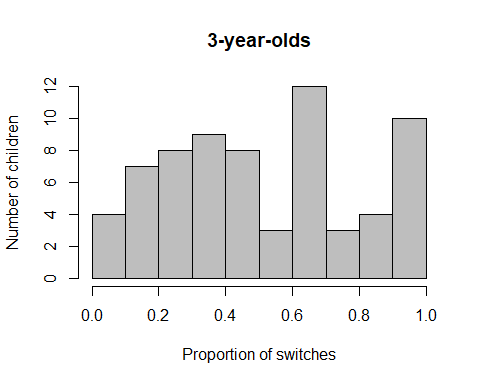
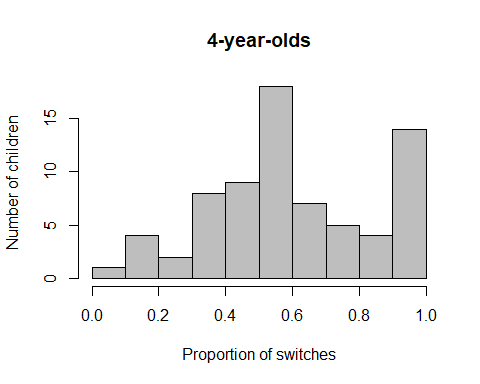
* Fife (n = 86): 35.86 (SD = 0.70, range 31-36)
* Edinburgh (n = 66): 35.91 (SD = 0.42, range 33-36)

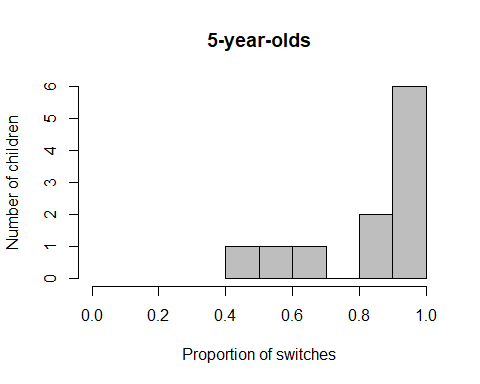
Children from Edinburgh and Fife did not differ in the number of administered test trials.

## DV: What was the proportion of switches the children achieved out of the number of possible switches?



Of the 152 children, the proportion of switches achieved out of the possible number of switches (given the number of trials administered) was 0.57 (SD = 0.26, range 0-1). The variable is not normally distributed, W = 0.959, p < .001. Chance level: 0.2242245 Children’s performance is significantly above chance level, V = 11425, p < .001.



* 3 years (n = 68): 0.51 (SD = 0.26, range 0.09-1), not normally distributed, performance better than chance, V = 2248, p < .001
* 4 years (n = 72): 0.59 (SD = 0.24, range 0-1), not normally distributed, performance better than chance, V = 2603, p < .001
* 5 years (n = 11): 0.82 (SD = 0.19, range 0.45-1), not normally distributed, performance better than chance, V = 66, p = .003
* 6 years (n = 1): 0.82

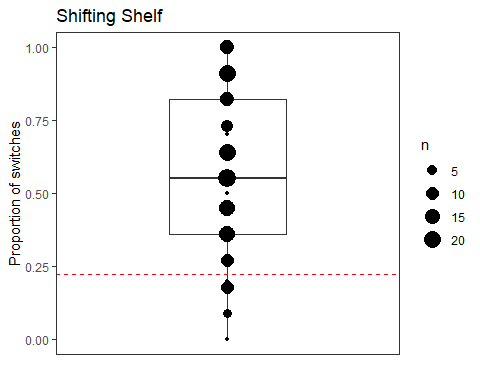
3-year-olds performed fewer switches than 4-year-olds, W = 1985, p = .026.

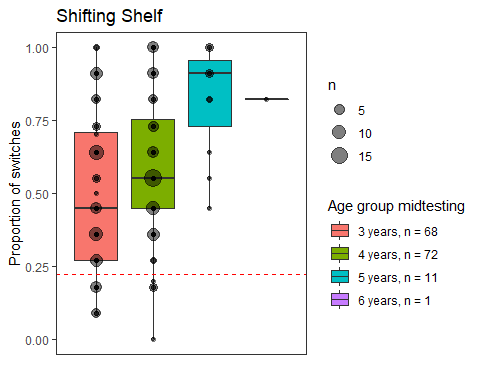
* Young (n = 76): 0.52 (SD = 0.26, range 0.09-1), not normally distributed, performance better than chance, V = 2816, p < .001
* Old (n = 76): 0.63 (SD = 0.24, range 0-1), not normally distributed, performance better than chance, V = 2908, p < .001

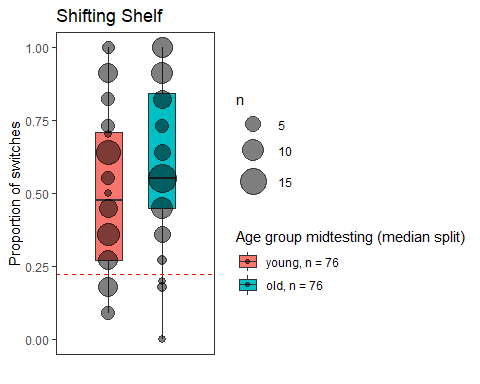
Young performed fewer switches than older children, W = 2145.5, p = .003.

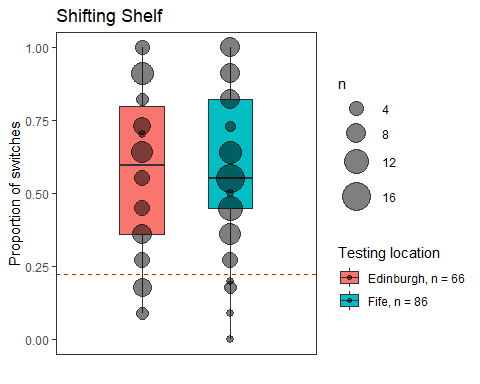
* Fife (n = 86): 0.58 (SD = 0.25, range 0-1), not normally distributed, performance better than chance, V = 3701, p < .001
* Edinburgh (n = 66): 0.56 (SD = 0.27, range 0.09-1), not normally distributed, performance better than chance, V = 2141, p < .001

Children from Edinburgh and Fife did not differ in the proportion of switches.

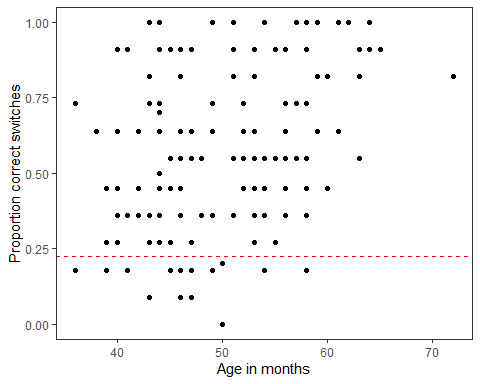








## Plot age as continuous variable against proportion of switches



## How many switches were possible?

The number of possible switches was either 10 or 11. 149 children had 11 possible switches, 3 children had 10 possible switches.

## Number of achieved switches

Of the 152 children, the mean number of switches achieved was 6.32 (SD = 2.83, range 0-11).

* 3 years: 5.62 (SD = 2.88, range 1-11)
* 4 years: 6.64 (SD = 2.62, range 0-11)
* 5 years: 9.00 (SD = 2.10, range 5-11)
* 6 years: 9

## Number of errors

Of the 152 children, the mean number of errors achieved was 11.89 (SD = 7.10, range 0-36).

* 3 years (n = 68): 13.72 (SD = 7.33, range 0-33), normally distributed, performance better than chance (18 mistakes), t(67) = -4.810, p < .001
* 4 years (n = 72): 11.39 (SD = 6.46, range 0-36), not normally distributed, performance better than chance, V = 121, p < .001
* 5 years: 4.64 (SD = 3.93, range 1-12), not normally distributed, performance better than chance, V = 0, p = .004
* 6 years: 4

4-year-olds made more mistakes than 5-year-olds, W = 635, p = .001.

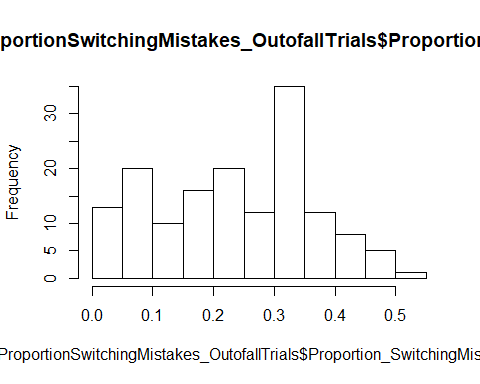
* Young (n = 76): 13.50 (SD = 7.20, range 0-33), normally distributed, performance better than chance (18 mistakes), t(75) = -5.449, p < .001
* Old (n = 76): 12.26 (SD = 7.54, range 0-31), not normally distributed, performance better than chance, V = 110, p < .001

Young made more mistakes than older children, W = 2667.5, p = .002.

* Fife (n = 86): 11.62 (SD = 6.77, range 0-36), not normally distributed, performance better than chance, V = 189, p < .001
* Edinburgh (n = 66): 12.26 (SD = 7.54, range 0-31), normally distributed, performance better than chance, t(65) = -6.18, p < .001.

Children from Edinburgh and Fife did not differ in the number of mistakes.

## Proportion of switching mistakes out of all administered trials



The mean proportion of switching mistakes out of all trials is 0.23 (SD = 0.13, range 0-0.53). Note that only those children are included where we could exactly determine the number of switching mistakes. (1 child missing, ID 26, a 4-year-old).

* 3 years (n = 68): 0.26 (SD = 0.13, range 0-0.5)
* 4 years (n = 71!): 0.23 (SD = 0.12, range 0-0.53)
* 5 years (n = 11): 0.11 (SD = 0.09, range 0.03-0.30)
* 6 years: 0.08
* Young (n = 76): 0.26 (SD = 0.13, range 0-0.5)
* Old (n = 75!): 0.21 (SD = 0.12, range 0-0.53)

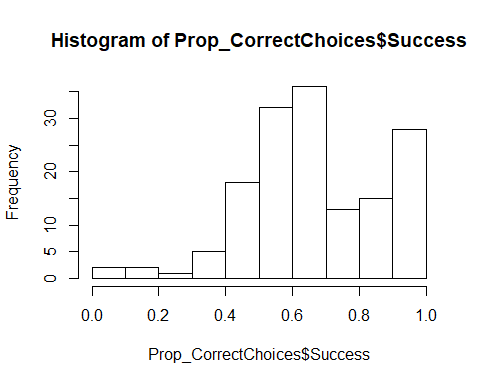
## Number of switching mistakes

Of the 152 children, the mean number of switching mistakes is 8.39 (SD = 4.56, range 0-19).

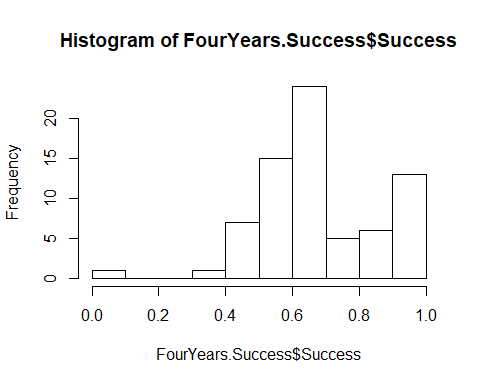
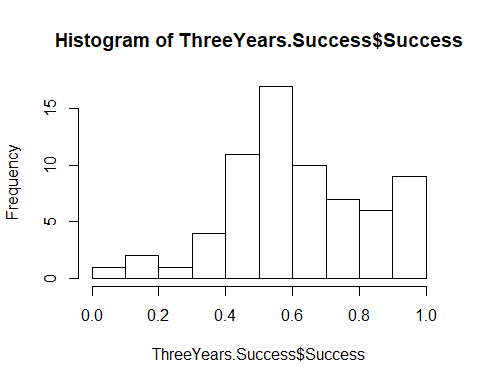
* 3 years: 9.37 (SD = 4.59, range 0-18)
* 4 years: 8.24 (SD = 4.30, range 0-19)
* 5 years: 3.91 (SD = 3.18, range 1-11)
* 6 years: 3

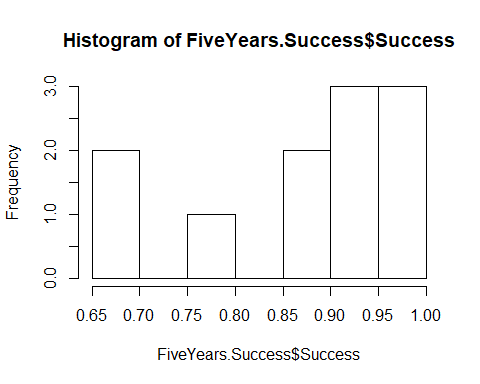
## Mean number of correct choices

The mean number of correct trials is 0.67 (SD = 0.20, range 0-1).



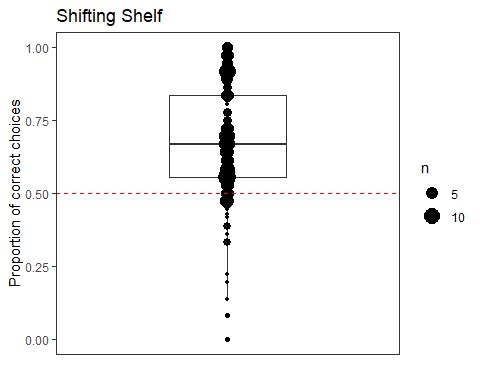
Children’s proportion of correct trials (M = 0.67, SD = 0.20, range 0-1) is significantly higher than what would be expected if we assumed that children would choose randomly between the two cups on the winning platform, V = 9671, p <. 001.

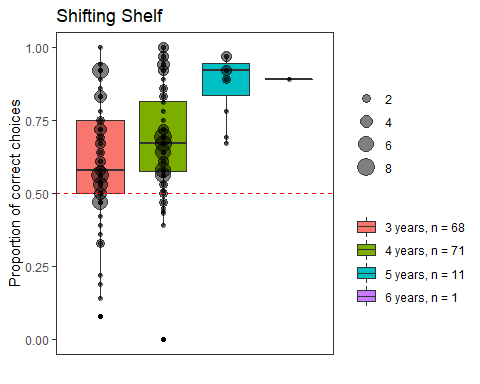


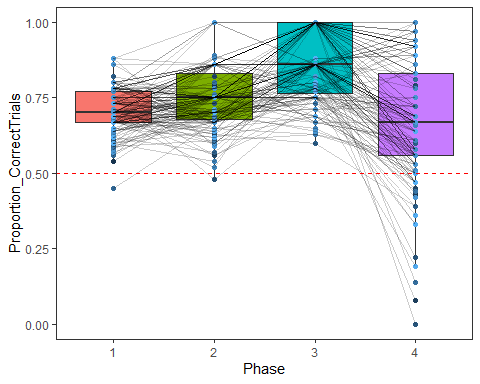


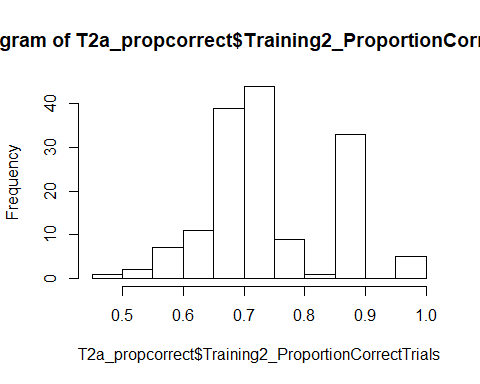
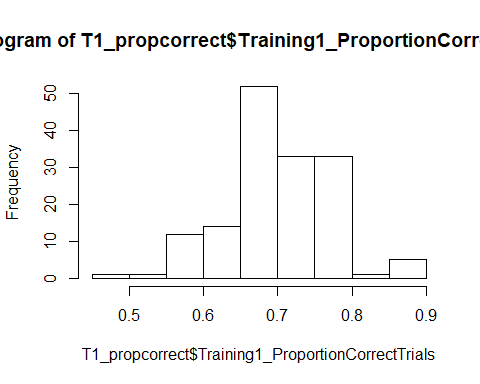
Mean number of correct choices:

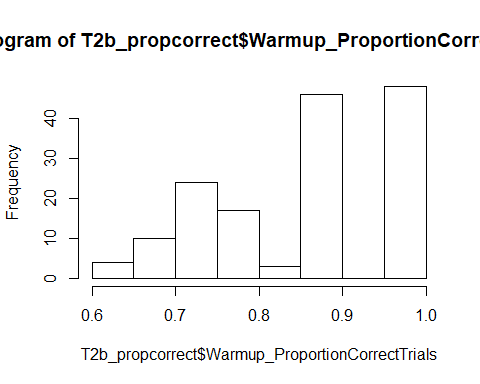
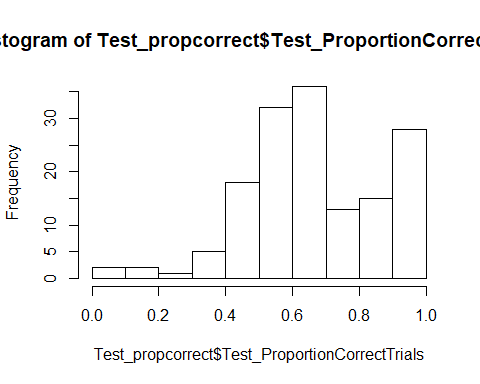
* 3 years (n = 68): 0.61 (SD = 0.20, range 0.08-1), perform significantly different from chance (0.5), V = 1692.5, p < .001
* 4 years (n = 72): 0.68 (SD = 0.18, range 0-1), perform significantly different from chance (0.5), V = 2358.5, p < .001
* 5 years (n = 11): 0.87 (SD = 0.11, range 0.67-0.97), perform significantly different from chance (0.5), V = 66, p = .004
* 6 years (n = 1): 0.89









Training 1: Performance significantly above chance (0.5), V = 11626, p < .001

Training 2a: Performance significantly above chance (0.5), V = 11627, p < .001

Training 2b: Performance significantly above chance (0.5), V = 11628, p < .001

Test: Performance significantly above chance (0.5), V = 9705.5, p < .001

# In the test phase, can age and trial number and the interaction between age and trial number predict success?

#Full model  
control=glmerControl(optimizer="bobyqa",optCtrl=list(maxfun=100000))  
  
res<-glmer(Success ~ z.age.midtesting + z.Trial\_nr + z.age.midtesting:z.Trial\_nr + (1+z.Trial\_nr|ID), data=Shelf.clean.test, family = binomial, control = control)

The full model explains the data significantly better than the null model only consisting of the intercept, X2(3) = 50.714, p < .001.

There is only a marginally significant interaction effect, X2(1) = 3.557, p = .059, so we remove it from the model.

res<-glmer(Success ~ z.age.midtesting + z.Trial\_nr + (1+z.Trial\_nr|ID), data=Shelf.clean.test, family = binomial, control=control)

The full model explains the data significantly better than the null model only consisting of the intercept, X2(2) = 47.157, p < .001.

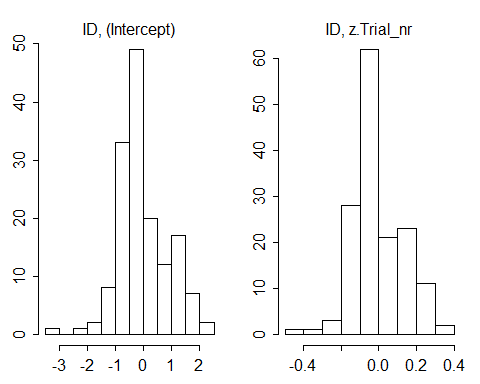
Trial number has a significant effect, X2(1) = 28.231, p < .001.

Age has a significant effect, X2(1) = 16.943, p < .001.

summary(res)

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: Success ~ z.age.midtesting + z.Trial\_nr + (1 + z.Trial\_nr | ID)  
## Data: Shelf.clean.test  
## Control: control  
##   
## AIC BIC logLik deviance df.resid   
## 6297.8 6337.4 -3142.9 6285.8 5452   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -4.3465 -0.9851 0.3782 0.7331 2.2350   
##   
## Random effects:  
## Groups Name Variance Std.Dev. Corr  
## ID (Intercept) 1.01081 1.005   
## z.Trial\_nr 0.02658 0.163 0.94  
## Number of obs: 5458, groups: ID, 152  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.87999 0.08903 9.884 < 2e-16 \*\*\*  
## z.age.midtesting 0.35794 0.08539 4.192 2.77e-05 \*\*\*  
## z.Trial\_nr 0.19145 0.03654 5.239 1.61e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) z.g.md  
## z.g.mdtstng 0.013   
## z.Trial\_nr 0.374 -0.086

## Model assumptions



#collinearity

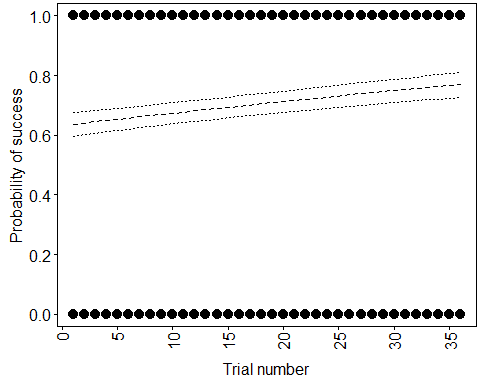
## z.Trial\_nr z.age.midtesting   
## 1 1

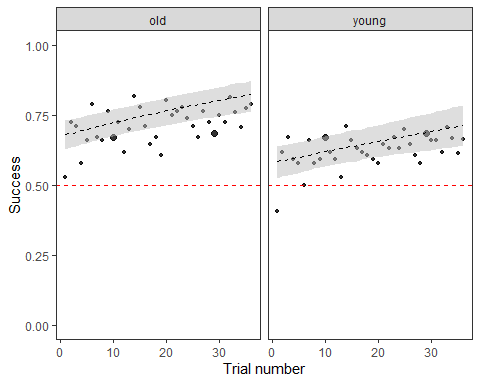
m.stab$summary

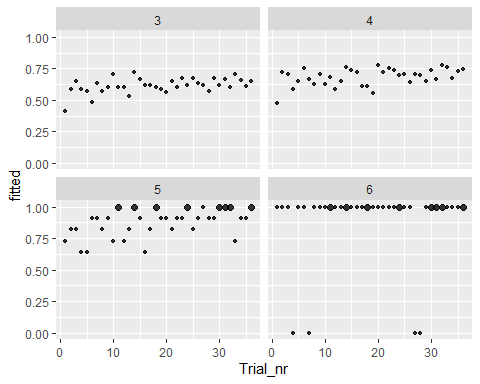
## what orig min max  
## (Intercept) (Intercept) 0.8799864 0.8585752 0.8979880  
## z.age.midtesting z.age.midtesting 0.3579433 0.3393107 0.3775191  
## z.Trial\_nr z.Trial\_nr 0.1914479 0.1827306 0.1983459  
## ID@(Intercept) ID@(Intercept) 1.0026909 0.9737495 1.0056344  
## ID@z.Trial\_nr ID@z.Trial\_nr 0.4037608 0.3653676 0.4204540

boot.res$ci.estimates

## orig X2.5. X97.5.  
## (Intercept) 0.8799864 0.7141163 1.0575051  
## z.age.midtesting 0.3579433 0.1926725 0.5217537  
## z.Trial\_nr 0.1914479 0.1221226 0.2602241







# Mean success by trial

