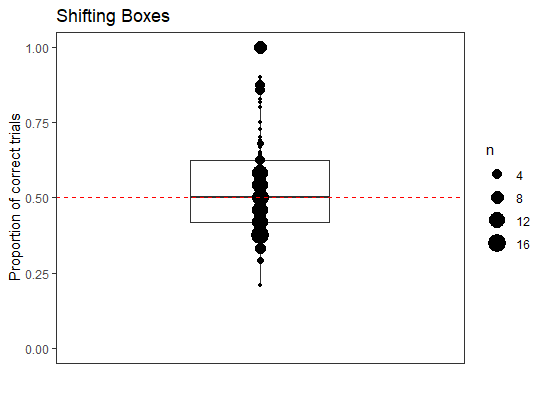
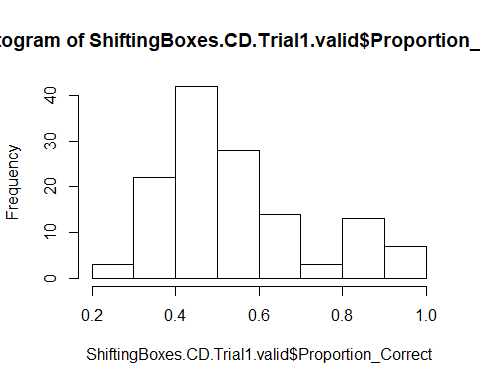
Shifting Boxes

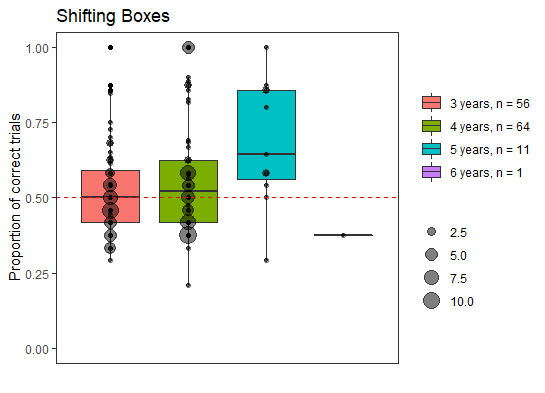
Eva Reindl

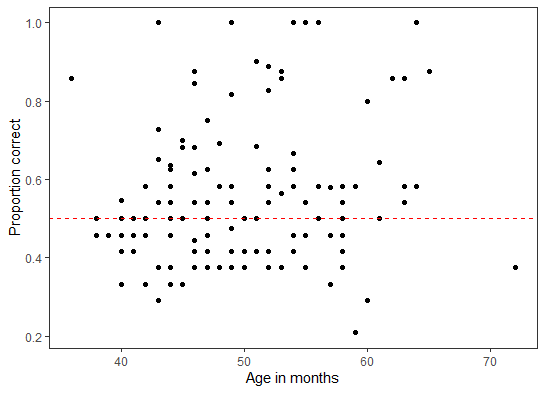
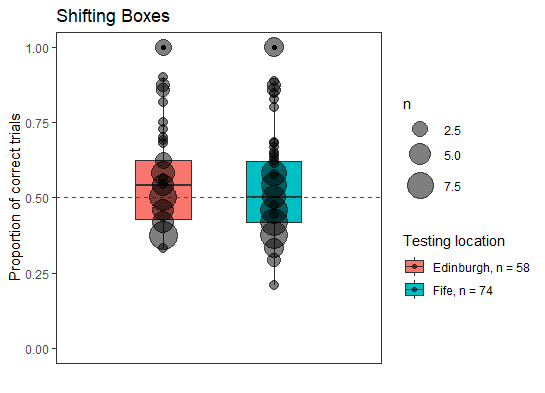
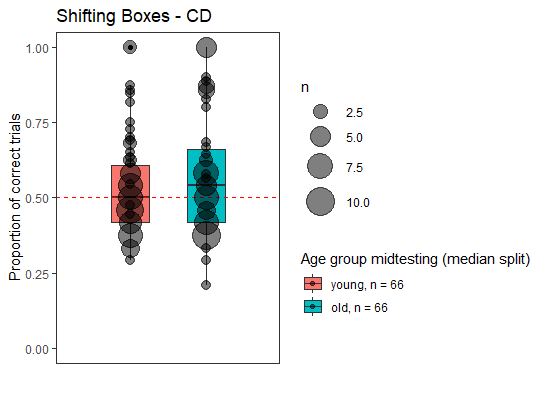
16 6 2020

Key findings:

* 132 have valid data on the CD (regardless of whether they reached criterion in SD)
  + Mean proportion of correct trials in CD: **0.56 (SD = 0.18, range 0.21-1)**
  + **not normally distributed,** W = 0.908, p < .001
  + Performance is significantly **above chance across groups**, V = .438, p = .020, **for 5-year-olds** and older **children**, but not for 3- and 4-year-olds nor young children
  + **No difference between young and old children, no difference between Edinburgh and Fife children**
  + No effect of age or trial number on success

**



****

# Summary table

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **SD** | **Reversal** | **Compound Discrimination – all children who started CD** | **Compound Discrimination – only children who reached criterion in SD** | **EDS for all children who started CD** | **EDS – only children who reached criterion in SD** |
| Started | 144 | 111 | 139 | 106 | 36 | 26 |
| Completed/valid | 141 | 102 (1 of them only did 22 trials, unclear whether reached criterion) | 132 (1 of them stopped after 18 trials, 1 after 19 trials, and with 1 child there was an experimenter error, E switched to EDS after 22 trials even though only 5 in a row were correct (ID38); and for all it is unclear whether they would have reached criterion, so they are removed those variables) | 101 (1 child completed 19 trials) | 35 (including 1 child where E accidentally switched to EDS after only 5 consecutive trials in CD, 1 child who stopped after 33 trials, 1 child where E accidentally only administered 35 trials, 1 child with 35 trials because E accidentally skipped one trial, 1 child stopped after 27 trials) | 25 (including 1 child where E accidentally switched to EDS after only 5 consecutive correct in CD, 1 child who stopped after 33 trials, and 2 children with 35 trials (one child stopped, with one child E missed a trial), 1 child stopped after 27 trials |
| Reached criterion | 111 (79%) | 92 (90%) | 36 (27%) (plus there are 3 children for whom it is unclear) | 27 (27%) | / | / |
| Trials administered (excluding dropouts) | 13.73 (SD = 6.62, range 6-24)  3y (n = 60): 13.95 (SD = 6.65, range 6-24)  4y (n = 69): 13.72 (SD = 6.65, range 6-24)  5y (n = 11): 13 (SD = 7.56, range 6-24)  6y (n = 1): 9  young (n = 70): 14.40 (SD = 6.64, range 6-24)  old (n = 71): 13.20 (SD = 6.43, range 6-24) | 12.50 (SD = 5.71, range 6-24)  3y (n = 38): 13.79 (SD = 5.77, range 6-24)  4y (n = 55): 11.85 (SD = 5.56, range 6-24)  5y (n = 8): 11.50 (SD = 6.19, range 7-24)  6y (n =1): 7  young (n = 44): 13.79 (SD = 5.78, range 6-24)  old (n = 58): 11.52 (SD = 5.51, range 6-24) | 20.70 (SD = 6.01, range 6-24)  3y (n = 56): 21.68 (SD = 5.06, range 6-24)  4y (n = 64): 20.72 (SD = 6.16, range 6-24)  5y (n = 11): 15.36 (SD = 7.38, range 6-24)  6y (n = 1): 24  young (n = 66): 21.32 (SD = 5.37, range 6-24)  old (n = 66): 20.09 (SD = 6.56, range 6-24) | 20.69 (SD = 6.11, range 6-24)  3y (n = 39) : 22.02 (SD = 4.83, range 6-24)  4y (n = 53) : 20.49 (SD = 6.33, range 6-24)  5y (n = 8) : 15.12 (SD = 7.92, range 6-24)  6y (n = 1) : 24  young (n = 46): 21.69 (SD = 4.96, range 6-24)  old (n = 55): 19.85 (SD = 6.86, range 6-24) | 35.60 (SD = 1.59, range 27-36)  3y (n = 11): 35.64 (SD = 0.92, range 33-36)  4y (n = 17): 35.41 (SD = 2.18, range 27-36)  5y (n = 7): 36  young (n = 14): 35.71 (SD = 0.82, range 33-36)  old (n = 21): 35.52 (SD = 1.96, range 27-36) | 35.60 (SD = 1.80, range 27-36)  3y (n = 6): 36  4y (n = 14): 35.28 (SD = 2.40, range 27-36)  5y (n = 5): 36  young (n = 8): 36  old (n = 17): 35.41 (SD = 2.18, range 27-36) |
| Number of correct trials (excluding dropouts) | 9.30 (SD = 3.10, range 6-17)  3y : 9.10 (SD = 2.77, range 6-15)  4y : 9.62 (SD = 3.38, range 6-17)  5y : 8.64 (SD = 3.11, range 6-15)  6y: 7  young (n = 70): 9.48 (SD = 3.48, range 6-17)  old (n = 71): 9.13 (SD = 3.20, range 6-17) | 8.80 (SD = 3.14, range 6-17)  3y : 9.37 (SD = 3.36, range 6-17)  4y : 8.64 (SD = 3.07, range 6-17)  5y: 7.62 (SD = 2.20, range 6-11)  6y: 6  young (n = 44): 9.43 (SD = 3.34, range 6-17)  old (n = 58): 8.33 (SD = 2.92, range 6-17) | 10.61 (SD = 2.71, range 5-19)  3y : 10.84 (SD = 2.54, range 6-15)  4y : 10.65 (SD = 2.77, range 5-19)  5y : 9.36 (SD = 3.23, range 6-14)  6y: 9  young (n = 66): 10.74 (SD = 2.54, range 6-15)  old (n = 66): 10.48 (SD = 2.89, range 5-19) | 10.63 (SD = 2.82, range 5-19)  3y: 11.02 (SD = 2.60, range 6-15)  4y: 10.68 (SD = 2.87, range 5-19)  5y: 8.62 (SD = 3.16, range 6-14)  6y: 9  young (n = 46): 11.04 (SD = 2.55, range 6-15)  old (n = 55): 10.29 (SD = 3.02, range 5-19) | 29.63 (SD = 4.84, range 18-36)  3y: 27.64 (SD = 5.50, range 18-35)  4y: 29.88 (SD = 4.38, range 21-34)  5y: 32.14 (SD = 4.02, range 24-36)  young (n = 14): 27.71 (SD = 4.94, range 18-35)  old (n = 21): 30.90 (SD = 4.44, range 21-36) | 30.72 (SD = 4.06, range 21-36)  3y: 30.33 (SD = 3.72, range 24-35)  4y: 30.28 (SD = 4.08, range 21-34)  5y: 32.40 (SD = 4.77, range 24-36)  young (n = 8): 29.50 (SD = 3.66, range 24-35)  old (n = 17): 31.29 (SD = 4.21, range 21-36) |
| Errors (excluding dropouts) | 4.42 (SD = 4.23, range 0-16)  3y: 4.85 (SD = 4.63, range 0-16)  4y: 4.10 (SD = 3.79, range 0-14)  5y: 4.36 (SD = 5.04, range 2-15)  6y: 2  young (n = 70): 4.91 (SD = 4.52 range 0-16)  old (n = 71): 3.94 (SD = 3.93, range 0-15) | 3.70 (SD = 3.25, range 0-15)  3y: 4.42 (SD = 3.32, range 0-15)  4y: 3.22 (SD = 3.03, range 0-15)  5y: 3.87 (SD = 4.15, range 1-13)  6y: 1  young (n = 44): 4.36 (SD = 3.31, range 0-15)  old (n = 58): 3.19 (SD = 3.13, range 0-15) | 10.09 (SD = 4.94, range 0-19)  3y: 10.84 (SD = 4.28, range 0-17)  4y: 10.06 (SD = 5.12, range 0-19)  5y: 6.00 (SD = 5.46, range 0-17)  6y: 15  young (n = 66): 10.57 (SD = 4.46, range 0-17)  old (n = 66): 9.61 (SD = 5.37, range 0-19) | 10.06 (SD = 5.05, range 0-19)  3y: 11.00 (SD = 4.32, range 0-17)  4y: 9.81 (SD = 5.25, range 0-19)  5y: 6.50 (SD = 5.95, range 0-17)  6y: 15  young (n = 46): 10.65 (SD = 4.33, range 0-17)  old (n = 55): 9.56 (SD = 5.57, range 0-19) | 5.97 (SD = 4.67, range 0-17)  3y: 8.00 (SD = 5.14, range 1-17)  4y: 5.53 (SD = 4.33 range 2-15)  5y: 3.86 (SD = 4.02, range 0-12)  young (n = 14): 8.00 (SD = 4.62, range 1-17)  old (n = 21): 4.62 (SD = 4.28, range 0-15) | 4.88 (SD = 3.94, range 0-15)  3y: 5.67 (SD = 3.72, range 1-12)  4y: 5.00 (SD = 3.92, range 2-15)  5y: 3.60 (SD = 4.77, range 0-12)  young (n = 8): 6.50 (SD = 3.66, range 1-12)  old (n = 17): 4.12 (SD = 3.93, range 0-15) |
| Proportion of correct trials (excluding dropouts) | 0.74 (SD = 0.15, range 0.33-1)  3y: 0.72 (SD = 0.16, range 0.33-1)  4y: 0.75 (SD = 0.14, range 0.42-1)  5y: 0.75 (SD = 0.18, range 0.37-1)  6y: 0.78  young (n = 70): 0.72 (SD = 0.16, range 0.33-1)  old (n = 71): 0.76 (SD = 0.14, range 0.37-1) | 0.74 (SD = 0.12, range 0.29-1)  3y: 0.71 (SD = 0.12, range 0.29-1)  4y: 0.77 (SD = 0.11, range 0.37-1)  5y: 0.72 (SD = 0.13, range 0.46-.86)  6y: 0.86  young (n = 44): 0.71 (SD = 0.12, range 0.29-1)  old (n = 58): 0.76 (SD = 0.11, range 0.37-1) | 0.56 (SD = 0.18, range 0.21-1)  3y: 0.53 (SD = 0.15, range 0.29-1)  4y: 0.56 (SD = 0.20, range 0.21-1)  5y: 0.68 (SD = 0.21, range 0.29-1)  6y: 0.37  young (n = 66): 0.53 (SD = 0.16, range 0.29-1)  old (n = 66): 0.58 (SD = 0.20, range 0.21-1) | 0.56 (SD = 0.19, range 0.21-1)  3y: 0.53 (SD = 0.16, range 0.29-1)  4y: 0.57 (SD = 0.20, range 0.21-1)  5y: 0.66 (SD = 0.23, range 0.29-1)  6y: 0.37  young (n = 46): 0.53 (SD = 0.15, range 0.29-1)  old (n = 55): 0.58 (SD = 0.21, range 0.21-1) | 0.83 (SD = 0.13, range 0.5-1)  3y: 0.77 (SD = 0.15, range 0.51-0.97)  4y: 0.84 (SD = 0.12, range 0.58-0.94)  5y: 0.89 (SD = 0.11, range 0.67-1)  young (n = 14): 0.77 (SD = 0.13, range 0.51-0.97)  old (n = 21): 0.87 (SD = 0.12, range 0.57-1) | 0.86 (SD = 0.11, range 0.58-1)  3y: 0.84 (SD = 0.10, range 0.67-0.97)  4y: 0.86 (SD = 0.11, range 0.58-.94)  5y: 0.90 (SD = 0.13, range 0.67-1)  young (n = 8): 0.82 (SD = 0.10, range 0.67-0.97)  old (n = 17): 0.88 (SD = 0.11, range 0.58-1) |
| Trials needed to reach criterion | 10.95 trials (SD = 4.39, range 6-24)  3y (n = 45): 10.60 (SD = 3.65, range 6-20)  4y (n = 57): 11.56 (SD = 4.99, range 6-24)  5y (n = 8): 8.75 (SD = 3.19, range 6-16)  6y (n = 1): 9  young (n = 52): 11.08 (SD = 3.99, range 6-22)  old (n = 59): 10.85 (SD = 4.74, range 6-24) | 11.41 trials (SD = 4.70, range 6-24)  3y (n = 33): 12.65 (SD = 4.95, range 6-24)  4y (n = 51): 10.90 (SD = 4.54, range 6-24)  5y (n = 7): 9.71 (SD = 3.86, range 7-18)  6y (n = 1): 7  young (n = 38): 12.54 (SD = 4.84, range 6-24)  old (n = 54): 10.59 (SD = 4.47, range 6-24) | 13.31 trials (SD = 6.51, range 6-24)  3y (n = 11): 14.61 (SD = 6.61, range 6-24)  4y (n = 18): 13.47 (SD = 7.18, range 6-24)  5y (n = 7): 10.43 (SD = 3.55, range 6-14)  young (n = 14): 13.57 (SD = 6.34, range 6-24)  old (n = 22): 13.00 (SD = 6.75, range 6-24) | 13.03 trials (SD = 6.46, range 6-24)  3y (n = 7): 15.44 (SD = 6.89, range 6-24)  4y (n = 15): 12.69 (SD = 6.64, range 6-24)  5y (n = 5): 9.80 (SD = 3.90, range 6-14)  young (n = 9): 14.82 (SD = 6.34, range 6-24)  old (n = 18): 12.00 (SD = 6.46, range 6-24) | / | / |
| Errors made to criterion | 2.47 errors (SD = 1.92, range 0-9)  3y: 2.42 (SD = 1.79, range 0-9)  4y: 2.63 (SD = 2.08, range 0-8)  5y: 1.62 (SD = 1.50, range 0-5)  6y: 2  young (n = 52): 2.52 (SD = 1.84, range 0-9)  old (n = 59): 2.42 (SD = 2.01, range 0-8) | 2.98 errors (SD = 2.31, range 0-15)  3y: 3.79 (SD = 2.90, range 0-15)  4y: 2.53 (SD = 1.74, range 0-7)  5y: 2.57 (SD = 2.07, range 1-7)  6y: 1  young (n = 38): 3.64 (SD = 2.76, range 0-15)  old (n = 54): 2.50 (SD = 1.81, range 0-7) | 3.90 errors (SD = 3.90, range 0-15)  3y: 4.92 (SD = 4.21, range 0-15)  4y: 3.68 (SD = 4.08, range 0-11)  5y: 2.57 (SD = 2.51, range 0-7)  young (n = 14): 4.37 (SD = 4.01, range 0-15)  old (n = 22): 3.35 (SD = 3.87, range 0-11) | 3.67 errors (SD = 3.86, range 0-15)  3y: 5.11 (SD = 4.57, range 0-15)  4y: 3.12 (SD = 3.67, range 0-11)  5y: 2.80 (SD = 3.03, range 0-7)  young (n = 9): 4.73 (SD = 4.20, range 0-15)  old (n = 18): 3.05 (SD = 3.63, range 0-11) | / | / |

"R version 3.6.1 (2019-07-05)"

**144 children started the Shifting Boxes Task**.

# Dropouts in SD

Out of the 144 children who started the Shifting Boxes Task, **3 children** (2 f, 1 m, all 3 years old) dropped out in the SD phase: **1 child stopped after 12 trials, and 2 children stopped after 14 trials**. We remove these children from further analyses, as none of these children completed 75% of the maximum number of possible trials.

# Valid data in SD

**141 children** have valid data on the SD phase.

# Description of sample on SD

## Description of gender distribution in SD

There were **76 girls** and **65 boys**.

## Description of age distribution in SD

### Age at beginning of testing

The 141 children who had valid data on the SD phase were on average 48.82 months old (SD = 6.97, range 36-70) at the beginning of testing. There were 63 3-year-olds, 70 4-year-olds, and 8 5-year-olds.

* 3-year-olds: 36 f, 27 m
* 4-year-olds: 36 f, 34 m
* 5-year-olds: 4 f, 4 m

### Age in the middle of testing

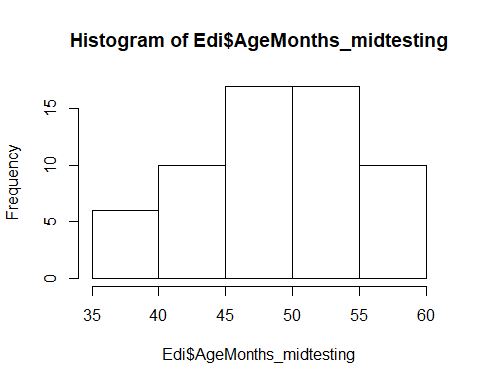
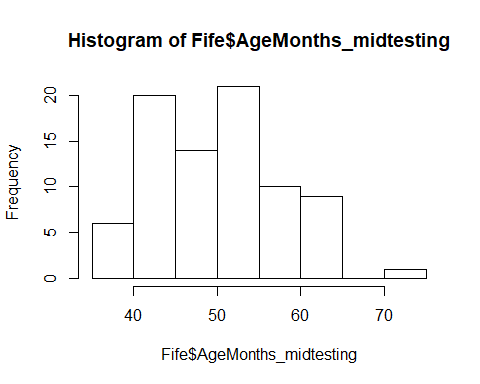
The 141 children who had valid data on the SD phase were on average **49.95 months old (SD = 6.91, range 36-72)** in the middle of testing. There were **60 3-year-olds**, **69 4-year-olds**, **11 5-year-olds**, and **1 6-year-old**.

* 3-year-olds: 35 f, 25 m
* 4-year-olds: 34 f, 35 m
* 5-year-olds: 7 f, 4 m
* 6-year-old: 1 m

### Age at mediansplit (based on entire sample)

There are **70 young** children (38 f, 32 m) and **71 old** children (38 f, 33 m).

## Testing location

81 children were from Fife, 60 children from Edinburgh.

Fife: M = 50.49 (SD = 7.62, range 38-72)

* 35 3y
* 34 4y
* 11 5y
* 1 6y

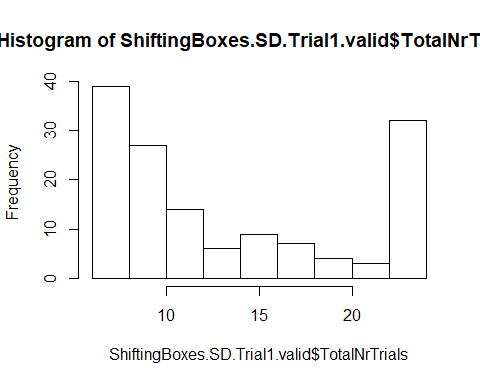
Edinburgh: M = 49.22 (SD = 5.80, range 36-58)

* 25 3y
* 35 4y

There is no difference in the age distribution between the two testing locations, two-sided Wilcoxon test, W = 2298, p = .583.

# Performance in SD

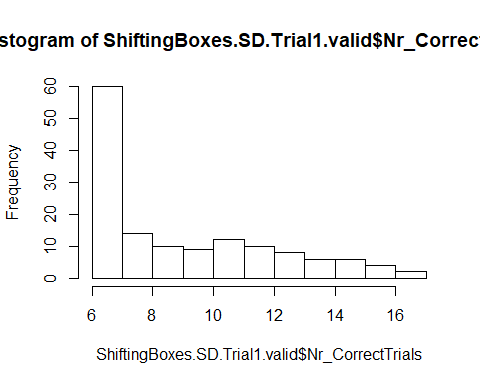
## Total number of trials



In the SD phase, the mean number of administered trials was **13.73 (SD = 6.62, range 6-24)**. The distribution of this variable is **bimodal**, with children either solving the task relatively quickly or maxing out on the total number of possible trials. 50% of the children were administered 11 or fewer trials, 75% were administered 20 or fewer trials.

* 3y (n = 60): 13.95 (SD = 6.65, range 6-24)
* 4y (n = 69): 13.72 (SD = 6.56, range 6-24)
* 5y (n = 11): 13.00 (SD = 7.56, range 6-24)
* 6y (n = 1): 9
* young (n = 70): 14.40 (SD = 6.64, range 6-24)
* old (n = 71): 13.07 (SD = 6.58, range 6-24)
* Fife (n = 81): 13.20 (SD = 6.43, range 6-24)
* Edinburgh (n = 60): 14.45 (SD = 6.86, range 6-24)

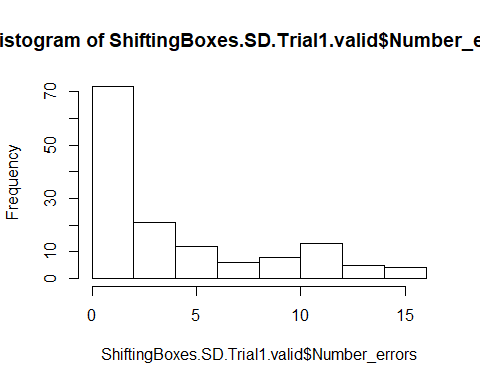
## Number of correct trials



In the SD phase, the mean number of correct trials was 9.30 (SD = 3.10, range 6-17). The distribution of this variable is very left-skewed.

* 3y (n = 60): 9.10 (SD = 2.77, range 6-15)
* 4y (n = 69): 9.62 (SD = 3.38, range 6-17)
* 5y (n = 11): 8.64 (SD = 3.11, range 6-15)
* 6y (n = 1): 7
* young (n = 70): 9.48 (SD = 3.00, range 6-17)
* old (n = 71): 9.13 (SD = 3.20, range 6-17)
* Fife (n = 81): 9.05 (SD = 3.08, range 6-17)
* Edinburgh (n = 60): 9.65 (SD = 3.12, range 6-17)

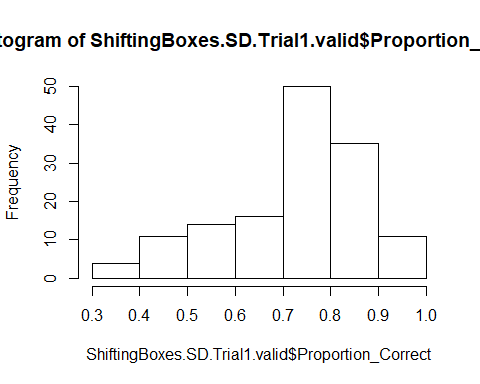
## Number of errors

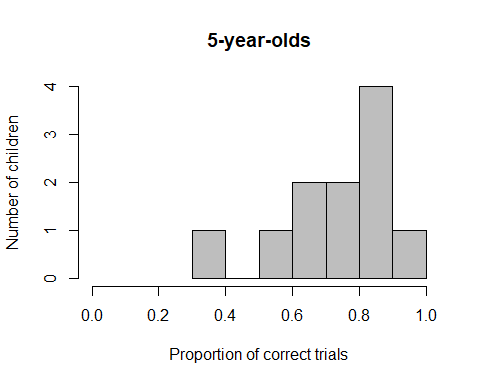
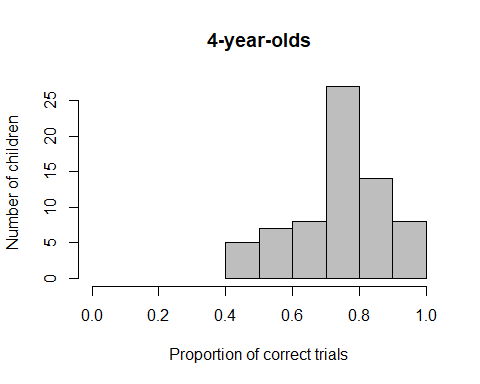
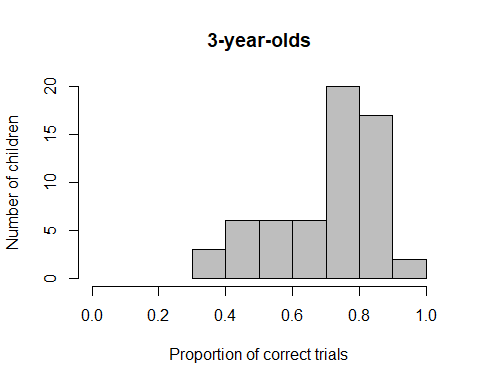


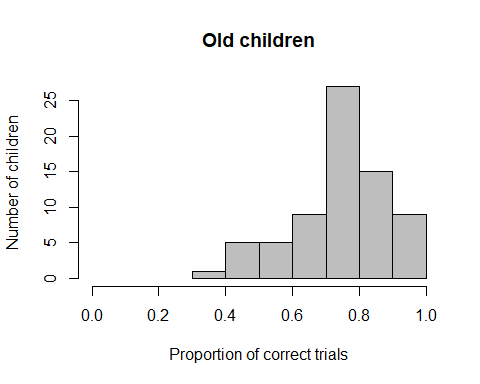
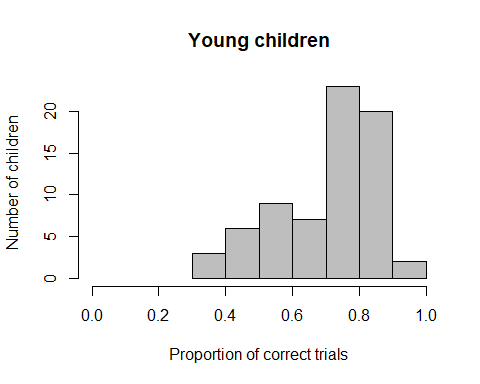
In the SD phase, the mean number of errors was 4.42 (SD = 4.23, range 0-16). The distribution of this variable is very left-skewed.

* 3y (n = 60): 4.85 (SD = 4.63, range 0-16)
* 4y (n = 69): 4.10 (SD = 3.79, range 0-14)
* 5y (n = 11): 4.36 (SD = 5.04, range 0-15)
* 6y (n = 1): 2
* young (n = 70): 4.91 (SD = 4.52, range 0-16)
* old (n = 71): 3.94 (SD = 3.93, range 0-15)
* Fife (n = 81): 4.91 (SD = 4.52, range 0-16)
* Edinburgh (n = 60): 3.94 (SD = 3.93, range 0-15)

## Proportion of correct trials







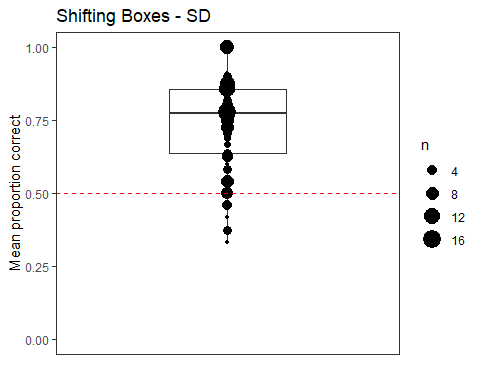
In the SD phase, the mean proportion of correct trials was **0.74 (SD = 0.15, range 0.33-1)**. The distribution of this variable is slightly right-skewed and **not normally distributed**, W = 0.957, p < .001. **Performance is significantly above chance**, V = 9047.5, p < .001.

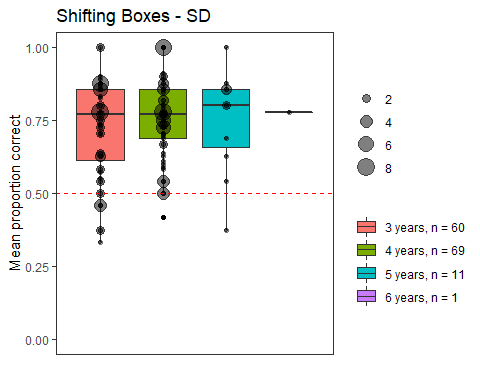
* 3y (n = 60): **0.72 (SD = 0.16, range 0.33-1)**, not normally distributed, W = 0.936, p = .003, performance above chance, V = 1657, p < .001
* 4y (n = 69): **0.75 (SD = 0.14, range 0.42-1)**, not normally distributed, W = 0.963, p = .041, performance above chance, V = 2139, p < .001
* 5y (n = 11): **0.75 (SD = 0.18, range 0.37-1)**, normally distributed, W = 0.912, p = .261, performance above chance, t(10) = 4.673, p = .001
* 6y (n =1): **0.78**
* young (n = 70): **0.72 (SD = 0.16, range 0.33-1)**, not normally distributed, W = 0.942, p = .003, performance above chance, V = 2281, p < .001
* old (n = 71): **0.76 (SD = 0.14, range 0.37-1)**, not normally distributed, W = 0.957, p = .016, performance above chance, V = 2264.5, p < .001

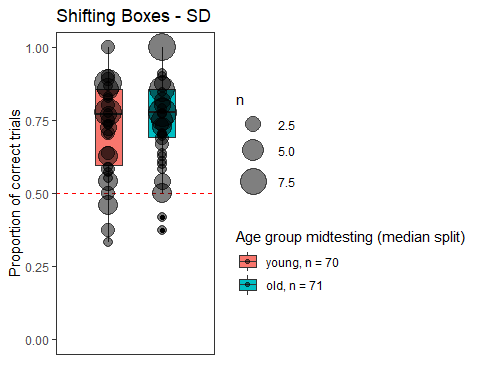
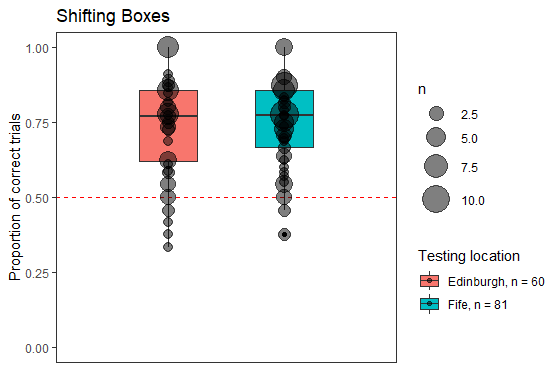
There is **no difference between young and old children** in their performance, one-sided Wilcoxon test, W = 2742, p = .145.

* Fife (n = 81): 0.74 (SD = 0.14, range 0.37-1)
* Edinburgh (n = 60): 0.73 (SD = 0.17, range 0.33-1)

There is **no difference between Fife and Edinburgh children**.







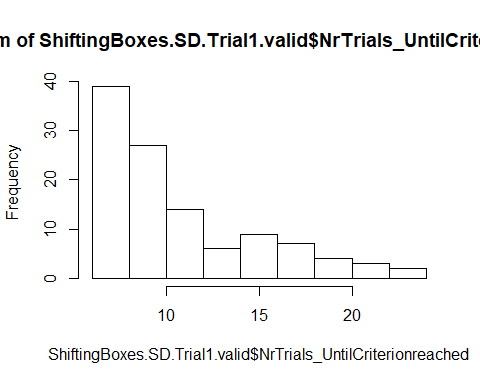
## Criterion reached

In the SD phase, out of the 141 children who had valid data in this phase, **111 children (79%) reached the criterion**, while 30 children did not reach the criterion and thus did not move on to the reversal phase.

* 3y: 45 out of 60 children reached criterion (75%)
* 4y: 57 out of 69 children reached criterion (83%)
* 5y: 8 out of 11 children reached criterion (73%)
* 6y: child reached criterion (100%)
* young: 52 out of 70 children reached criterion (74%)
* old: 59 out of 71 children reached criterion (83%)

Young and old children did not significantly differ in whether or not they reached criterion, X2(1) = 1.151, p = .283.

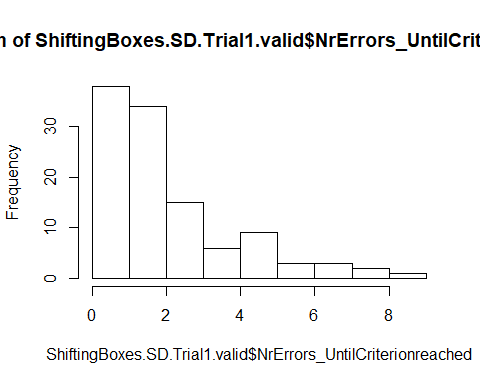
## Number of trials until criterion reached



In the SD phase, out of the 111 children who reached the criterion, children needed on average **10.95 trials (SD = 4.39, range 6-24)** to reach the criterion.

* 3y (n = 45): 10.60 (SD = 3.65, range 6-20)
* 4y (n = 57): 11.56 (SD = 4.99, range 6-24)
* 5y (n = 8): 8.75 (SD = 3.23, range 6-16)
* 6y (n =1): 9
* young (n = 52): 11.08 (SD = 3.99, range 6-22)
* old (n = 59): 10.85 (SD = 4.74, range 6-24)

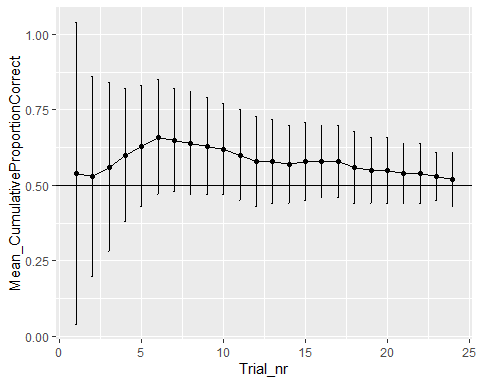
## Number of errors until criterion reached



In the SD phase, out of the 111 children who reached the criterion, children made on average 2.47 errors (SD = 1.92, range 0-9) until they reached the criterion.

* 3y (n = 45): 2.42 (SD = 1.79, range 0-9)
* 4y (n = 57): 2.63 (SD = 2.08, range 0-8)
* 5y (n = 8): 1.62 (SD = 1.50, range 0-5)
* 6y (n =1: 2
* young (n = 52): 2.52 (SD = 1.84, range 0-9)
* old (n = 59): 2.42 (SD = 2.01, range 0-8)

## Cumulative Proportion Correct



# Reversal

How many children started the task?

**111 children** continued to the Reversal task.

## Dropouts in Reversal

Out of the 111 children who started the Reversal Task, 9 children (5 f, 4 m, 8 3 years old, 1 4 years) dropped out in the Reversal phase. **1 child stopped after 1 trial**, **2 children after 5** trials, **1 after 7 trials, 1 after 9 trials, 1 after 10 trials, 1 after 14 trials, 1 after 15 trials,** and for **1 child E accidentally stopped the testing after 5 instead of 6 trials in a row correct (12 trials** in total).

We remove the dropouts from further analyses, as none of these children completed 75% of the maximum number of possible trials.

## Valid trials in Reversal

Thus, **102 children have valid data on the Reversal task** (with 1 child (ID 162) having completed 22 trials).

## Description of gender distribution

There were 53 girls and 49 boys.

## Description of age distribution

### Age at beginning of testing

The 102 children who had valid data on the Reversal phase were on average 49.80 months old (SD = 6.66, range 37-70) at the beginning of testing. There were 39 3-year-olds, 57 4-year-olds, and 6 5-year-olds.

* 3-year-olds: 24 f, 15 m
* 4-year-olds: 26 f, 31 m
* 5-year-olds: 3 f, 3 m
* young: 25 f, 19 m
* old: 28 f, 30 m

### Age in the middle of testing

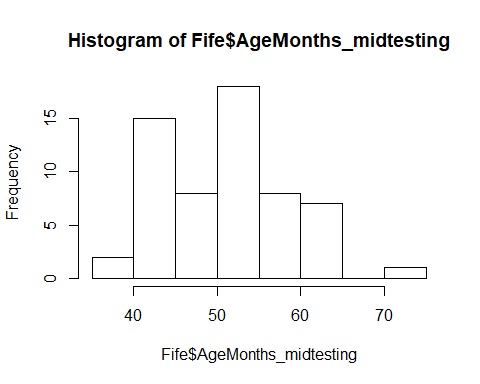
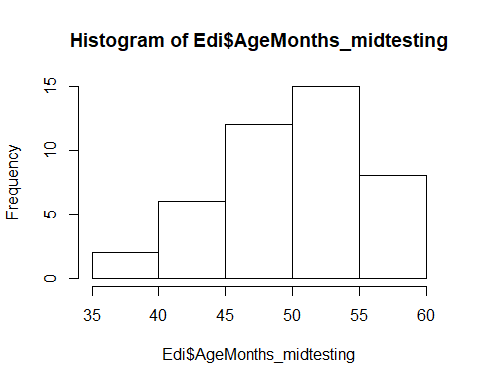
The 103 children who had valid data on the Reversal phase were on average **50.89 months old (SD = 6.59, range 38-72)** in the middle of testing. There were **38 3-year-olds, 55 4-year-olds, 8 5-year-olds, and 1 6-year-old**.

* 3-year-olds: 24 f, 14 m
* 4-year-olds: 24 f, 31 m
* 5-year-olds: 5 f, 3 m
* 6-year-old: 1 m

### Mediansplit age (based on entire sample)

There are **44 young** children (25 f, 19 m) and **58 old** children (28 f, 30 m).

## Testing location



59 children were from Fife, 43 children from Edinburgh.

Fife: M = 51.25 (SD = 7.48, range 39-72)

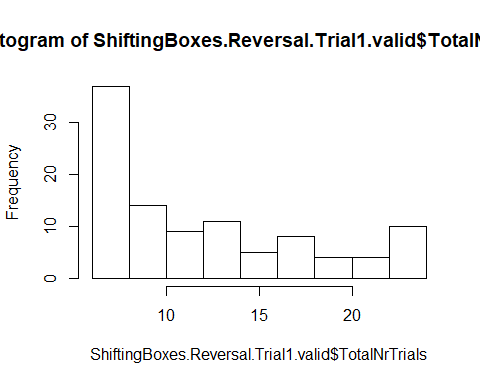
* 23 3y
* 27 4y
* 8 5y
* 1 6y

Edinburgh: M = 50.39 (SD = 5.16, range 38-58)

* 15 3y
* 28 4y

There is **no difference in the age distribution between the two testing locations**, two-sided t-test, t(99.741) = -0.686, p = .494.

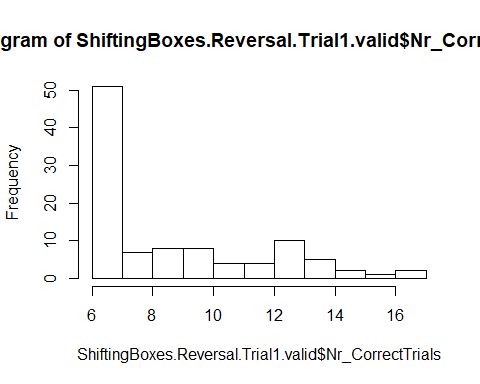
## Total number of trials



In the Reversal phase, the mean number of administered trials was **12.50 (SD = 5.71, range 6-24)**. The distribution of this variable is left-skewed, with most children solving the task relatively quickly. 50% of the children were administered 10.5 or fewer trials, 75% were administered 16.75 or fewer trials.

* 3y (n = 38): 13.79 (SD = 5.77, range 6-24)
* 4y (n = 55): 11.85 (SD = 5.56, range 6-24)
* 5y (n = 8): 11.50 (SD = 6.19, range 7-24)
* 6y (n = 1): 7
* young (n = 44): 13.79 (SD = 5.78, range 6-24)
* old (n = 58): 11.52 (SD = 5.51, range 6-24)
* Fife (n = 59): 13.22 (SD = 5.76, range 6-24)
* Edinburgh (n = 43): 11.51 (SD = 5.56, range 6-24)

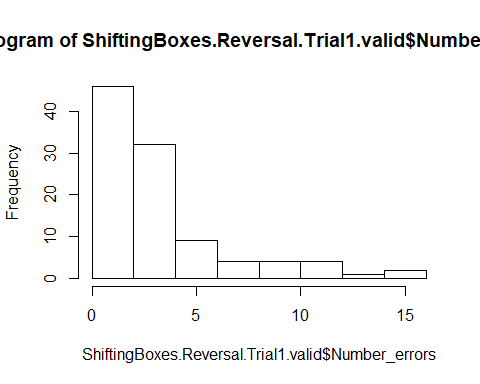
## Number of correct trials



In the Reversal phase, the mean number of correct trials was 8.80 (SD = 3.14, range 6-17). The distribution of this variable is very left-skewed.

* 3y (n = 38): 9.37 (SD = 3.36, range 6-17)
* 4y (n = 55): 8.64 (SD = 3.07, range 6-17)
* 5y (n = 8): 7.62 (SD = 2.20, range 6-11)
* 6y (n = 1): 6
* young (n = 44): 9.43 (SD = 3.34, range 6-17)
* old (n = 58): 8.33 (SD = 2.92, range 6-17)
* Fife (n = 59): 9.03 (SD = 3.26, range 6-17)
* Edinburgh (n = 43): 8.49 (SD = 2.98, range 6-15)

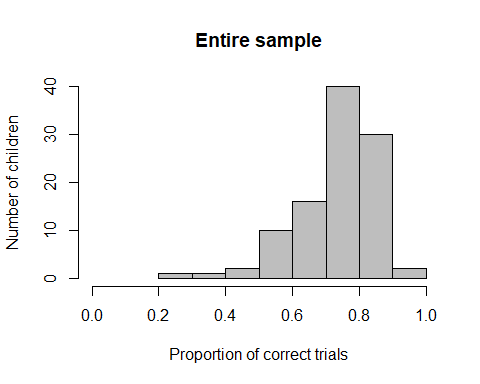
## Number of errors

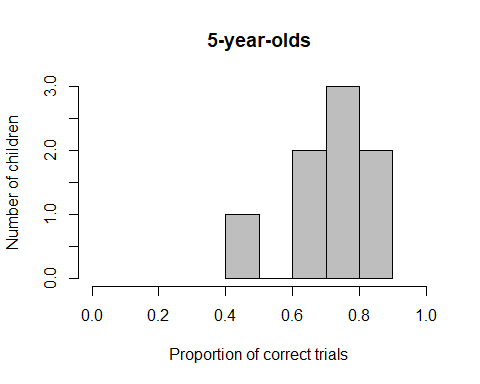
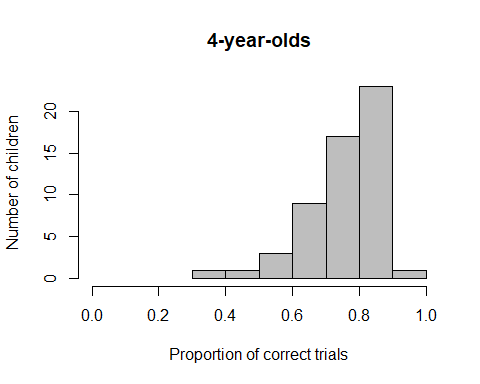
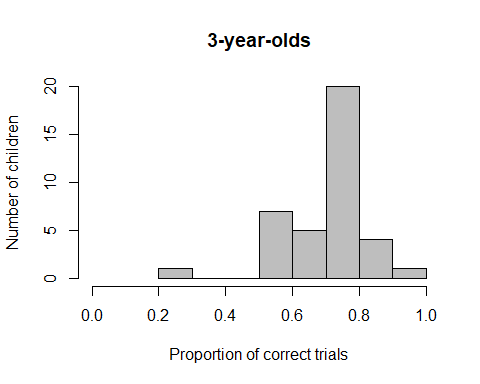


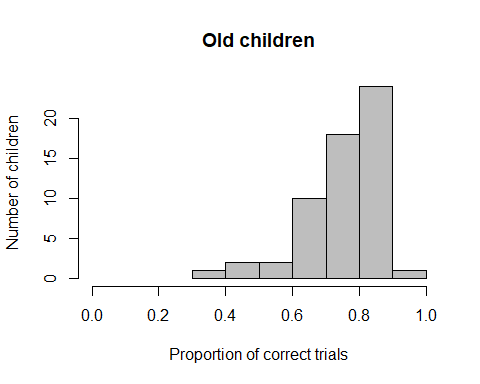
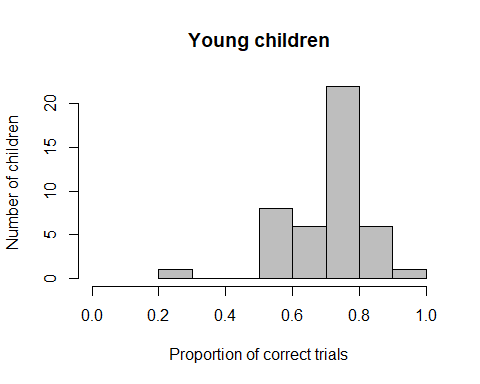
In the Reversal phase, the mean number of errors was 3.70 (SD = 3.25, range 0-15). The distribution of this variable is very left-skewed.

* 3y (n = 38): 4.42 (SD = 3.32, range 0-15)
* 4y (n = 55): 3.22 (SD = 3.03, range 0-15)
* 5y (n = 8): 3.87 (SD = 4.15, range 1-13)
* 6y (n = 1): 1
* young (n = 44): 4.36 (SD = 3.31, range 0-15)
* old (n = 58): 3.19 (SD = 3.13, range 0-15)
* Fife (n = 59): 4.19 (SD = 3.41, range 0-15)
* Edinburgh (n = 43): 3.02 (SD = 2.91, range 0-11)

## Proportion of correct trials







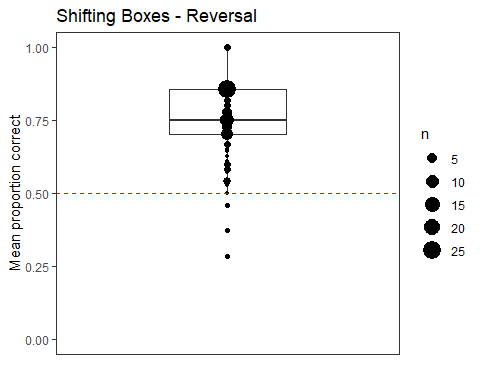
In the Reversal phase, the mean proportion of correct trials was **0.74 (SD = 0.12, range 0.29-1)**. The distribution of this variable is right-skewed and not normally distributed, W = 0.909, p < .001.

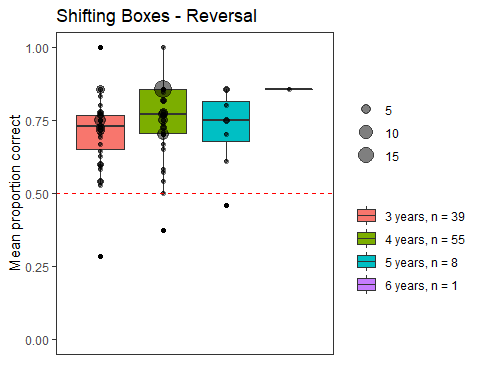
* 3y (n = 38): 0.71 (SD = 0.12, range 0.29-1), not normally distributed, W = 0.914, p = .007, performance significantly above chance, V = 725, p < .001
* 4y (n = 55): 0.77 (SD = 0.11, range 0.37-1), not normally distributed, W = 0.871, p < .001, performance significantly above chance, V = 1481, p < .001
* 5y (n = 8): 0.72 (SD = 0.13, range 0.46-.86), normally distributed, W = 0.896, p = .264, performance significantly above chance, t(7) = 4.687, p = .002
* 6y (n = 1): 0.86
* young (n = 44): 0.71 (SD = 0.12, range 0.29-1), not normally distributed, W = 0.919, p = .004, performance significantly above chance, V = 972, p < .001
* old (n = 58): 0.76 (SD = 0.11, range 0.37-1), not normally distributed, W = 0.863, p < .001, performance significantly above chance, V = 1647, p < .001

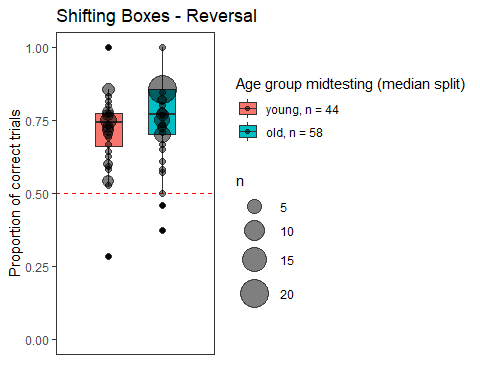
**Older children perform significantly better than younger** children, W = 1655.5, p = .005.

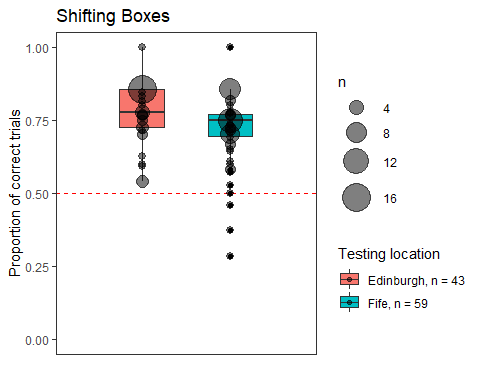
* Fife (n = 59): 0.72 (SD = 0.12, range 0.29-1), not normally distributed, W = 0.905, p < .001, performance significantly above chance, V = 1677, p < .001
* Edinburgh (n = 43): 0.78 (SD = 0.10, range 0.54-1), not normally distributed, W = 0.871, p < .001, performance significantly above chance, V = 946, p < .001

**Children from Edinburgh perform significantly better than children from Fife**, W = 1705.5, p = .003.







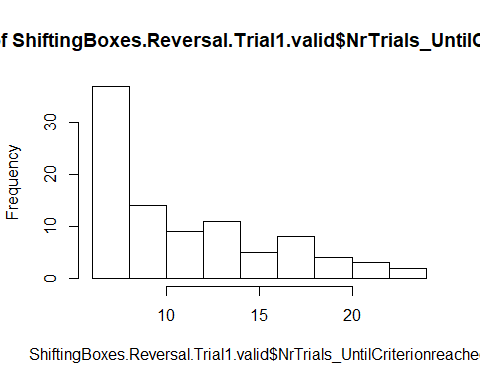


## Criterion reached

In the Reversal phase, out of the 102 children who had valid data in this phase, **92 children (90%) reached the criterion**, while 10 children did not reach the criterion.

* 3y: 33 out of 38 children reached criterion (87%)
* 4y: 51 out of 55 children reached criterion (93%)
* 5y: 7 out of 8 children reached criterion (87%)
* 6y: child reached criterion
* young: 38 out of 44 children reached criterion (86%)
* old: 54 out of 58 children reached criterion (93%)
* Edinburgh: 38 out of 43 children reached criterion (88%)
* Fife: 54 out of 59 children reached criterion (91%)

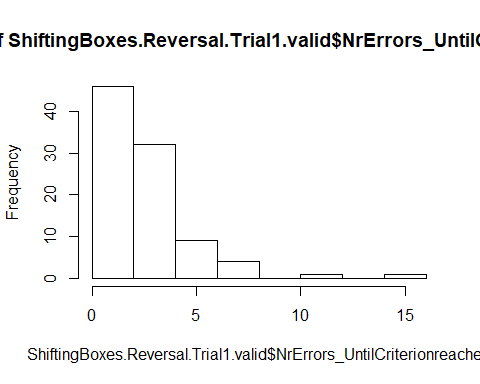
## Number of trials until criterion reached



In the Reversal phase, out of the 92 children who reached the criterion, children needed on average **11.41 trials (SD = 4.70, range 6-24) to reach the criterion**.

* 3y (n = 33): 12.65 (SD = 4.95, range 6-24)
* 4y (n = 51): 10.90 (SD = 4.54, range 6-24)
* 5y (n = 7): 9.71 (SD = 3.86, range 7-18)
* 6y (n = 1): 7
* young (n = 38): 12.54 (SD = 4.84, range 6-24)
* old (n = 54): 10.59 (SD = 4.47, range 6-24)
* Edinburgh (n = 38): 10.28 (SD = 4.17, range 6-24)
* Fife (n = 54): 12.22 (SD = 4.93, range 6-24)

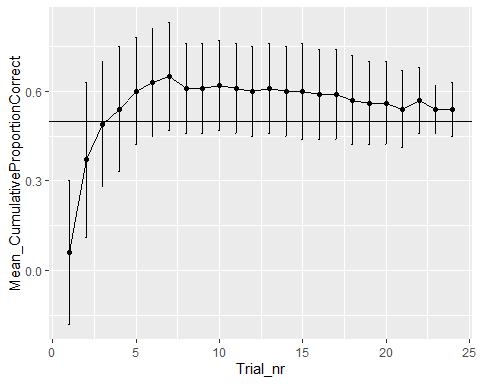
## Number of errors until criterion reached



In the Reversal phase, out of the 92 children who reached the criterion, children made on average 2.98 errors (SD = 2.31, range 0-15) until they reached the criterion.

* 3y (n = 33): 3.79 (SD = 2.90, range 0-15)
* 4y (n = 51): 2.53 (SD = 1.74, range 0-7)
* 5y (n = 7): 2.57 (SD = 2.07, range 1-7)
* 6y (n = 1): 1
* young (n = 38): 3.64 (SD = 2.76, range 0-15)
* old (n = 54): 2.50 (SD = 1.81, range 0-7)
* Edinburgh (n = 38): 2.31 (SD = 1.91, range 0-11)
* Fife (n = 54): 3.46 (SD = 2.47, range 0-15)

## Cumulative Proportion Correct



# CD: All children who started that phase, independent of whether or not they reached criterion in SD (this will not be used for the actual analysis, is here jst for sake of completion)

## How many children started the task?

139 children started the CD task.

## Dropouts in CD

Out of the 139 children who started the CD Task, **7 children** (5 f, 2 m, 5 3 years old, 2 4 years) **dropped out** in the CD phase. **1 child stopped after 1 trial, 2 children after 7 trials, 1 after 9 trials, 1 after 14 trials, 1 after 16 trials, and for 1 child** (ID 24) E accidentally put a sticker in the wrong box on trial 13, so there were only **12 valid trials**.

We decided to keep all those children who had completed at least 18 trials, i.e., we included IDs 34 and 67. ID34 stopped the game after 18 trials and ID 67 stopped the game after 19 trials.

There was also 1 case that was a bit unclear. Here (ID 38), E accidentally switched to EDS after only 5 consecutive trials correct instead of 6. As the child completed 22 trials and had done quite well, we decided to include the child, but we do not know whether or not the child reached criterion, so the child had to be dropped from this analysis.

We removed the remaining 7 clear dropout children from further analyses.

## Valid data in CD

Thus, **132 children had valid data on the CD task (with 1 child having completed 18 trials, 1 19 trials).**

### Description of gender distribution in CD

There were 70 girls and 62 boys.

### Description of age distribution

#### Age at beginning of testing

The 132 children who had valid data on the CD phase were on average 48.94 months old (SD = 6.99, range 36-70) at the beginning of testing. There were 59 3-year-olds, 65 4-year-olds, and 8 5-year-olds.

* 3-year-olds: 33 f, 26 m
* 4-year-olds: 33 f, 32 m
* 5-year-olds: 4 f, 4 m

#### Age in the middle of testing

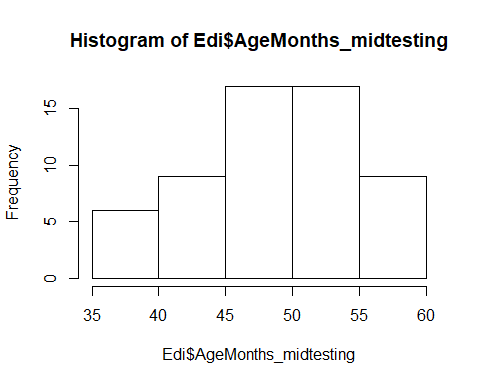
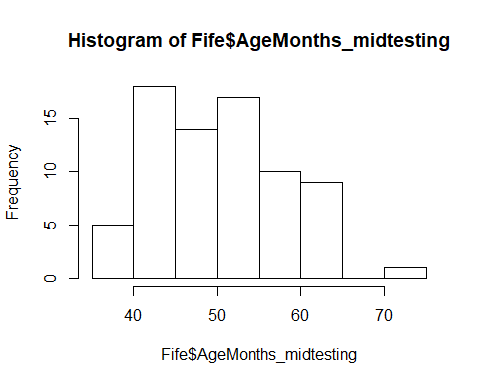
The 132 children who had valid data on the CD phase were on average 50.04 months old (SD = 6.95, range 36-72) in the middle of testing. There were 56 3-year-olds, 64 4-year-olds, 11 5-year-olds, and 1 6-year-old.

* 3-year-olds: 32 f, 24 m
* 4-year-olds: 31 f, 33 m
* 5-year-olds: 7 f, 4 m
* 6-year-old: 1 m

#### Age at mediansplit

There were 66 young children (35 f, 31 m) and 66 old children (35 f, 31 m).

### Testing location

74 children were from Fife, 58 children from Edinburgh.

Fife: M = 50.73 (SD = 7.70, range 39-72)

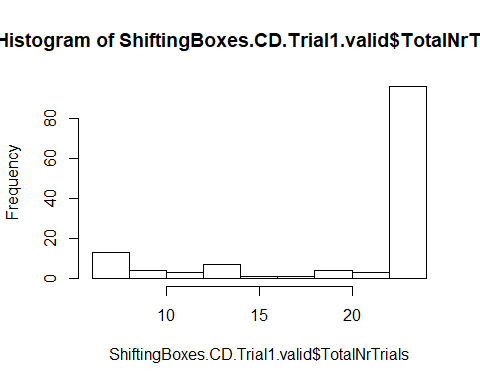
* 32 3y
* 30 4y
* 11 5y
* 1 6y

Edinburgh: M = 49.17 (SD = 5.80, range 36-58)

* 24 3y
* 34 4y

There is no difference in the age distribution between the two testing locations, two-sided Wilcoxon test, W = 1984.5, p = .460.

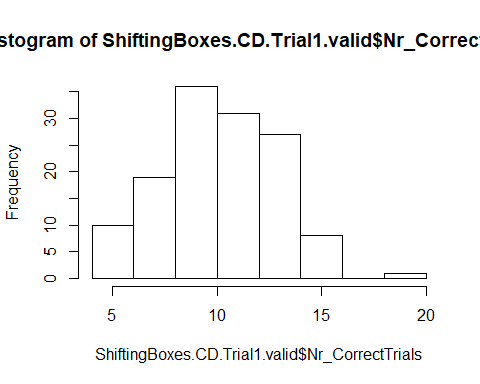
## Total number of trials



In the CD phase, the mean number of administered trials was 20.70 (SD = 6.01, range 6-24). The distribution of this variable is very right-skewed. 50% of the children were administered 24 of fewer trials.

* 3y (n = 56): 21.68 (SD = 5.06, range 6-24)
* 4y (n = 64): 20.72 (SD = 6.16, range 6-24)
* 5y (n = 11): 15.36 (SD = 7.38, range 6-24)
* 6y (n = 1): 24
* young (n = 66): 21.32 (SD = 5.37, range 6-24)
* old (n = 66): 20.09 (SD = 6.56, range 6-24)
* Edinburgh (n = 58): 20.64 (SD = 6.15, range 6-24)
* Fife (n = 74): 20.76 (SD = 5.93, range 6-24)

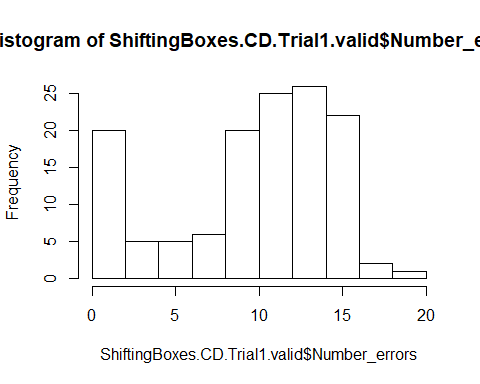
## Number of correct trials



In the CD phase, the mean number of correct trials was 10.61 (SD = 2.71, range 5-19). The distribution of this variable is slightly left-skewed.

* 3y (n = 56): 10.84 (SD = 2.54, range 6-15)
* 4y (n = 64): 10.65 (SD = 2.77, range 5-19)
* 5y (n = 11): 9.36 (SD = 3.23, range 6-14)
* 6y (n = 1): 9
* young (n = 66): 10.74 (SD = 2.54, range 6-15)
* old (n = 66): 10.48 (SD = 2.89, range 5-19)
* Edinburgh (n = 58): 10.79 (SD = 2.55, range 6-15)
* Fife (n = 74): 10.47 (SD = 2.85, range 5-19)

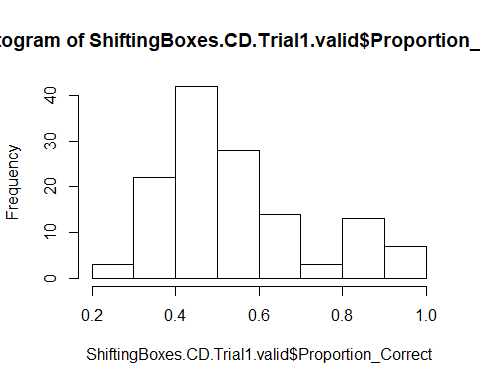
## Number of errors

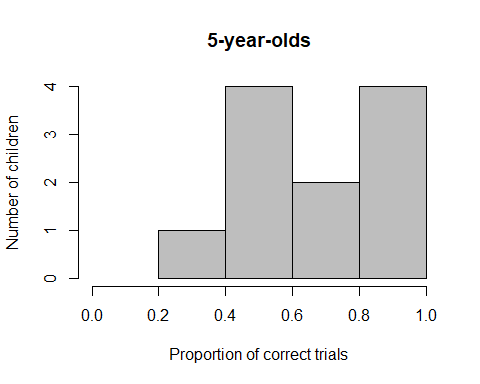
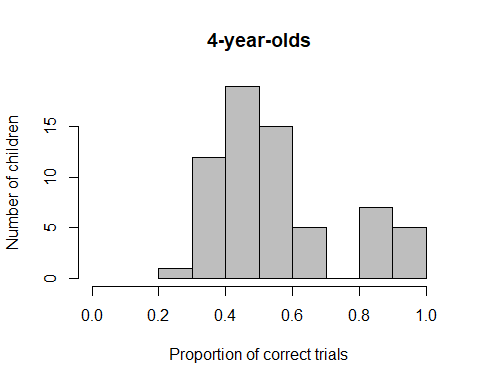
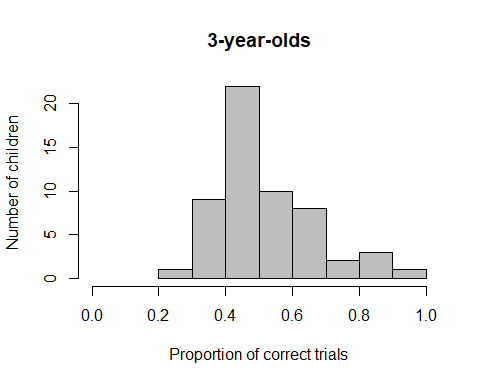


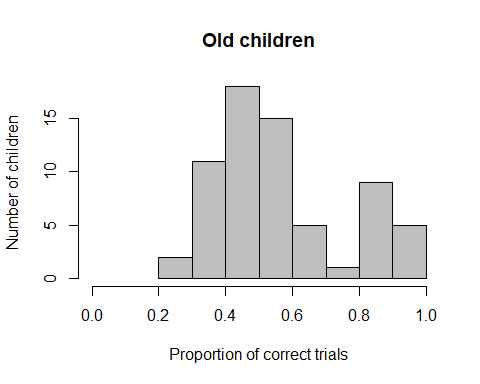
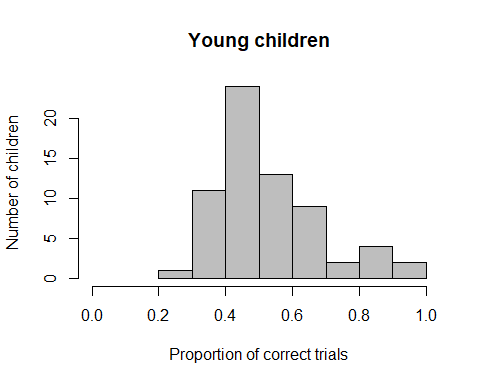
In the CD phase, the mean number of errors was 10.09 (SD = 4.94, range 0-19). The distribution of this variable is bimodal, with some children having very few errors and others quite many.

* 3y (n = 56): 10.84 (SD = 4.28, range 0-17)
* 4y (n = 64): 10.06 (SD = 5.12, range 0-19)
* 5y (n =11): 6.00 (SD = 5.46, range 0-17)
* 6y (n = 1): 15
* young (n = 66): 10.57 (SD = 4.46, range 0-17)
* old (n = 66): 9.61 (SD = 5.37, range 0-19)
* Edinburgh (n = 58): 9.84 (SD = 4.88, range 0-16)
* Fife (n = 74): 10.28 (SD = 5.00, range 0-19)

## Proportion of correct trials







In the CD phase, the mean proportion of correct trials was **0.56 (SD = 0.18, range 0.21-1)**. The distribution of this variable has 1 large peak on the left and a smaller one on the right and is not normally distributed, W = 0.908, p < .001. Performance is significantly **above chance**, V = .438, p = .020.

* 3y (n = 56): **0.53 (SD = 0.15, range 0.29-1)**, not normally distributed, W = 0.926, p = .002, performance **at chance** level, V = 705, p = .359
* 4y (n = 64): **0.56 (SD = 0.20, range 0.21-1)**, not normally distributed, W = 0.882, p < .001, performance **at chance** level, V = 1027, p = .185
* 5y (n =11): **0.68 (SD = 0.21, range 0.29-1)**, normally distributed, W = 0.951, p = .655, performance significantly **above chance**, t(10) = 2.923, p = .015
* 6y (n = 1): 0.37
* young (n = 66): **0.53 (SD = 0.16, range 0.29-1)**, not normally distributed, W = 0.917, p < .001, performance **at chance** level, V = 1050, p = .214
* old (n = 66): **0.58 (SD = 0.20, range 0.21-1)**, not normally distributed, W = 0.904, p < .001, performance significantly **above chance**, V = 1147, p = .048

**Younger and older children do not differ** in their performance, W = 2389.5, p = .168.

* Edinburgh (n = 58): 0.57 (SD = 0.18, range 0.33-1), not normally distributed, W = 0.891, p < .001, performance significantly above chance, V = 853.5, p = .037
* Fife (n = 74): 0.55 (SD = 0.19, range 0.21-1), not normally distributed, W = 0.912, p < .001, performance at chance level, V = 1377, p = .213

Children in Fife and Edinburgh do not differ in their performance in CD, W = 2326.5, p = .407

## Criterion reached

In the CD phase, out of the 132 children who had valid data in this phase, **36 children (27%) reached the criterion**, while 95 children did not reach the criterion. For the 1 child mentioned above (ID 38), it was unclear whether or not they reached the criterion (E stopped too early, but we asumed the child would have passed criterion).

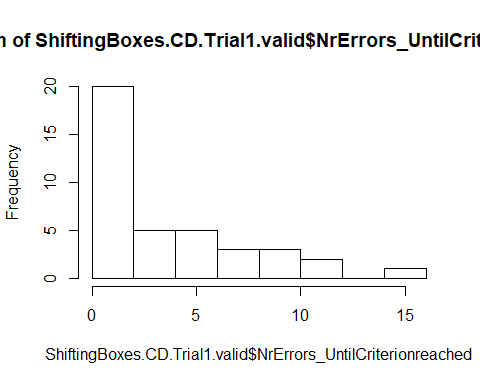
* 3y: 11 out of 56 children reached criterion (20%), 1 child unclear
* 4y: 18 out of 64 children reached criterion (28%)
* 5y: 7 out of 11 children reached criterion (63%)
* 6y: child did not reach criterion
* young: 14 out of 66 children reached criterion (21%), 1 child unclear
* old: 22 out of 66 children reached criterion (33%)
* Edinburgh: 17 out of 58 children reached criterion (29%), i child unclear
* Fife: 19 out of 64 children reached criterion (30%)

## Number of trials until criterion reached

In the CD phase, out of the 36 children who reached the criterion, children needed on average 13.31 trials (SD = 6.51, range 6-24) to reach the criterion.

* 3y (n = 11): 14.61 (SD = 6.61, range 6-24)
* 4y (n = 18): 13.47 (SD = 7.18, range 6-24)
* 5y (n = 7): 10.43 (SD = 3.55, range 6-14)
* young (n = 14): 13.75 (SD = 6.34, range 6-24)
* old (n = 22): 13.00 (SD = 6.75, range 6-24)
* Fife (n = 19): 12.80 (SD = 6.29, range 6-24)
* Edinburgh (n = 17): 13.84 (SD = 6.86, range 6-24)

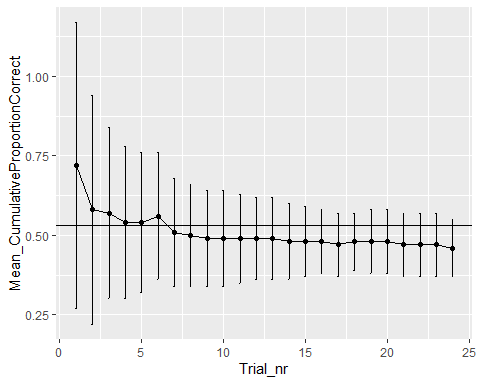
## Number of errors until criterion reached



In the CD phase, out of the 36 children who reached the criterion, children made on average 3.90 errors (SD = 3.90, range 0-15) until they reached the criterion.

* 3y (n = 11): 4.92 (SD = 4.21, range 0-15)
* 4y (n = 18): 3.68 (SD = 4.08, range 0-11)
* 5y (n = 7): 2.57 (SD = 2.51, range 0-7)
* young (n = 14): 4.37 (SD = 4.01, range 0-15)
* old (n = 22): 3.56 (SD = 3.87, range 0-11)
* Fife (n = 19): 3.35 (SD = 3.18, range 0-11)
* Edinburgh (n = 17): 4.47 (SD = 4.55, range 0-15)

## Cumulative Proportion Correct



# CD: Only those children who reached criterion in SD

## How many children reached criterion in SD and started the CD task?

**106 children reached criterion in SD and started the CD task**. 111 children reached criterion in SD, but 5 of them could not be tested on the second phase.

## Dropouts in CD

Out of the 106 children who started the CD Task, **5 children** (3 f, 2 m, 4 3 years old, 1 4 years) dropped out in the CD phase. **1 child (3y, f) stopped after 1 trial**, **1 child (m, 3y) stopped after 7 trials**, for **1 child** (f, 4 y) there was an experimenter error (sticker in wrong box) and there were only **12 valid trials**, **1 child** (m, 3y) stopped after **9 trials**, **1 child** (m, 4y) stopped after **16 trials**.

We decided to keep all those children who had completed at least 18 trials, i.e., we included ID 67. ID 67 stopped the game after 19 trials.

There was also 1 case that was a bit unclear. Here (ID 38), E accidentally switched to EDS after only 5 consecutive trials correct instead of 6. As the child completed 22 trials and had done quite well, we decided to include the child, but we do not know whether or not the child reached criterion, so the child had to be dropped from this analysis.

We remove all 5 dropout children from further analyses, as none of these children completed 75% of the maximum number of possible trials.

## Valid data on CD

Thus, **101 children have valid data on the CD task** (with 1 child having completed 19 trials).

## Description of gender distribution

There were 52 girls and 49 boys.

## Description of age distribution

### Age at beginning of testing

The 101 children who had valid data on the CD phase were on average 49.72 months old (SD = 6.65, range 37-70) at the beginning of testing. There were 41 3-year-olds, 54 4-year-olds, and 6 5-year-olds.

* 3-year-olds: 25 f, 16 m
* 4-year-olds: 24 f, 30 m
* 5-year-olds: 3 f, 3 m

### Age in the middle of testing

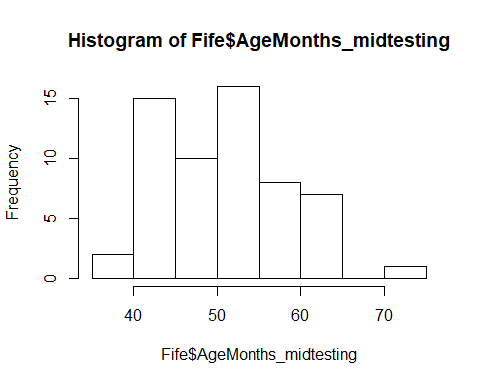
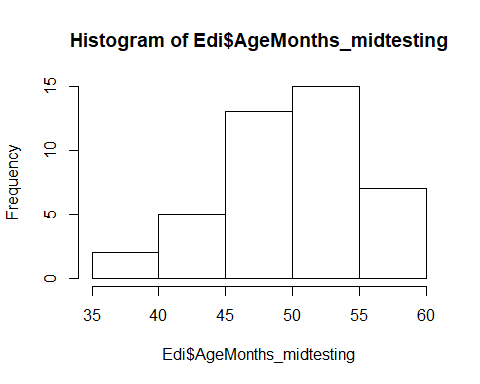
The 101 children who had valid data on the CD phase were on average **50.80 months old (SD = 6.60, range 38-72)** in the middle of testing. There were **39 3-year-olds, 53 4-year-olds, 8 5-year-olds, and 1 6-year-old**.

* 3-year-olds: 25 f, 14 m
* 4-year-olds: 22 f, 31 m
* 5-year-olds: 5 f, 3 m
* 6-year-old: 1 m

### Mediansplit age

There were 46 young children (26 f, 20 m) and 55 old children (26 f, 29 m).

## Testing location



42 children were from Edinburgh, 59 children from Fife

Fife: M = 51.15 (SD = 7.51, range 39-72)

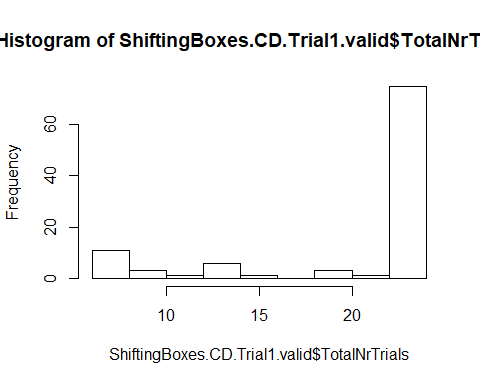
* 24 3y
* 26 4y
* 8 5y
* 1 6y

Edinburgh: M = 50.31 (SD = 5.11, range 38-58)

* 15 3y
* 27 4y

There is no difference in the age distribution between the two testing locations, two-sided t-test, t(98.83) = -0.671, p = .504.

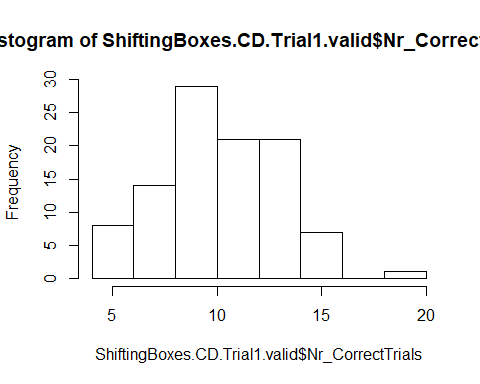
## Total number of trials



In the CD phase, the mean number of administered trials was **20.69 (SD = 6.11, range 6-24)**. The distribution of this variable is very right-skewed. 50% of the children were administered 24 of fewer trials.

* 3y (n = 39): 22.02 (SD = 4.83, range 6-24)
* 4y (n = 53): 20.49 (SD = 6.33, range 6-24)
* 5y (n = 8): 15.12 (SD = 7.92, range 6-24)
* 6y (n = 1): 24
* young (n = 46): 21.69 (SD = 4.96, range 6-24)
* old (n = 55): 19.85 (SD = 6.86, range 6-24)
* Edinburgh (n = 42): 20.62 (SD = 6.13, range 6-24)
* Fife (n = 59): 20.74 (SD = 6.15, range 6-24)

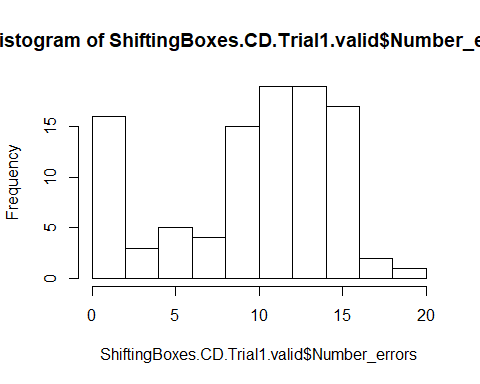
## Number of correct trials



In the CD phase, the mean number of correct trials was 10.63 (SD = 2.82, range 5-19). The distribution of this variable is slightly left-skewed.

* 3y (n = 39): 11.02 (SD = 2.60, range 6-15)
* 4y (n = 53): 10.68 (SD = 2.87, range 5-19)
* 5y (n = 11): 8.62 (SD = 3.16, range 6-14)
* 6y (n = 1): 9
* young (n = 46): 11.04 (SD = 2.55, range 6-15)
* old (n = 55): 10.29 (SD = 3.02, range 5-19)
* Edinburgh (n = 42): 10.86 (SD = 2.53, range 6-15)
* Fife (n = 59): 10.47 (SD = 3.02, range 5-19)

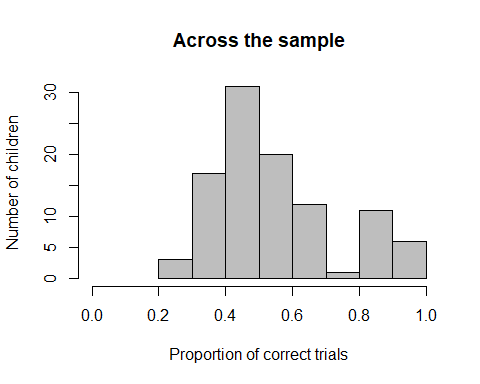
## Number of errors

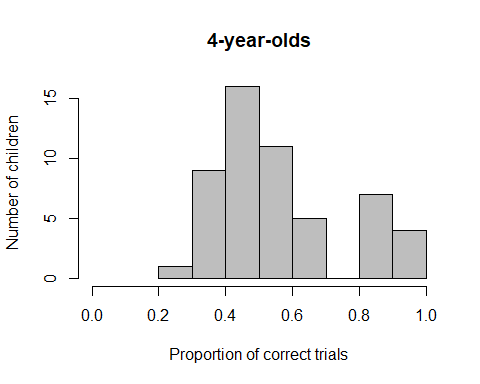
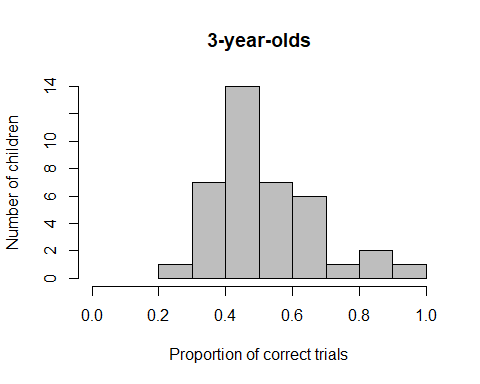


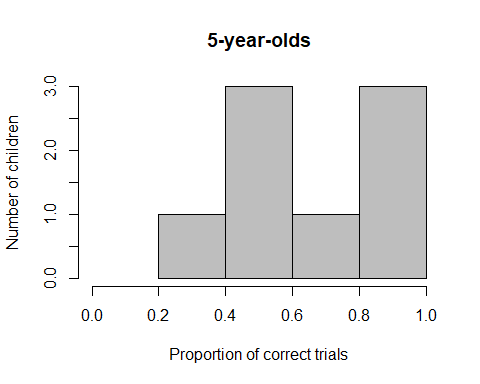
In the CD phase, the mean number of errors was 10.06 (SD = 5.05, range 0-19). The distribution of this variable is bimodal, with some children having very few errors and others quite many.

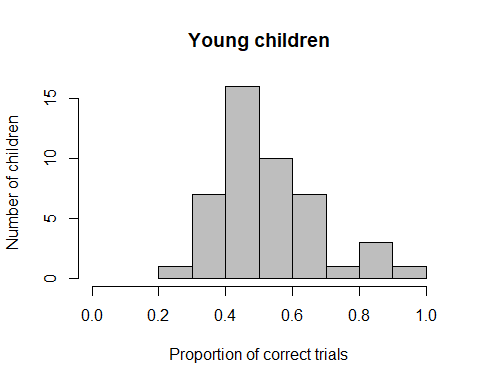
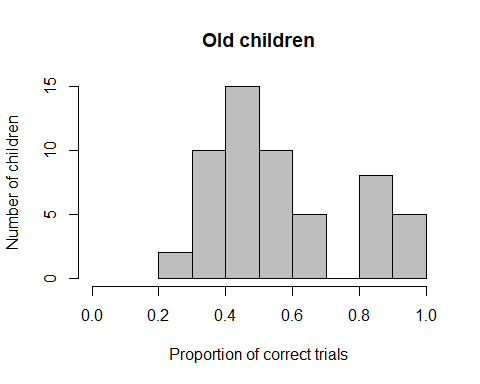
* 3y (n = 39): 11.00 (SD = 4.32, range 0-17)
* 4y (n = 53): 9.81 (SD = 5.25, range 0-19)
* 5y (n = 11): 6.50 (SD = 5.95, range 0-17)
* 6y (n = 1): 15
* young (n = 46): 10.65 (SD = 4.33, range 0-17)
* old (n = 55): 9.56 (SD = 5.57, range 0-19)
* Edinburgh (n = 42): 9.76 (SD = 4.94, range 0-16)
* Fife (n = 59): 10.27 (SD = 5.16, range 0-19)

## Proportion of correct trials







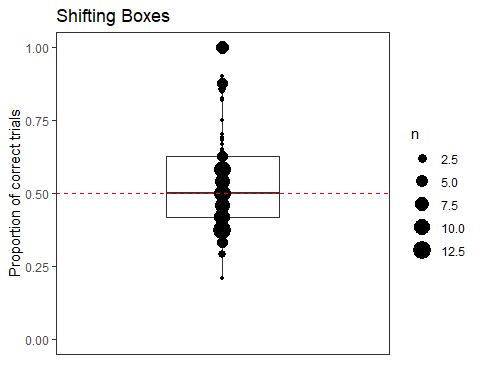
In the CD phase, the mean proportion of correct trials was **0.56 (SD = 0.19, range 0.21-1)**. The distribution of this variable has 1 large peak on the left and a smaller one on the right and is **not normally distributed, W = 0.911, p < .001**. Performance is **above chance** level, V = 5151, p < .001.

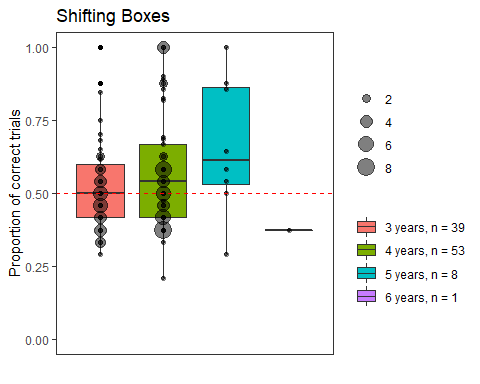
* 3y (n = 39): **0.53 (SD = 0.16, range 0.29-1)**, not normally distributed, W = 0.927, p = .014, performance **at chance** level, V = 332, p = .560
* 4y (n = 53): **0.57 (SD = 0.20, range 0.21-1)**, not normally distributed, W = 0.901, p < .001, performance **at chance** level, V = 754, p = .089
* 5y (n =11): **0.66 (SD = 0.23, range 0.29-1)**, normally distributed, W = 0.960, p = .814, performance **at chance** level, t(7) = 1.956, p = .091
* 6y (n = 1): 0.37
* young (n = 46): **0.53 (SD = 0.15, range 0.29-1)**, not normally distributed, W = 0.940, p = .020, performance **at chance** level, V = 514, p = .281
* old (n = 55): **0.58 (SD = 0.21, range 0.21-1)**, not normally distributed, W = 0.900, p < .001, performance **at chance** level, V = 802, p = .060

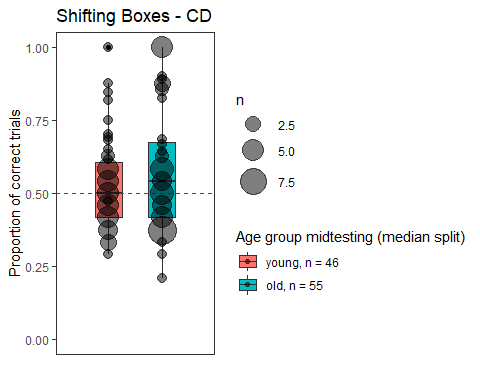
**Younger and older children do not differ** in their performance, W = 1363.5, p = .251.

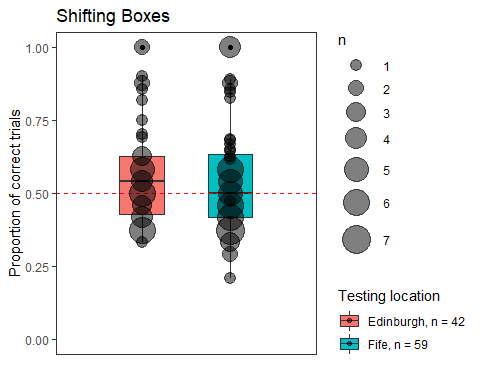
* Edinburgh (n = 42): 0.57 (SD = 0.18, range 0.33-1), not normally distributed, W = 0.899, p = .001, performance at chance level, V = 450, p = .066
* Fife (n = 59): 0.55 (SD = 0.20, range 0.21-1), not normally distributed, W = 0.912, p < .001, performance at chance level, V = 881.5, p = .233

**Children in Fife and Edinburgh do not differ** in their performance in CD, W = 1344.5, p = .468

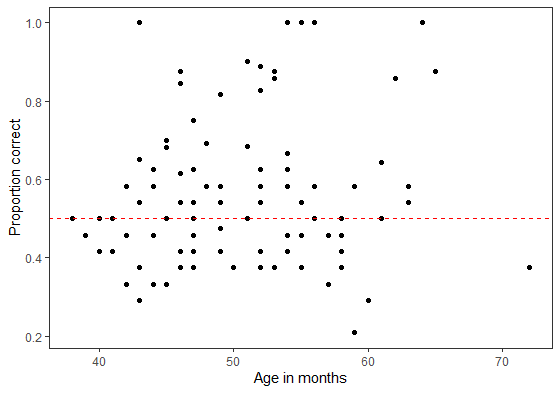








## Plot age as continuous variable against proportion correct

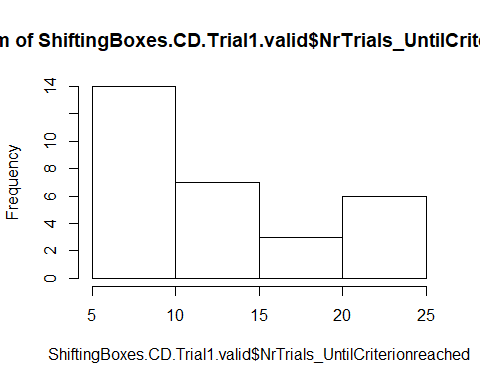


## Criterion reached

In the CD phase, out of the 101 children who had valid data in this phase, 27 children (27%) reached the criterion, while 74 children did not reach the criterion.

* 3y: 7 out of 39 children reached criterion (18%)
* 4y: 15 out of 53 children reached criterion (28%)
* 5y: 5 out of 11 children reached criterion (45%)
* 6y: child did not reach criterion
* young: 9 out of 46 children reached criterion (19%)
* old: 18 out of 55 children reached criterion (33%)
* Edinburgh: 11 out of 42 children reached criterion (26%)
* Fife: 16 out of 59 children reached criterion (27%)

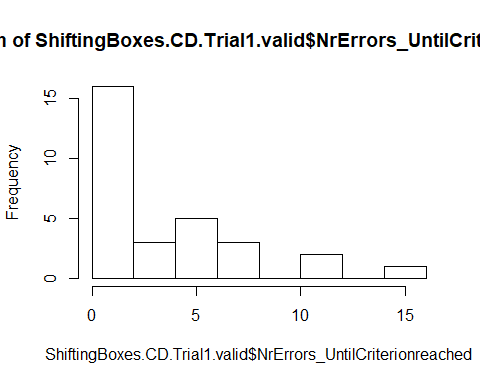
## Number of trials until criterion reached



In the CD phase, out of the 27 children who reached the criterion, children needed on average 13.03 trials (SD = 6.46, range 6-24) to reach the criterion.

* 3y (n = 7): 15.44 (SD = 6.89, range 6-24)
* 4y (n = 15): 12.69 (SD = 6.64, range 6-24)
* 5y (n = 5): 9.80 (SD = 3.90, range 6-14)
* young (n = 9): 14.82 (SD = 6.34, range 6-24)
* old (n = 18): 12.00 (SD = 6.46, range 6-24)
* Edinburgh (n = 11): 13.08 (SD = 6.24, range 6-24)
* Fife (n = 16): 13.00 (SD = 6.79, range 6-24)

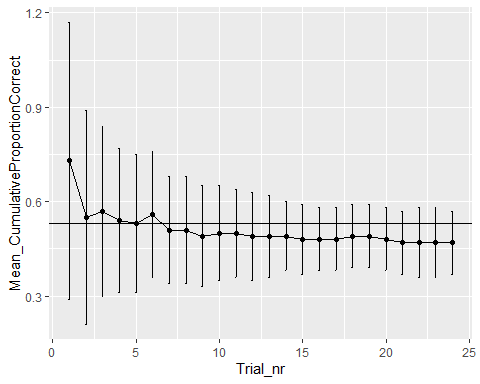
## Number of errors until criterion reached



In the CD phase, out of the 27 children who reached the criterion, children made on average 3.67 errors (SD = 3.86, range 0-15) until they reached the criterion.

* 3y (n = 7): 5.11 (SD = 4.57, range 0-15)
* 4y (n = 15): 3.12 (SD = 3.67, range 0-11)
* 5y (n = 5): 2.80 (SD = 3.03, range 0-7)
* young (n = 9): 4.73 (SD = 4.20, range 0-15)
* old (n = 18): 3.05 (SD = 3.63, range 0-11)
* Edinburgh (n = 11): 3.92 (SD = 4.50, range 0-15)
* Fife (n = 16): 3.47 (SD = 3.43, range 0-11)

## Cumulative Proportion Correct



# EDS: All children who reached criterion in CD, regardless of performance in SD

**36 children reached criterion in CD and started the EDS task.** There was **an additional child (ID 38) where E stopped after 5 in a row instead of 6 in a row**. We included this child in the analysis for CD, and also included it in EDS, but here did not label them as having reached the criterion, because we just cannot be sure.

## Dropouts in EDS

1 child was a dropout (4y, m), stopped after 24 trials. We decided to remove this child, as they did not reach 75% of the trials (i.e., 27 trials). There is one child (ID 38) where E accidentally switched to EDS after only 5 instead of 6 consecutive trials correct in CD. For now, we decided to keep this child. There is also 1 child who stopped the game after 33 trials, and this child is also included in the dataset. There is also 1 child where E accidentally only administered 35 trials, this child is also included. There is also 1 child who has 35 trials as E accidentally skipped one trial.

## Valid data

There are 35 valid datapoints in EDS.

## Description of gender distribution in EDS

There were 19 girls and 16 boys.

## Description of age distribution in EDS

### Age at beginning of testing

The 35 children who had valid data on the EDS phase were on average 50.77 months old (SD = 6.70, range 36-64) at the beginning of testing. There were 11 3-year-olds, 20 4-year-olds, and 4 5-year-olds.

* 3-year-olds: 6 f, 5 m
* 4-year-olds: 12 f, 8 m
* 5-year-olds: 1 f, 3 m

### Age in the middle of testing

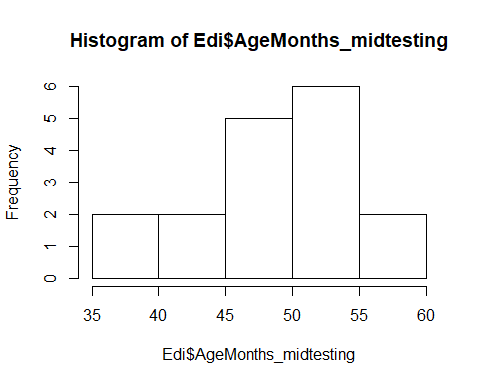
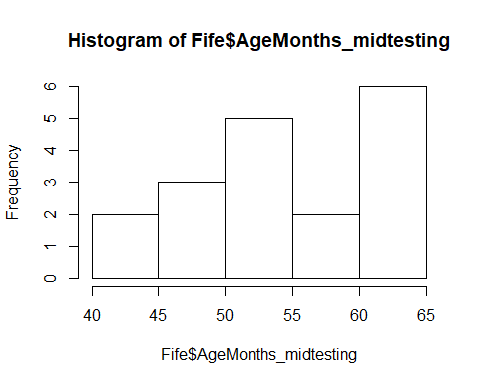
The 35 children who had valid data on the EDS phase were on average 51.77 months old (SD = 7.12, range 36-65) in the middle of testing. There were 11 3-year-olds, 17 4-year-olds, and 7 5-year-olds.

* 3-year-olds: 6 f, 5 m
* 4-year-olds: 9 f, 8 m
* 5-year-olds: 4 f, 3 m

### Mediansplit age

There were 14 young children (7 f, 7 m) and 21 old children (12 f, 9 m).

## Testing location

17 children were from Edinburgh, 18 children from Fife

Fife: M = 54.33 (SD = 7.42, range 43-65)

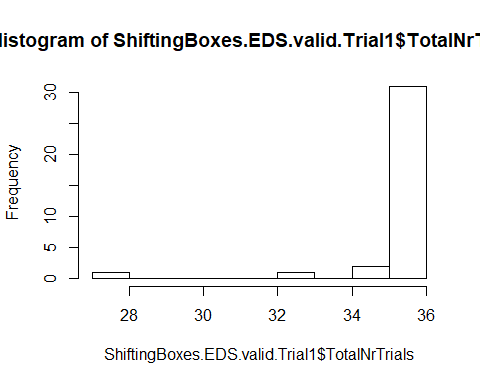
* 5 3y
* 6 4y
* 7 5y

Edinburgh: M = 49.06 (SD = 5.82, range 36-58)

* 6 3y
* 11 4y

**Fife children are significantly older than Edinburgh children**, two-sided t-test, t(31.96) = -2.35, p = .025.

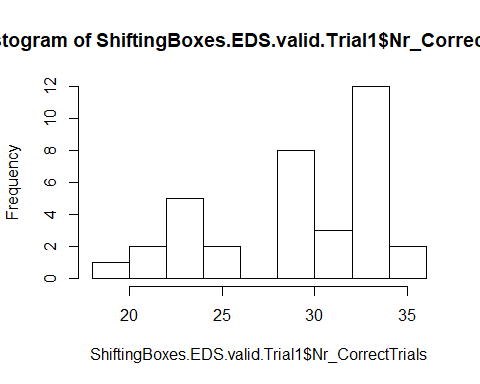
## Total number of trials



In the EDS phase, the mean number of administered trials was 35.60 (SD = 1.59, range 27-36).

* 3y (n = 11): 35.64 (SD = 0.92, range 33-36)
* 4y (n = 17): 35.41 (SD = 2.18, range 27-36)
* 5y (n = 7): 36
* young (n = 14): 35.71 (SD = 0.82, range 33-36)
* old (n = 21): 35.52 (SD = 1.96, range 27-36)
* Edinburgh (n = 17): 35.35 (SD = 2.18, range 27-36)
* Fife (n = 18): 35.83 (SD = 0.71, range 33-36)

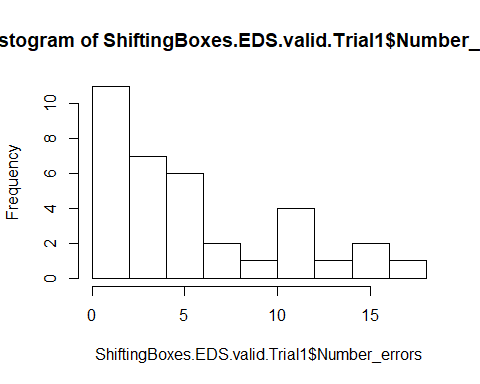
## Number of correct trials



In the EDS phase, the mean number of correct trials was 29.63 (SD = 4.84, range 18-36). The distribution of this variable has 2 peaks, one around 34, the other around 23.

* 3y (n = 11): 27.64 (SD = 5.50, range 18-35)
* 4y (n = 17): 29.88 (SD = 4.38, range 21-34)
* 5y (n = 7): 32.14 (SD = 4.02, range 24-36)
* young (n = 14): 27.71 (SD = 4.94, range 18-35)
* old (n = 21): 30.90 (SD = 4.44, range 21-36)
* Edinburgh (n = 17): 27.76 (SD = 5.42, range 18-35)
* Fife (n = 18): 31.39 (SD = 3.52, range 23-36)

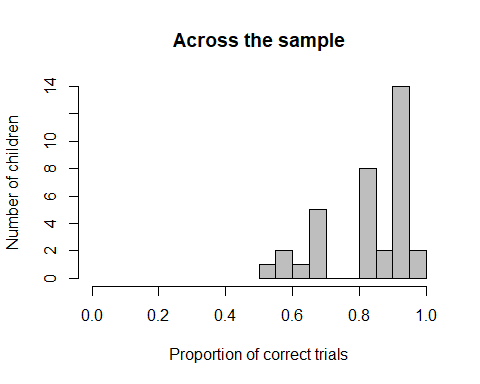
## Number of errors

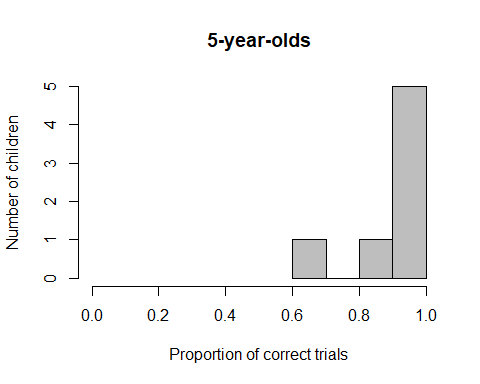
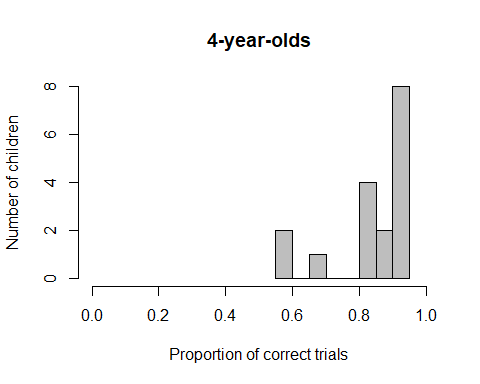
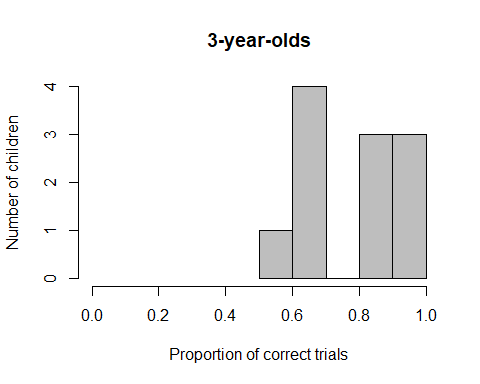


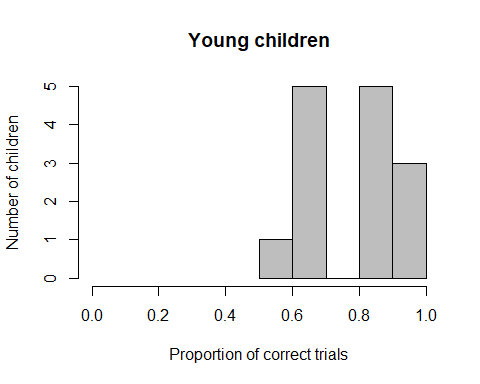
In the EDS phase, the mean number of errors was 5.97 (SD = 4.67, range 0-17). The distribution of this variable is left-skewed.

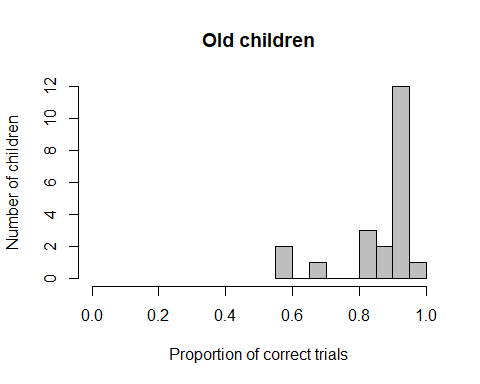
* 3y (n = 11): 8.00 (SD = 5.14, range 1-17)
* 4y (n = 17): 5.53 (SD = 4.33, range 2-15)
* 5y (n = 7): 3.86 (SD = 4.02, range 0-12)
* young (n = 14): 8.00 (SD = 4.62, range 1-17)
* old (n = 21): 4.62 (SD = 4.28, range 0-15)
* Edinburgh (n = 17): 7.59 (SD = 5.51, range 1-17)
* Fife (n = 18): 4.44 (SD = 3.15, range 0-12)

## Proportion of correct trials









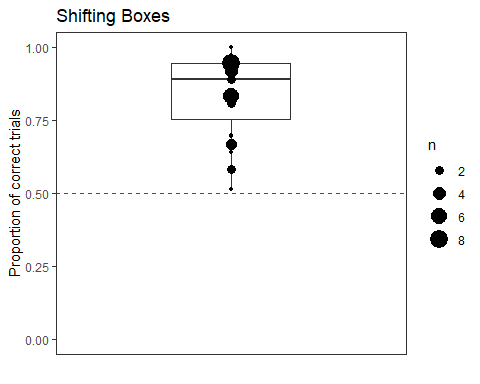
In the EDS phase, the mean proportion of correct trials was **0.83 (SD = 0.13, range 0.51-1)**. The distribution of this variable is right-skewed and **not normally distributed**, W = 0.867, p < .001. Performance is significantly **above chance**, V = 630, p < .001.

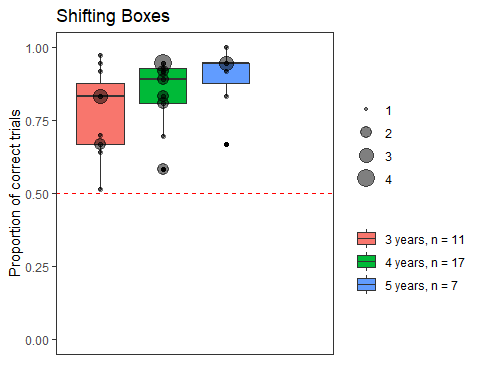
* 3y (n = 11): **0.77 (SD = 0.15, range 0.51-.97)**, normally distributed, W = 0.937, p = .484, performance above chance, t(10) = 6.22, p < .001
* 4y (n = 17): **0.84 (SD = 0.12, range 0.58-.94)**, not normally distributed, W = 0.784, p = .001, performance above chance, V = 153, p < .001
* 5y (n =7): **0.89 (SD = 0.11, range 0.67-1),** normally distributed, W = 0.805, p = .046, performance above chance, t(6) = 9.324, p < .001
* young (n = 14): **0.77 (SD = 0.13, range 0.51-.97)**, normally distributed, W = 0.946, p = .499, performance **above chance**, t(13) = 7.827, p < .001
* old (n = 21): **0.87 (SD = 0.12, range 0.58-1)**, not normally distributed, W = 0.939, p < .001, performance **above chance,** V = 231, p < .001

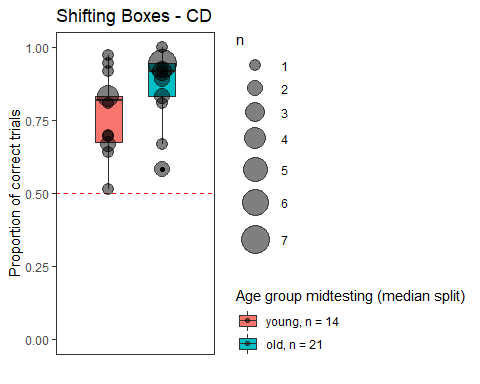
**Older children are significantly better than younger** children, W = 210.5, p = .016.

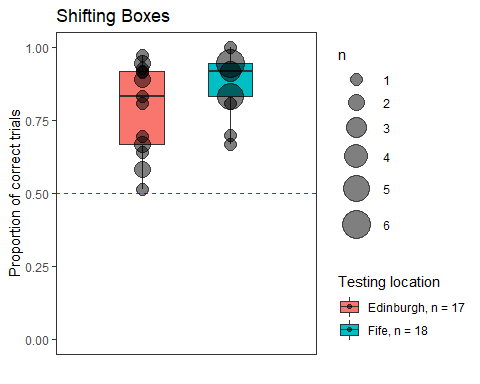
* Edinburgh (n = 17): 0.79 (SD = 0.15, range 0.51-.97), not normally distributed, W = 0.863, p = .014, performance above chance, V = 153, p < .001
* Fife (n = 18): 0.87 (SD = 0.09, range 0.67-1), not normally distributed, W = 0.912, p < .001, performance above chance, V = 171, p < .001

**Children in Fife and Edinburgh do not differ** in their performance in CD, W = 99.5, p = .078

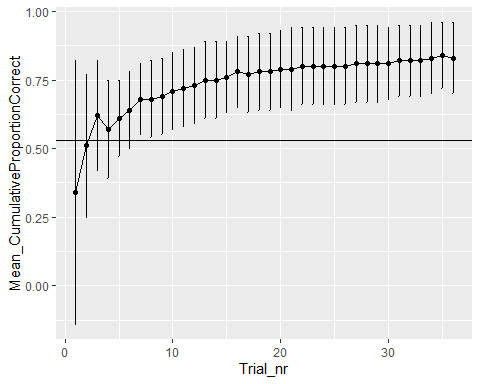








## Cumulative Proportion Correct



# EDS: Only those children who reached criterion in SD and then proceeded to CD

111 children reached the criterion in SD and moved on to CD. Of these, 27 children reached criterion in CD and qualified for EDS. 26 children were tested in EDS, 1 child is missing as E accidentally did not administer EDS.

## Dropouts

1 child was a dropout (4y, m), stopped after 24 trials. We decided to remove this child, as they did not reach 75% of the trials (i.e., 27 trials). There is also 1 child where E accidentally switched to EDS after only 5 consecutive trials in CD, 1 child who stopped after 33 trials, 1 child where E accidentally only administered 35 trials, and 1 child with 35 trials because E accidentally skipped one trial) - these are all included.

## Valid data

There are **25 valid datapoints** in EDS.

## Description of gender distribution in EDS

There were 14 girls and 11 boys.

## Description of age distribution in EDS

### Age at beginning of testing

The 24 children who had valid data on the EDS phase were on average 51.80 months old (SD = 6.10, range 41-64) at the beginning of testing. There were 6 3-year-olds, 16 4-year-olds, and 3 5-year-olds.

* 3-year-olds: 5 f, 1 m
* 4-year-olds: 8 f, 8 m
* 5-year-olds: 1 f, 2 m

### Age in the middle of testing

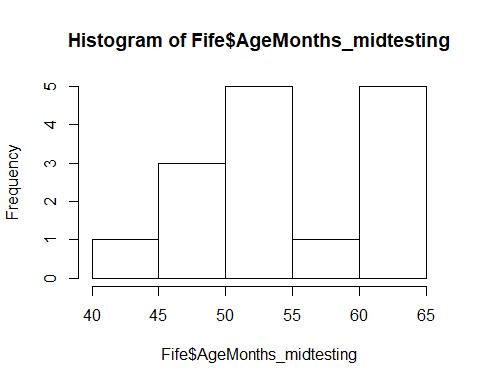
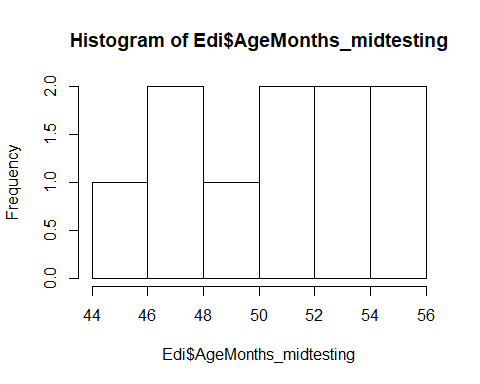
The 25 children who had valid data on the EDS phase were on average 52.80 months old (SD = 6.11, range 43-65) in the middle of testing. There were 6 3-year-olds, 14 4-year-olds, and 5 5-year-olds.

* 3-year-olds: 5 f, 1 m
* 4-year-olds: 6 f, 8 m
* 5-year-olds: 3 f, 2 m

### Mediansplit age

There are 8 young children (5 f, 3 m) and 17 old children (9 f, 8 m).

## Testing location



10 children were from Edinburgh, 15 children from Fife

Fife: M = 54.07 (SD = 7.18, range 43-65)

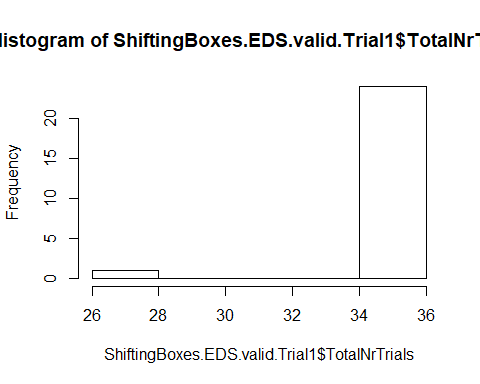
* 4 3y
* 6 4y
* 5 5y

Edinburgh: M = 50.90 (SD = 3.57, range 45-56)

* 2 3y
* 8 4y

There is no age difference between the two testing locations, two-sided t-test, t(21.68) = -1.46, p = .159.

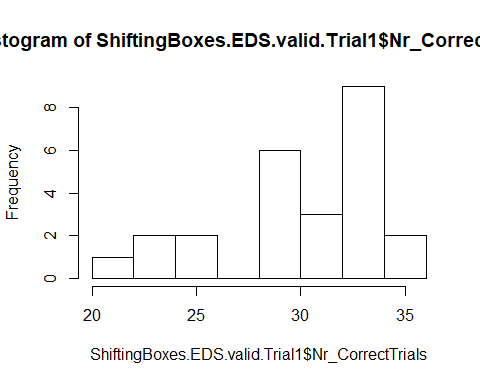
## Total number of trials



In the EDS phase, the mean number of administered trials was 35.60 (SD = 1.80, range 27-36).

* 3y (n = 6): 36
* 4y (n = 14): 35.28 (SD = 2.40, range 27-36)
* 5y (n = 5): 36
* young (n = 8): 36
* old (n = 17): 35.41 (SD = 2.18, range 27-36)
* Edinburgh (n = 10): 35 (SD = 2.83, range 27-36)
* Fife (n = 15): 36 (SD = 0)

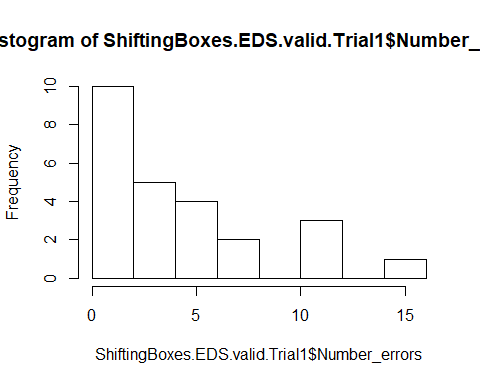
## Number of correct trials



In the EDS phase, the mean number of correct trials was 30.72 (SD = 4.06, range 21-36). The distribution of this variable has 2 peaks, one around 34, the other around 23.

* 3y (n = 6): 30.33 (SD = 3.72, range 24-35)
* 4y (n = 14): 30.28 (SD = 4.08, range 21-34)
* 5y (n = 5): 32.40 (SD = 4.77, range 24-36)
* young (n = 8): 29.50 (SD = 3.66, range 24-35)
* old (n = 17): 31.29 (SD = 4.21, range 21-36)
* Edinburgh (n = 10): 28.90 (SD = 4.82, range 21-35)
* Fife (n = 15): 31.93 (SD = 3.06, range 24-36)

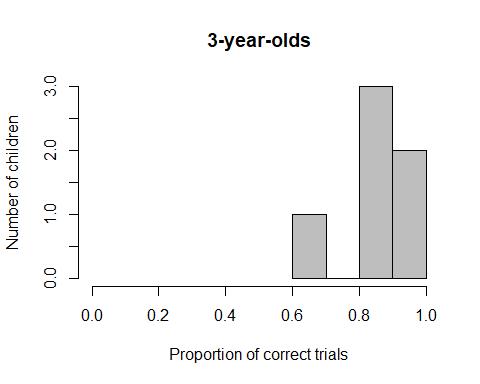
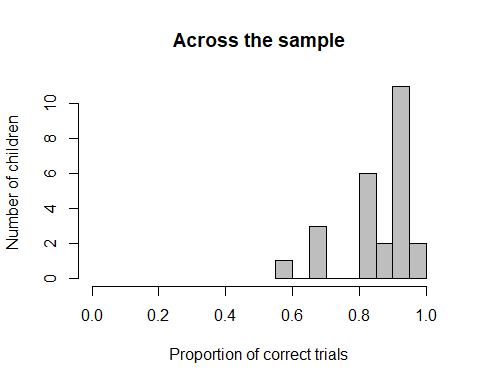
## Number of errors

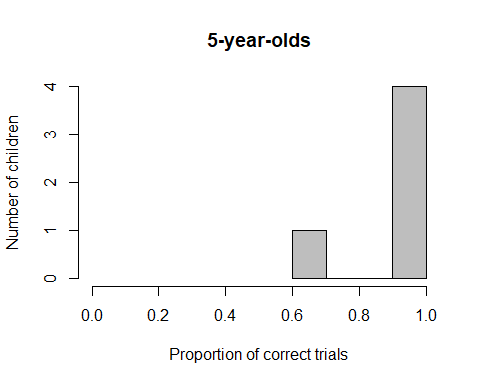
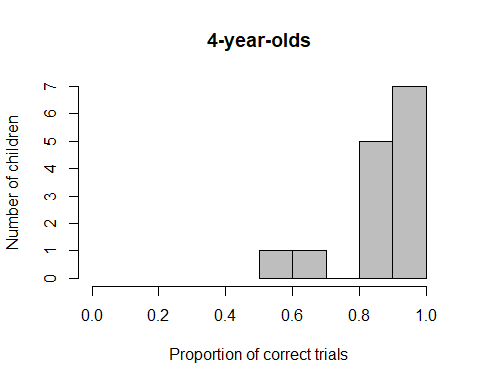


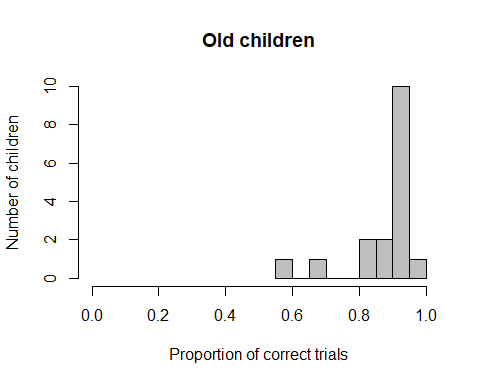
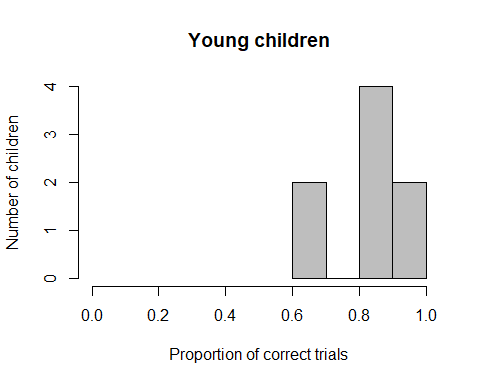
In the EDS phase, the mean number of errors was 4.88 (SD = 3.94, range 0-15). The distribution of this variable is left-skewed.

* 3y (n = 6): 5.67 (SD = 3.72, range 1-12)
* 4y (n = 14): 5.00 (SD = 3.92, range 2-15)
* 5y (n = 5): 3.60 (SD = 4.77, range 0-12)
* young (n = 8): 6.50 (SD = 3.66, range 1-12)
* old (n = 17): 4.12 (SD = 3.93, range 0-15)
* Edinburgh (n = 10): 6.10 (SD = 4.91, range 1-15)
* Fife (n = 15): 4.07 (SD = 3.06, range 0-12)

## Proportion of correct trials







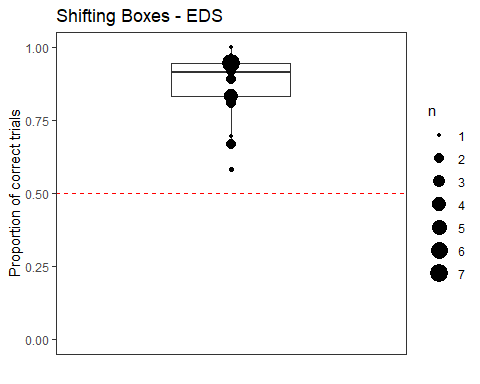
In the EDS phase, the mean proportion of correct trials was **0.86 (SD = 0.11, range 0.58-1)**. The distribution of this variable is right-skewed and **not normally distributed**, W = 0.852, p = .002. Performance is significantly **above chance**, V = 325, p < .001.

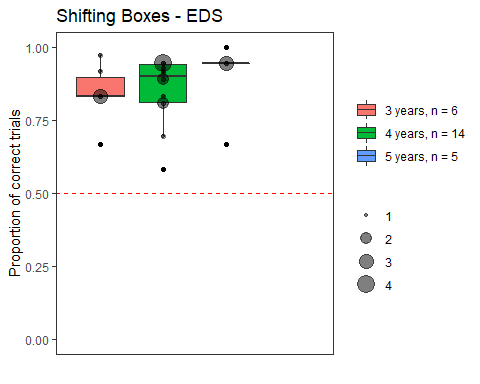
* 3y (n = 6): **0.84 (SD = 0.10, range 0.67-.97)**, normally distributed, W = 0.904, p = .398, performance above chance, t(5) = 8.119, p < .001
* 4y (n = 14): **0.86 (SD = 0.11, range 0.58-.94)**, not normally distributed, W = 0.789, p = .004, performance above chance, V = 105, p = .001
* 5y (n =5): **0.90 (SD = 0.13, range 0.67-1)**, not normally distributed, W = 0.699, p = .009, **performance at chance**, V = 15, p = .054
* young (n = 8): **0.82 (SD = 0.10, range 0.67-.97),** normally distributed, W = 0.931, p = .526, performance above chance, t(7) = 8.876, p < .001
* old (n = 17): **0.88 (SD = 0.11, range 0.58-1)**, not normally distributed, W = 0.735, p < .001, performance above chance, V = 153, p < .001

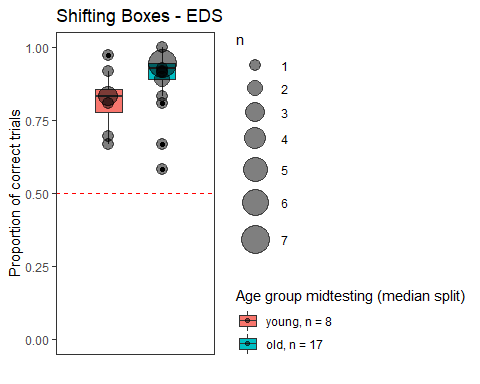
**Older children are not better than younger** children, W = 96, p = .052.

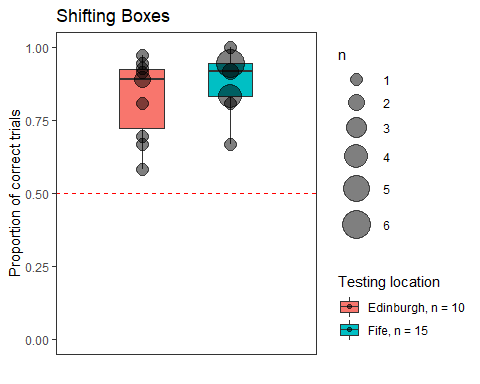
* Edinburgh (n = 10): 0.83 (SD = 0.13, range 0.58-.97), normally distributed, W = 0.873, p = .109, performance above chance, t(9) = 7.716, p < .001
* Fife (n = 15): 0.89 (SD = 0.08, range 0.67-1), not normally distributed, W = 0.845, p = .015, performance above chance, V = 120, p < .001

**Children in Fife and Edinburgh do not differ** in their performance, W = 54, p = .249

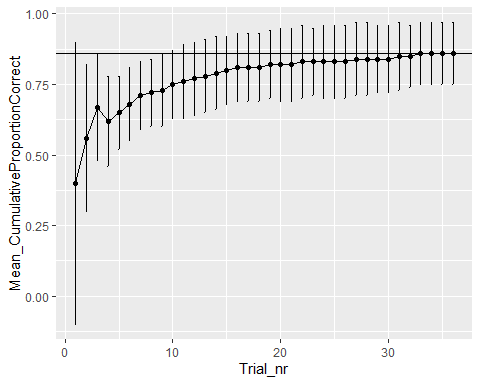




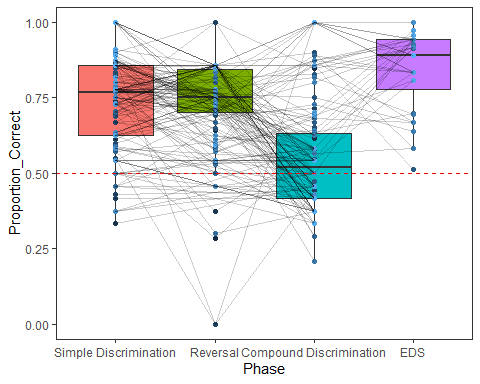




## Cumulative Proportion Correct



# Boxplots: Mean proportion correct in SD, Reversal, CD (only those children who reached criterion in SD) and EDS (only those children who reached criterion in SD)



# What is the effect of age and trial number on performance in CD (only children who reached criterion in CD)?

#Full model  
contr<-glmerControl(optimizer="bobyqa", optCtrl=list(maxfun=10000000))  
  
res<-glmer(success ~ z.age.midtesting + z.Trial\_nr + z.age.midtesting:z.Trial\_nr + (1+z.Trial\_nr|ID), data=ShiftingBoxes.CD.valid, family = binomial, control = contr)#singular fit

## boundary (singular) fit: see ?isSingular

res<-glmer(success ~ z.age.midtesting + z.Trial\_nr + z.age.midtesting:z.Trial\_nr + (1|ID) + (0+z.Trial\_nr|ID), data=ShiftingBoxes.CD.valid, family = binomial, control = contr)#converges, singular fit

The initial full model yields a singular fit. This is because the variance for the random slope of trial on ID is basically 0, so we remove this term from the model.

res<-glmer(success ~ z.age.midtesting + z.Trial\_nr + z.age.midtesting:z.Trial\_nr + (1|ID), data=ShiftingBoxes.CD.valid, family = binomial, control = contr)#converges, no singular fit warning

Together, age, trial number and the interaction between age and trial number do not explain the data significantly better than a null model containing only the intercept, X2(3) = 6.291, p = .098.