Ripped Foil - Children

Zeynep Civelek

June 10, 2020

#PREPARE  
R.Version()#for referencing, shows you which R version you are using

## $platform  
## [1] "x86\_64-w64-mingw32"  
##   
## $arch  
## [1] "x86\_64"  
##   
## $os  
## [1] "mingw32"  
##   
## $system  
## [1] "x86\_64, mingw32"  
##   
## $status  
## [1] ""  
##   
## $major  
## [1] "3"  
##   
## $minor  
## [1] "6.3"  
##   
## $year  
## [1] "2020"  
##   
## $month  
## [1] "02"  
##   
## $day  
## [1] "29"  
##   
## $`svn rev`  
## [1] "77875"  
##   
## $language  
## [1] "R"  
##   
## $version.string  
## [1] "R version 3.6.3 (2020-02-29)"  
##   
## $nickname  
## [1] "Holding the Windsock"

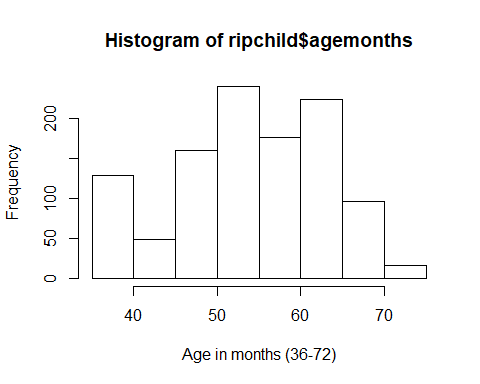
rm(list=ls())#removes any other items in your workspace  
ls()#check whether workspace is empty

## character(0)

#### The structure of the data is shown below with all the variable names and levels of measurement.

## 'data.frame': 1088 obs. of 15 variables:  
## $ id : Factor w/ 68 levels "14","15","16",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ age : Factor w/ 3 levels "3","4","5": 3 3 3 3 3 3 3 3 3 3 ...  
## $ agemonths: num 62.9 62.9 62.9 62.9 62.9 ...  
## $ ageyears : num 5.24 5.24 5.24 5.24 5.24 ...  
## $ agecat : Factor w/ 3 levels "five","four",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ sex : Factor w/ 2 levels "1","2": 1 1 1 1 1 1 1 1 1 1 ...  
## $ order : Factor w/ 2 levels "penfirst","stickerfirst": 1 1 1 1 1 1 1 1 1 1 ...  
## $ boxtype : Factor w/ 2 levels "blue-right","pink-right": 1 1 1 1 1 1 1 1 1 1 ...  
## $ trialtype: Factor w/ 2 levels "pen","sticker": 1 1 1 1 1 1 1 1 2 2 ...  
## $ order2 : int 1 1 1 1 1 1 1 1 2 2 ...  
## $ order3 : Factor w/ 2 levels "test","transfer": 1 1 1 1 1 1 1 1 2 2 ...  
## $ trialno : int 1 2 3 4 5 6 7 8 1 2 ...  
## $ side : Factor w/ 2 levels "L","R": 1 2 2 1 2 1 1 2 2 1 ...  
## $ response : Factor w/ 3 levels "","L","R": 3 2 2 3 2 3 3 2 3 2 ...  
## $ correct : int 1 1 1 1 1 1 1 1 1 1 ...

#### We decided to use age as a continuous variable (age in months) in our main analyses rather than categorical to have more power.The distribution of age:



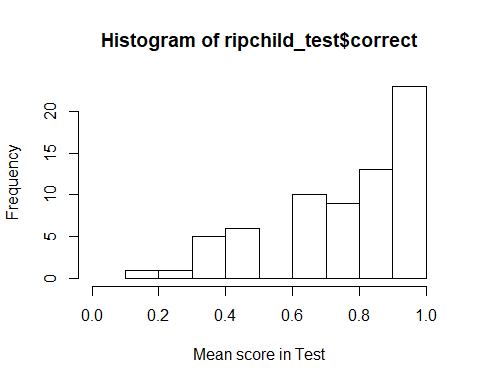
#### I aggregated the trial-by-trial data to create one score per child for Test and Transfer phases so I can check whether they are normally distributed or not.

## `summarise()` regrouping output by 'id', 'age' (override with `.groups` argument)

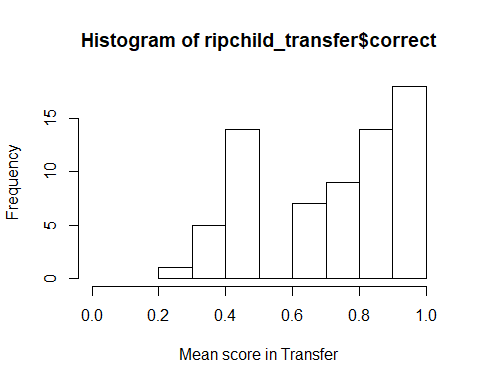
## `summarise()` regrouping output by 'id' (override with `.groups` argument)  
## `summarise()` regrouping output by 'id' (override with `.groups` argument)

## Histograms for Test and Transfer

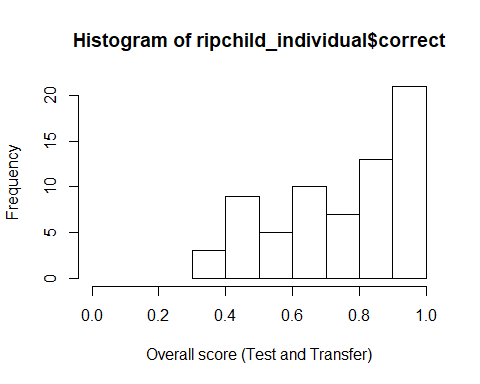
hist(ripchild\_test$correct, xlab="Mean score in Test", xlim=c(0,1))



hist(ripchild\_transfer$correct, xlab="Mean score in Transfer", xlim=c(0,1))



hist(ripchild\_individual$correct, xlab="Overall score (Test and Transfer)", xlim=c(0,1))



## The normality tests

shapiro.test(ripchild\_individual$correct)

##   
## Shapiro-Wilk normality test  
##   
## data: ripchild\_individual$correct  
## W = 0.92081, p-value = 0.0003523

shapiro.test(ripchild\_test$correct)

##   
## Shapiro-Wilk normality test  
##   
## data: ripchild\_test$correct  
## W = 0.86957, p-value = 3.891e-06

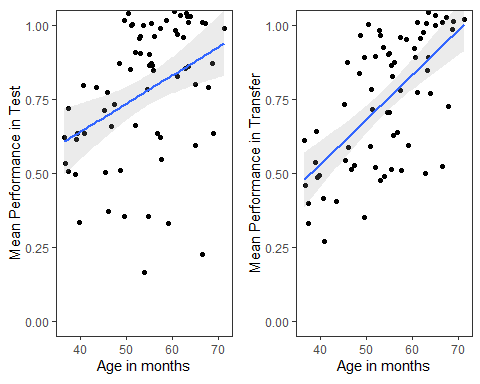
shapiro.test(ripchild\_transfer$correct)

##   
## Shapiro-Wilk normality test  
##   
## data: ripchild\_transfer$correct  
## W = 0.88645, p-value = 1.522e-05

#### The Shapiro-Wilk normality tests confirm that the data are not normally distributed for Test and Transfer phases (neither for an overall score).

#### Below is the how children in different ages perform in test and transfer.

## `geom\_smooth()` using formula 'y ~ x'  
## `geom\_smooth()` using formula 'y ~ x'



#### The regression line (with a 95% CI) for the Transfer phase is steeper than the Test phase. Younger children seem to be doing worse in Transfer than in Test.

#### Next is the pre-processing of data: z-transformations for 'trial no' and 'age in months', coding dummy variables for trialtype, sex and phase (order3), and finally centering the random slope components for these variables.

## id age agemonths ageyears agecat sex   
## 14 : 16 3:272 Min. :36.46 Min. :3.038 five :336 1:576   
## 15 : 16 4:480 1st Qu.:48.12 1st Qu.:4.010 four :480 2:512   
## 16 : 16 5:336 Median :54.84 Median :4.570 three:272   
## 17 : 16 Mean :54.00 Mean :4.500   
## 18 : 16 3rd Qu.:61.26 3rd Qu.:5.105   
## 19 : 16 Max. :71.26 Max. :5.939   
## (Other):992   
## order boxtype trialtype order2 order3   
## penfirst :544 blue-right:512 pen :544 Min. :1.0 test :544   
## stickerfirst:544 pink-right:576 sticker:544 1st Qu.:1.0 transfer:544   
## Median :1.5   
## Mean :1.5   
## 3rd Qu.:2.0   
## Max. :2.0   
##   
## trialno side response correct z.trialno   
## Min. :1.00 L:542 : 48 Min. :0.0000 Min. :-1.5268   
## 1st Qu.:2.75 R:546 L:537 1st Qu.:1.0000 1st Qu.:-0.7634   
## Median :4.50 R:503 Median :1.0000 Median : 0.0000   
## Mean :4.50 Mean :0.7583 Mean : 0.0000   
## 3rd Qu.:6.25 3rd Qu.:1.0000 3rd Qu.: 0.7634   
## Max. :8.00 Max. :1.0000 Max. : 1.5268   
##   
## z.age   
## Min. :-1.92987   
## 1st Qu.:-0.64675   
## Median : 0.09227   
## Mean : 0.00000   
## 3rd Qu.: 0.79909   
## Max. : 1.90012   
##

#Coding dummy variables before centering the slopes  
ripchild$trialtype.sticker<-as.numeric(ripchild$trialtype==levels(ripchild$trialtype)[2])  
ripchild$sex.m<-as.numeric(ripchild$sex==levels(ripchild$sex)[2])  
ripchild$order3.transfer<-as.numeric(ripchild$order3==levels(ripchild$order3)[2])  
  
#Centering the slopes  
ripchild$trialtype.sticker.c<-ripchild$trialtype.sticker-mean(ripchild$trialtype.sticker)  
ripchild$sex.m.c<-ripchild$sex.m -mean(ripchild$sex.m)  
ripchild$order3.transfer.c<-ripchild$order3.transfer-mean(ripchild$order3.transfer)

## Full model

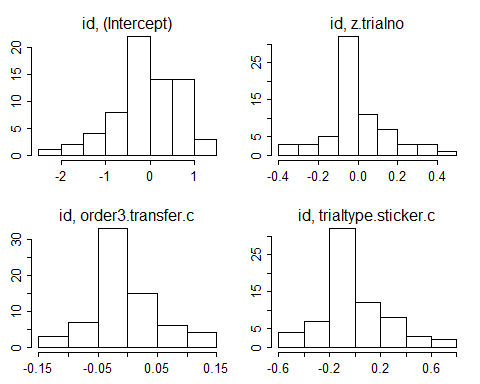
contr<-glmerControl(optimizer="bobyqa", optCtrl=list(maxfun=10000000))  
full<-glmer(correct ~ trialtype\*order3+z.trialno+z.age\*order3+sex+(1|id)+(0+z.trialno+order3.transfer.c+trialtype.sticker.c|id),data=ripchild, family=binomial, control=contr)

## boundary (singular) fit: see ?isSingular

#### The full model does not include box type (the location of the blue/pink box) as a random effect anymore due to convergence issues.

## Model assumptions

#### Distribution of random effects



#### Model stability

#### Multicollinearity

## trialtype order3 z.trialno z.age   
## 2.004 2.004 1.000 2.021   
## sex trialtype:order3 order3:z.age   
## 1.025 3.011 2.000

#### Overdispersion

## chisq df P dispersion.parameter  
## 1 814.953 1073 1 0.7595088

## Null model

null<-glmer(correct~ 1+(1|id)+(0+z.trialno+order3.transfer.c+trialtype.sticker.c|id),data=ripchild, family=binomial, control=contr)

## boundary (singular) fit: see ?isSingular

## Full and Null comparison

round(anova(null, full, test="Chisq"),3)

## Data: ripchild  
## Models:  
## null: correct ~ 1 + (1 | id) + (0 + z.trialno + order3.transfer.c +   
## null: trialtype.sticker.c | id)  
## full: correct ~ trialtype \* order3 + z.trialno + z.age \* order3 + sex +   
## full: (1 | id) + (0 + z.trialno + order3.transfer.c + trialtype.sticker.c |   
## full: id)  
## npar AIC BIC logLik deviance Chisq Df Pr(>Chisq)   
## null 8 1121.8 1161.8 -552.91 1105.8   
## full 15 1100.2 1175.1 -535.09 1070.2 35.646 7 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Model output

#### Coefficients

round(summary(full)$coefficients, 3)

## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) 1.517 0.282 5.370 0.000  
## trialtypesticker 0.256 0.335 0.764 0.445  
## order3transfer -0.184 0.329 -0.560 0.576  
## z.trialno 0.135 0.090 1.491 0.136  
## z.age 0.635 0.168 3.772 0.000  
## sex2 -0.135 0.288 -0.468 0.640  
## trialtypesticker:order3transfer 0.108 0.569 0.190 0.849  
## order3transfer:z.age 0.293 0.174 1.685 0.092

#### Individual predictor: Likelihood tests

xdrop1=drop1(full, test="Chisq",control=contr)

## boundary (singular) fit: see ?isSingular  
## boundary (singular) fit: see ?isSingular  
## boundary (singular) fit: see ?isSingular  
## boundary (singular) fit: see ?isSingular

round(xdrop1,3)

## Single term deletions  
##   
## Model:  
## correct ~ trialtype \* order3 + z.trialno + z.age \* order3 + sex +   
## (1 | id) + (0 + z.trialno + order3.transfer.c + trialtype.sticker.c |   
## id)  
## npar AIC LRT Pr(Chi)   
## <none> 1100.2   
## z.trialno 1 1100.4 2.228 0.136   
## sex 1 1098.4 0.218 0.641   
## trialtype:order3 1 1098.2 0.036 0.849   
## order3:z.age 1 1101.1 2.926 0.087 .  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#### Confidence intervals for the full model with the interaction

# model1CI=boot.glmm.pred(model.res=full, excl.warnings=T, nboots=1000, para=T)  
# round(model1CI$ci.estimates, 3)  
# save.image("model1\_CIs.RData")  
load("model1\_CIs.RData")  
round(model1CI$ci.estimates, 3)

## orig X2.5. X97.5.  
## (Intercept) 1.517 0.982 2.102  
## trialtypesticker 0.256 -0.454 0.944  
## order3transfer -0.184 -0.856 0.450  
## z.trialno 0.135 -0.055 0.322  
## z.age 0.635 0.321 0.976  
## sex2 -0.135 -0.718 0.449  
## trialtypesticker:order3transfer 0.108 -1.038 1.296  
## order3transfer:z.age 0.293 -0.047 0.679

#### The interactions terms are not significant (although there is a trend for scores to vary by age) so I will remove them from the model.

## Reduced model without the interactions

## Full model 2

full2=glmer(correct ~ trialtype+order3+z.trialno+z.age+sex+(1|id)+(0+z.trialno+order3.transfer.c+trialtype.sticker.c|id),data=ripchild, family=binomial, control=contr)

## boundary (singular) fit: see ?isSingular

summary(full2)

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: correct ~ trialtype + order3 + z.trialno + z.age + sex + (1 |   
## id) + (0 + z.trialno + order3.transfer.c + trialtype.sticker.c | id)  
## Data: ripchild  
## Control: contr  
##   
## AIC BIC logLik deviance df.resid   
## 1099.1 1164.0 -536.6 1073.1 1075   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.4044 0.1759 0.3170 0.5272 1.3195   
##   
## Random effects:  
## Groups Name Variance Std.Dev. Corr   
## id (Intercept) 0.83665 0.9147   
## id.1 z.trialno 0.09792 0.3129   
## order3.transfer.c 0.01593 0.1262 1.00   
## trialtype.sticker.c 0.28811 0.5368 1.00 1.00  
## Number of obs: 1088, groups: id, 68  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 1.53972 0.24073 6.396 1.59e-10 \*\*\*  
## trialtypesticker 0.29682 0.17560 1.690 0.091 .   
## order3transfer -0.21821 0.17202 -1.269 0.205   
## z.trialno 0.13694 0.09012 1.520 0.129   
## z.age 0.78166 0.14752 5.299 1.17e-07 \*\*\*  
## sex2 -0.12938 0.28758 -0.450 0.653   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) trltyp ordr3t z.trln z.age   
## trltypstckr -0.354   
## ordr3trnsfr -0.410 0.073   
## z.trialno -0.049 0.204 0.034   
## z.age 0.225 0.032 -0.014 0.036   
## sex2 -0.567 0.001 0.002 0.002 -0.171  
## convergence code: 0  
## boundary (singular) fit: see ?isSingular

## Null model

null<-glmer(correct~ 1+(1|id)+(0+z.trialno+order3.transfer.c+trialtype.sticker.c|id),data=ripchild, family=binomial, control=contr)

## boundary (singular) fit: see ?isSingular

#### Full 2 and null comparision

round(anova(null, full2, test="Chisq"),3)

## Data: ripchild  
## Models:  
## null: correct ~ 1 + (1 | id) + (0 + z.trialno + order3.transfer.c +   
## null: trialtype.sticker.c | id)  
## full2: correct ~ trialtype + order3 + z.trialno + z.age + sex + (1 |   
## full2: id) + (0 + z.trialno + order3.transfer.c + trialtype.sticker.c |   
## full2: id)  
## npar AIC BIC logLik deviance Chisq Df Pr(>Chisq)   
## null 8 1121.8 1161.8 -552.91 1105.8   
## full2 13 1099.1 1164.0 -536.57 1073.1 32.692 5 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Model output

#### Coefficients

## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) 1.540 0.241 6.396 0.000  
## trialtypesticker 0.297 0.176 1.690 0.091  
## order3transfer -0.218 0.172 -1.269 0.205  
## z.trialno 0.137 0.090 1.520 0.129  
## z.age 0.782 0.148 5.299 0.000  
## sex2 -0.129 0.288 -0.450 0.653

#### Individual predictors : Likelihod ratio tests

## boundary (singular) fit: see ?isSingular  
## boundary (singular) fit: see ?isSingular  
## boundary (singular) fit: see ?isSingular  
## boundary (singular) fit: see ?isSingular  
## boundary (singular) fit: see ?isSingular

## Single term deletions  
##   
## Model:  
## correct ~ trialtype + order3 + z.trialno + z.age + sex + (1 |   
## id) + (0 + z.trialno + order3.transfer.c + trialtype.sticker.c |   
## id)  
## npar AIC LRT Pr(Chi)   
## <none> 1099.1   
## trialtype 1 1100.0 2.856 0.091 .   
## order3 1 1098.7 1.605 0.205   
## z.trialno 1 1099.5 2.315 0.128   
## z.age 1 1123.2 26.077 <2e-16 \*\*\*  
## sex 1 1097.3 0.202 0.653   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#### Age emerges as a significant predictor of scores.

#### Confidence intervals for the reduced model without the interaction

# model2CI=boot.glmm.pred(model.res=full2, excl.warnings=T, nboots=1000, para=T)  
# round(model2CI$ci.estimates, 3)  
# save.image("model2\_CIs.RData")  
load("model2\_CIs.RData")  
round(model2CI$ci.estimates, 3)

## orig X2.5. X97.5.  
## (Intercept) 1.540 1.081 2.026  
## trialtypesticker 0.297 -0.045 0.662  
## order3transfer -0.218 -0.556 0.146  
## z.trialno 0.137 -0.037 0.317  
## z.age 0.782 0.480 1.097  
## sex2 -0.129 -0.685 0.406

#### Calculating effect sizes for the final model

## Warning: 'r.squaredGLMM' now calculates a revised statistic. See the help page.

## Warning: The null model is correct only if all variables used by the original  
## model remain unchanged.

## boundary (singular) fit: see ?isSingular

## R2m R2c  
## theoretical 0.13190079 0.3358856  
## delta 0.09146174 0.2329075

#### Running a separate model for the CI calculation. Age is entered as age in months (notz transformed).

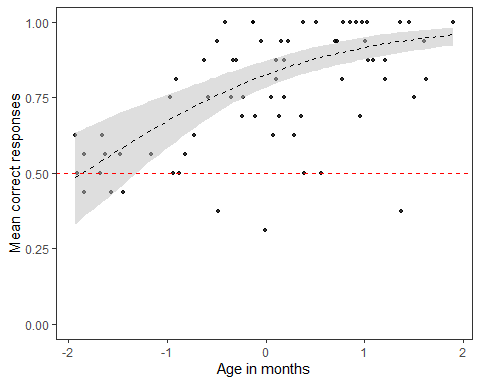
#running a new model for CI predictions  
full.CI=glmer(correct ~ z.trialno+order3.transfer.c+trialtype.sticker.c+z.age+sex.m.c+(1+z.trialno+order3.transfer.c+trialtype.sticker.c|id),data=ripchild, family=binomial, control=contr)

## boundary (singular) fit: see ?isSingular

#### Getting confidence intervals for fitted values

#### Overall performance across age.

## `summarise()` regrouping output by 'id' (override with `.groups` argument)



#### The plot is not informative as age in z-transformed. I correct this below.

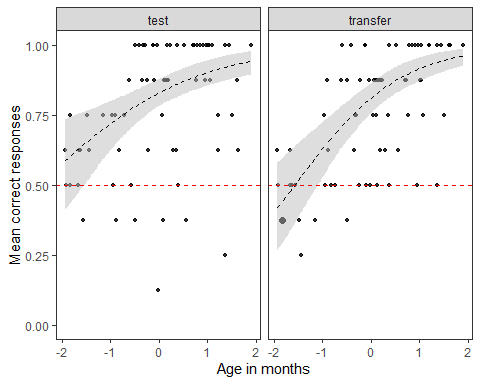
#### Running a new model for calculating CIs for the interaction model so we can plot the performance in test and transfer separately.

full.CI.int<-glmer(correct ~ trialtype.sticker.c+z.age\*order3+sex.m.c+z.trialno+(1|id)+(0+z.trialno+order3.transfer.c+trialtype.sticker.c|id),data=ripchild, family=binomial, control=contr)

## boundary (singular) fit: see ?isSingular

#### The plot for performance in Test and Transfer across age.

## `summarise()` regrouping output by 'id', 'z.age' (override with `.groups` argument)

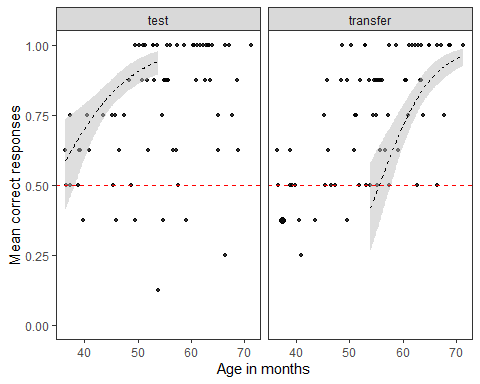


#### However, the age is z-transformed so the plot is not very informative. We change z-age to age in months.

#do the same scatter plot for interaction with age not z-tranformed so the graph is more informative.  
ripchild\_separate2 <- ripchild %>%  
 group\_by(id, agemonths, order3) %>%   
 summarize(correct= mean(correct))%>%  
 ungroup()%>%  
 group\_by(agemonths, order3)%>%  
 add\_count(correct)

## `summarise()` regrouping output by 'id', 'agemonths' (override with `.groups` argument)

plot.int=seq(from=min(ripchild\_separate2$agemonths),  
 to=max(ripchild\_separate2$agemonths),length.out=272)  
  
scatter.int2<-ggplot(ripchild\_separate2, aes(x=plot.int, y=boot.res.int$ci.predicted$fitted, group=order3)) +  
 geom\_point(aes(x=agemonths, y=correct), size=ripchild\_separate2$n, alpha=0.8) + facet\_wrap(~order3) + geom\_ribbon(data=boot.res.int$ci.predicted, aes(x=plot.int, y=boot.res.int$ci.predicted$fitted, ymin=boot.res.int$ci.predicted$lower.cl, ymax=boot.res.int$ci.predicted$upper.cl, group=boot.res.int$ci.predicted$order3),fill="grey", alpha=0.5) + geom\_line(data=boot.res.int$ci.predicted, aes(y=boot.res.int$ci.predicted$fitted), lty=2)+  
 theme\_bw() + theme(panel.grid.major = element\_blank(), panel.grid.minor = element\_blank())+ xlab("Age in months")+ ylab("Mean correct responses")+   
 ylim(0,1)+  
 geom\_hline(yintercept=0.50, color='red', linetype="dashed")  
   
 scatter.int2



#### The CIs do not cover the entire range here. How can we fix this?

#### Getting the data for the first trial of Transfer

ripchild\_1st\_trial <- ripchild %>%  
 filter(order3=="transfer" & trialno=="1") %>%  
 group\_by(id, age, agemonths, sex, trialtype, z.age, trialtype.sticker.c) %>%   
 summarize(correct)

## `summarise()` regrouping output by 'id', 'age', 'agemonths', 'sex', 'trialtype', 'z.age' (override with `.groups` argument)

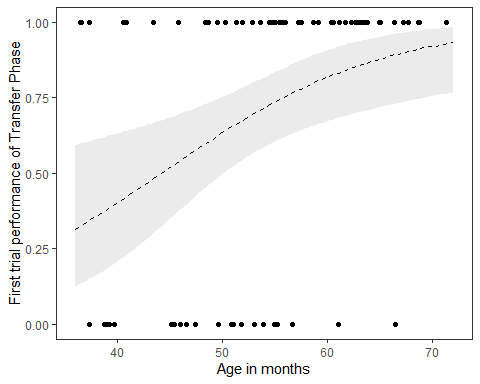
#### Trial 1 success in Transfer

firsttrial <- glm(correct~agemonths+sex+trialtype,data=ripchild\_1st\_trial,family=binomial(link = "logit"))  
summary(firsttrial)

##   
## Call:  
## glm(formula = correct ~ agemonths + sex + trialtype, family = binomial(link = "logit"),   
## data = ripchild\_1st\_trial)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.0658 -1.0328 0.5699 0.8060 1.5530   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -4.36231 1.80644 -2.415 0.01574 \*   
## agemonths 0.09633 0.03414 2.822 0.00478 \*\*  
## sex2 -0.02901 0.57703 -0.050 0.95990   
## trialtypesticker 0.22316 0.57090 0.391 0.69588   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 84.069 on 67 degrees of freedom  
## Residual deviance: 74.436 on 64 degrees of freedom  
## AIC: 82.436  
##   
## Number of Fisher Scoring iterations: 4

#### Age is a significant predictor of first trial performance of Transfer.

## `geom\_smooth()` using formula 'y ~ x'



#### Here is the plot for the first trial performance in Transfer.