Lab 4

Aniket Naik Desai

2019-02-11

# Abstract

This study investigated associations between working memory (measured by complex memory tasks) and both reading and mathematics abilities, as well as the possible mediating factors of fluid intelligence, verbal abilities, short-term memory (STM), and phonological awareness, in a sample of 6- to 11-year-olds with reading disabilities. As a whole, the sample was characterized by deficits in complex memory and visuospatial STM and by low IQ scores; language, phonological STM, and phonological awareness abilities fell in the low average range. Severity of reading difficulties within the sample was significantly associated with complex memory, language, and phonological awareness abilities, whereas poor mathematics abilities were linked with complex memory, phonological STM, and phonological awareness scores. These findings suggest that working memory skills indexed by complex memory tasks represent an important constraint on the acquisition of skill and knowledge in reading and mathematics. Possible mechanisms for the contribution of working memory to learning, and the implications for educational practice, are considered.

*Citation:* Gathercole, S. E., Alloway, T. P., Willis, C., & Adams, A. M. (2006). Working memory in children with reading disabilities. Journal of Experimental Child Psychology, 93(3), 265-281.

# Dataset:

- Dependent variable (Y): Reading - reading skills of the 6 to 11 year olds  
- Independent variables (X):  
 - Verbal - a measure of verbal ability (spelling, phonetics, etc.)  
 - Math - a measure of math ability  
 - Work\_mem - working memory score

##import the datafile  
library(haven)  
Lab4 = read\_spss("Lab4.sav")

# Data screening:

## Accuracy

Assume the data is accurate with no missing values. You will want to screen the dataset using all the predictor variables to predict the outcome in a simultaneous multiple regression (all the variables at once). This analysis will let you screen for outliers and assumptions across all subsequent analyses/steps.

## Outliers

a. Leverage:  
 i. What is your leverage cut off score?  
 ii. How many leverage outliers did you have?

model1 = lm(reading ~ verbal + math + work\_mem, data = Lab4)  
  
summary(model1)

##   
## Call:  
## lm(formula = reading ~ verbal + math + work\_mem, data = Lab4)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.8663 -0.7895 -0.3433 0.5104 7.8704   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.4413428 0.0409707 35.180 < 2e-16 \*\*\*  
## verbal -0.0068428 0.0049394 -1.385 0.166   
## math 0.0049810 0.0007971 6.249 4.72e-10 \*\*\*  
## work\_mem 0.0003228 0.0015434 0.209 0.834   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.09 on 3016 degrees of freedom  
## Multiple R-squared: 0.03238, Adjusted R-squared: 0.03141   
## F-statistic: 33.64 on 3 and 3016 DF, p-value: < 2.2e-16

#1.  
k = 3 ##number of Independet variables  
leverage = hatvalues(model1)  
cutleverage = (2\*k+2) / nrow(Lab4)  
cutleverage

## [1] 0.002649007

#2.  
badleverage = as.numeric(leverage > cutleverage)  
table(badleverage)

## badleverage  
## 0 1   
## 2773 247

b. Cook's:  
 i. What is your Cook's cut off score?  
 ii. How many Cook's outliers did you have?

#1.  
cooks = cooks.distance(model1)  
cutcooks= 4 / (nrow(Lab4) - k - 1)  
cutcooks

## [1] 0.00132626

#2.  
badcooks = as.numeric(cooks > cutcooks)  
table(badcooks)

## badcooks  
## 0 1   
## 2871 149

c. Mahalanobis:  
 i. What is your Mahalanobis df?  
 ii. What is your Mahalanobis cut off score?  
 iii. How many outliers did you have for Mahalanobis?

#1.  
mahal = mahalanobis(Lab4,   
 colMeans(Lab4),   
 cov(Lab4))  
  
#2.   
cutmahal = qchisq(1-.001, ncol(Lab4))  
cutmahal

## [1] 18.46683

#3.  
badmahal = as.numeric(mahal > cutmahal)   
table(badmahal)

## badmahal  
## 0 1   
## 2938 82

d. Overall:  
 i. How many total outliers did you have across all variables?  
 ii. Delete them!

#1.  
totalout = badmahal + badleverage + badcooks  
table(totalout)

## totalout  
## 0 1 2 3   
## 2690 218 76 36

#2.   
noout = subset(Lab4,totalout < 2)

# Hierarchical Regression:

a. In step 1, control for verbal ability of the participant predicting reading scores.   
b. In step 2, test if working memory is related to reading scores.  
c. In step 3, test if math score is related to reading scores.  
d. Include the summaries of each step, along with the ANOVA of the change between each step.

#a.  
Step1 = lm(reading ~ verbal, data = noout)  
summary(Step1)

##   
## Call:  
## lm(formula = reading ~ verbal, data = noout)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.1426 -0.7838 -0.3245 0.5361 4.3474   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.652610 0.029476 56.066 <2e-16 \*\*\*  
## verbal -0.009769 0.004728 -2.066 0.0389 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.015 on 2906 degrees of freedom  
## Multiple R-squared: 0.001467, Adjusted R-squared: 0.001124   
## F-statistic: 4.27 on 1 and 2906 DF, p-value: 0.03889

#b.  
Step2 = lm(reading ~ verbal+work\_mem , data = noout)  
summary(Step2)

##   
## Call:  
## lm(formula = reading ~ verbal + work\_mem, data = noout)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.5066 -0.7625 -0.3185 0.5052 4.2851   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.446605 0.039112 36.986 < 2e-16 \*\*\*  
## verbal -0.008749 0.004680 -1.869 0.0617 .   
## work\_mem 0.008699 0.001100 7.906 3.74e-15 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.004 on 2905 degrees of freedom  
## Multiple R-squared: 0.0225, Adjusted R-squared: 0.02183   
## F-statistic: 33.43 on 2 and 2905 DF, p-value: 4.412e-15

#c.  
Step3 = lm(reading ~ verbal+work\_mem+math , data = noout)  
summary(Step3)

##   
## Call:  
## lm(formula = reading ~ verbal + work\_mem + math, data = noout)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.4339 -0.7505 -0.3246 0.5056 4.3578   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.3984067 0.0399840 34.974 < 2e-16 \*\*\*  
## verbal -0.0078972 0.0046615 -1.694 0.0903 .   
## work\_mem 0.0011017 0.0018063 0.610 0.5420   
## math 0.0048806 0.0009227 5.289 1.32e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.9997 on 2904 degrees of freedom  
## Multiple R-squared: 0.03183, Adjusted R-squared: 0.03083   
## F-statistic: 31.82 on 3 and 2904 DF, p-value: < 2.2e-16

#d.  
anova(Step1, Step2,Step3)

## Analysis of Variance Table  
##   
## Model 1: reading ~ verbal  
## Model 2: reading ~ verbal + work\_mem  
## Model 3: reading ~ verbal + work\_mem + math  
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 2906 2993.1   
## 2 2905 2930.0 1 63.046 63.087 2.802e-15 \*\*\*  
## 3 2904 2902.1 1 27.959 27.978 1.318e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Moderation:

a. Examine the interaction between verbal and math scores predicting reading scores.  
b. Include the simple slopes for low, average, and high math levels (split on math) for verbal predicting reading.   
c. Include a graph of the interaction.

#a.  
  
noout$zverbal = scale(noout$verbal, scale = F)  
  
noout$zmath = scale(noout$math, scale = F)  
  
modmodel = lm(reading ~ zverbal\*zmath, data = noout)  
summary(modmodel)

##   
## Call:  
## lm(formula = reading ~ zverbal \* zmath, data = noout)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.5281 -0.7434 -0.3130 0.5061 4.4264   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.6034327 0.0185319 86.523 < 2e-16 \*\*\*  
## zverbal -0.0087881 0.0046667 -1.883 0.05978 .   
## zmath 0.0053216 0.0005588 9.524 < 2e-16 \*\*\*  
## zverbal:zmath -0.0004064 0.0001446 -2.810 0.00498 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.9984 on 2904 degrees of freedom  
## Multiple R-squared: 0.03433, Adjusted R-squared: 0.03333   
## F-statistic: 34.41 on 3 and 2904 DF, p-value: < 2.2e-16

#b.  
noout$zmathlow = noout$zmath + sd(noout$zmath)  
noout$zmathhigh = noout$zmath - sd(noout$zmath)  
summary(noout)

## reading verbal math work\_mem   
## Min. :0.510 Min. : 0.000 Min. : 0.00 Min. :-11.00   
## 1st Qu.:0.820 1st Qu.: 0.000 1st Qu.: 21.00 1st Qu.: 11.00   
## Median :1.280 Median : 5.000 Median : 36.00 Median : 18.00   
## Mean :1.606 Mean : 4.798 Mean : 45.03 Mean : 23.12   
## 3rd Qu.:2.143 3rd Qu.: 8.000 3rd Qu.: 61.00 3rd Qu.: 30.00   
## Max. :6.000 Max. :17.000 Max. :197.00 Max. :107.00   
## zverbal.V1 zmath.V1 zmathlow.V1   
## Min. :-4.798487 Min. :-45.02579 Min. :-11.85550   
## 1st Qu.:-4.798487 1st Qu.:-24.02579 1st Qu.: 9.14450   
## Median : 0.201513 Median : -9.02579 Median : 24.14450   
## Mean : 0.000000 Mean : 0.00000 Mean : 33.17029   
## 3rd Qu.: 3.201513 3rd Qu.: 15.97421 3rd Qu.: 49.14450   
## Max. :12.201513 Max. :151.97421 Max. :185.14450   
## zmathhigh.V1   
## Min. :-78.19608   
## 1st Qu.:-57.19608   
## Median :-42.19608   
## Mean :-33.17029   
## 3rd Qu.:-17.19608   
## Max. :118.80392

modmodellow = lm(reading ~ zverbal\*zmathlow, data = noout)  
modmodelhigh = lm(reading ~ zverbal\*zmathhigh, data = noout)  
  
# Low, Average , High Slope  
summary(modmodellow) #low slope

##   
## Call:  
## lm(formula = reading ~ zverbal \* zmathlow, data = noout)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.5281 -0.7434 -0.3130 0.5061 4.4264   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.4269127 0.0262074 54.447 < 2e-16 \*\*\*  
## zverbal 0.0046930 0.0064532 0.727 0.46714   
## zmathlow 0.0053216 0.0005588 9.524 < 2e-16 \*\*\*  
## zverbal:zmathlow -0.0004064 0.0001446 -2.810 0.00498 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.9984 on 2904 degrees of freedom  
## Multiple R-squared: 0.03433, Adjusted R-squared: 0.03333   
## F-statistic: 34.41 on 3 and 2904 DF, p-value: < 2.2e-16

summary(modmodel) #average slope

##   
## Call:  
## lm(formula = reading ~ zverbal \* zmath, data = noout)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.5281 -0.7434 -0.3130 0.5061 4.4264   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.6034327 0.0185319 86.523 < 2e-16 \*\*\*  
## zverbal -0.0087881 0.0046667 -1.883 0.05978 .   
## zmath 0.0053216 0.0005588 9.524 < 2e-16 \*\*\*  
## zverbal:zmath -0.0004064 0.0001446 -2.810 0.00498 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.9984 on 2904 degrees of freedom  
## Multiple R-squared: 0.03433, Adjusted R-squared: 0.03333   
## F-statistic: 34.41 on 3 and 2904 DF, p-value: < 2.2e-16

summary(modmodelhigh) #high slope

##   
## Call:  
## lm(formula = reading ~ zverbal \* zmathhigh, data = noout)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.5281 -0.7434 -0.3130 0.5061 4.4264   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.7799527 0.0262121 67.906 < 2e-16 \*\*\*  
## zverbal -0.0222693 0.0069236 -3.216 0.00131 \*\*   
## zmathhigh 0.0053216 0.0005588 9.524 < 2e-16 \*\*\*  
## zverbal:zmathhigh -0.0004064 0.0001446 -2.810 0.00498 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.9984 on 2904 degrees of freedom  
## Multiple R-squared: 0.03433, Adjusted R-squared: 0.03333   
## F-statistic: 34.41 on 3 and 2904 DF, p-value: < 2.2e-16

#c.  
library(ggplot2)  
  
cleanup = theme(panel.grid.major = element\_blank(),   
 panel.grid.minor = element\_blank(),   
 panel.background = element\_blank(),   
 axis.line.x = element\_line(color = "black"),  
 axis.line.y = element\_line(color = "black"),  
 legend.key = element\_rect(fill = "white"),  
 text = element\_text(size = 15))  
  
#use the z score variables  
modgraph = ggplot(noout, aes(zverbal, reading))  
modgraph +   
 xlab("Verbal Score") +   
 geom\_point(color = "gray") +  
   
 ##this part here assumes you named the models the same as above  
 ##and you did X\*M where M is the low medium high variable   
 ##change the labels for the slopes in BOTH places   
   
 geom\_abline(aes(intercept = modmodellow$coefficients[1],  
 slope = modmodellow$coefficients[2],   
 linetype = "-1SD math"), size = 1) +  
 geom\_abline(aes(intercept = modmodel$coefficients[1],  
 slope = modmodel$coefficients[2],   
 linetype = "Average math"), size = 1) +  
 geom\_abline(aes(intercept = modmodelhigh$coefficients[1],  
 slope = modmodelhigh$coefficients[2],   
 linetype = "+1SD math"), size = 1) +  
 scale\_linetype\_manual(values = c("dotted", "dashed", "solid"),  
 breaks = c("-1SD math", "Average math", "+1SD math"),  
 name = "Simple Slope") +  
 cleanup

