MaetzigEtAl2017 ICCM summary/analysis for SR/OR with reflexives

Paul Mätzig

5/11/2017

```
## Warning: NAs introduced by coercion
```

```r
model_data_refl_acc_SR <- subset(model_data, model_data$condition == "SR-REFL")
model_data_refl_acc_OR <- subset(model_data, model_data$condition == "OR-REFL")

head(model_data_refl_acc_OR)
```

```
##       X simset condition   GA  DAT  ANS sentCount corrCountEmbV
## 2    1      1   OR-REFL 0.25 0.05 0.15      2000           989
## 4    3      2   OR-REFL 0.50 0.05 0.15      2000           991
## 6    5      3   OR-REFL 0.75 0.05 0.15      2000           991
## 8    7      4   OR-REFL 1.00 0.05 0.15      2000           994
## 10   9      5   OR-REFL 0.25 0.06 0.15      2000           991
## 12  11      6   OR-REFL 0.50 0.06 0.15      2000           991
##     corrCountRefl acc_embV acc_refl
## 2             983    0.989    0.983
## 4             986    0.991    0.986
## 6             989    0.991    0.989
## 8             993    0.994    0.993
## 10            986    0.991    0.986
## 12            986    0.991    0.986
```

```r
head(model_data_refl_acc_SR)
```

```
##       X simset condition   GA  DAT  ANS sentCount corrCountEmbV
## 1    0      1   SR-REFL 0.25 0.05 0.15      2000           992
## 3    2      2   SR-REFL 0.50 0.05 0.15      2000           987
## 5    4      3   SR-REFL 0.75 0.05 0.15      2000           991
## 7    6      4   SR-REFL 1.00 0.05 0.15      2000           994
## 9    8      5   SR-REFL 0.25 0.06 0.15      2000           989
## 11  10      6   SR-REFL 0.50 0.06 0.15      2000           986
##     corrCountRefl acc_embV acc_refl
## 1             991    0.992    0.991
## 3             987    0.987    0.987
## 5             991    0.991    0.991
## 7             993    0.994    0.993
## 9             988    0.989    0.988
## 11            984    0.986    0.984
```
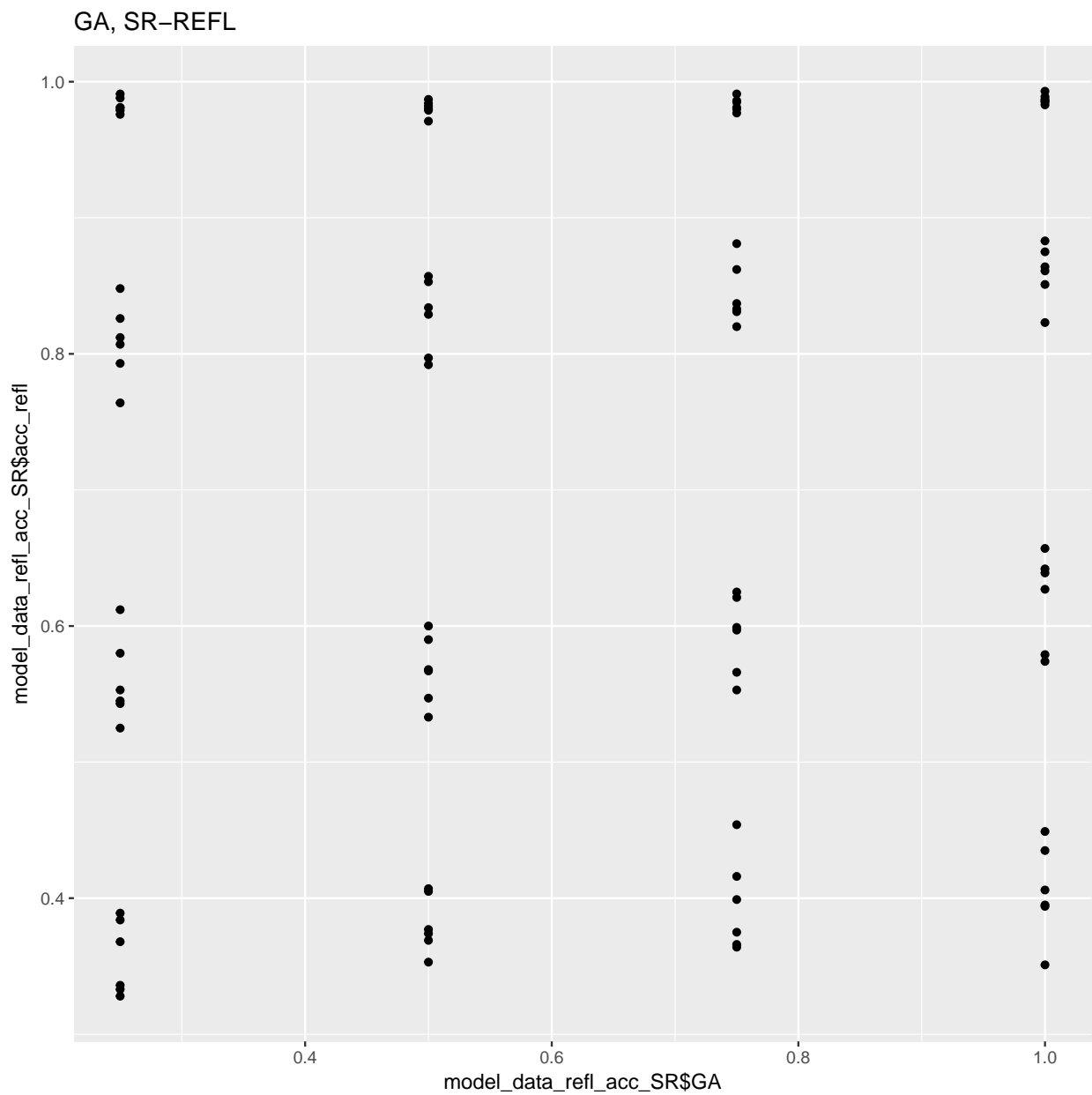
```r
plot(model_data_refl_acc_SR$acc_refl ~ model_data_refl_acc_SR$simset, main="SR-REFL")
```
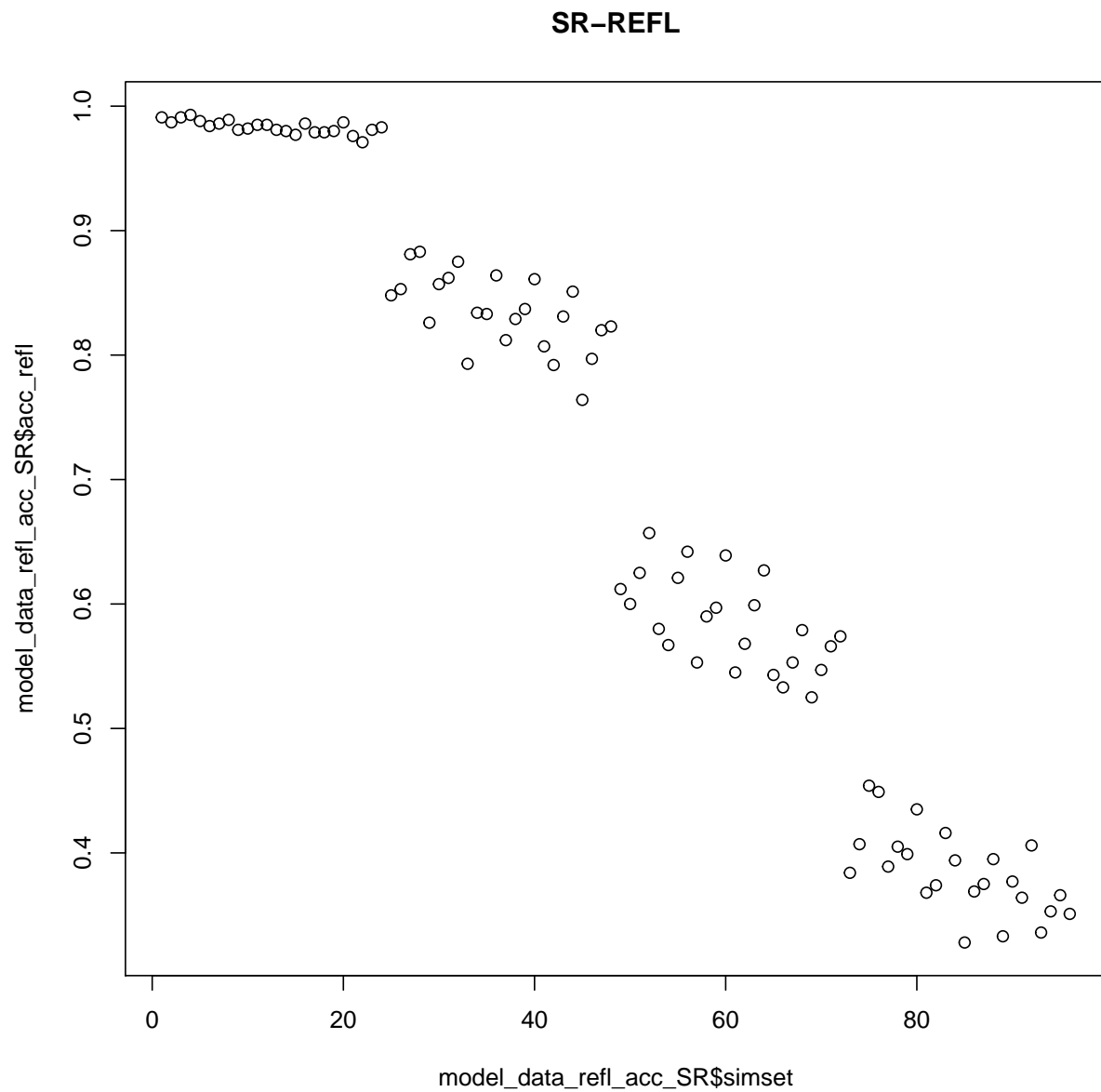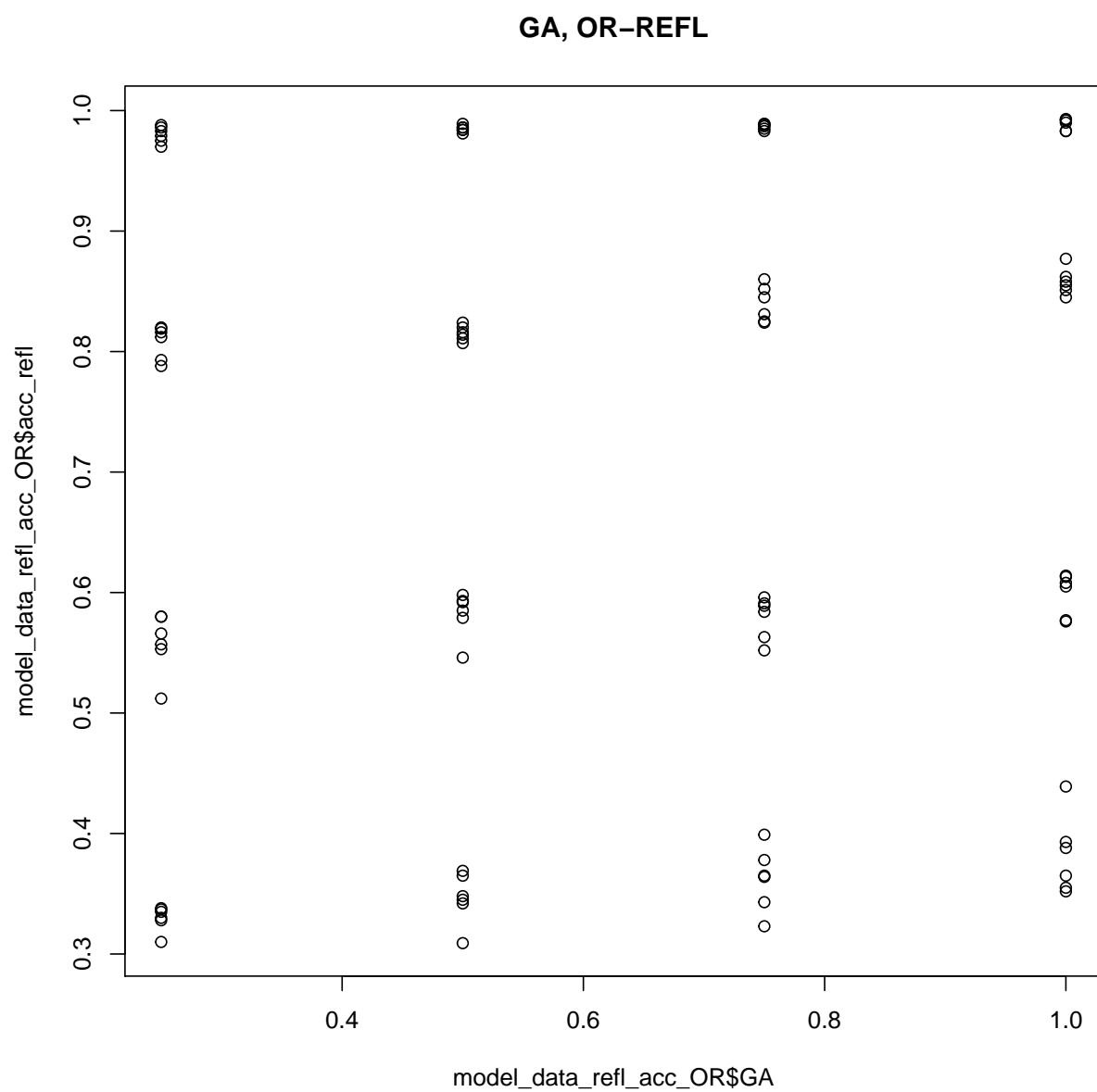
SR−REFL

```r
plot(model_data_refl_acc_OR$acc_refl ~ model_data_refl_acc_OR$simset, main="OR-REFL")
```
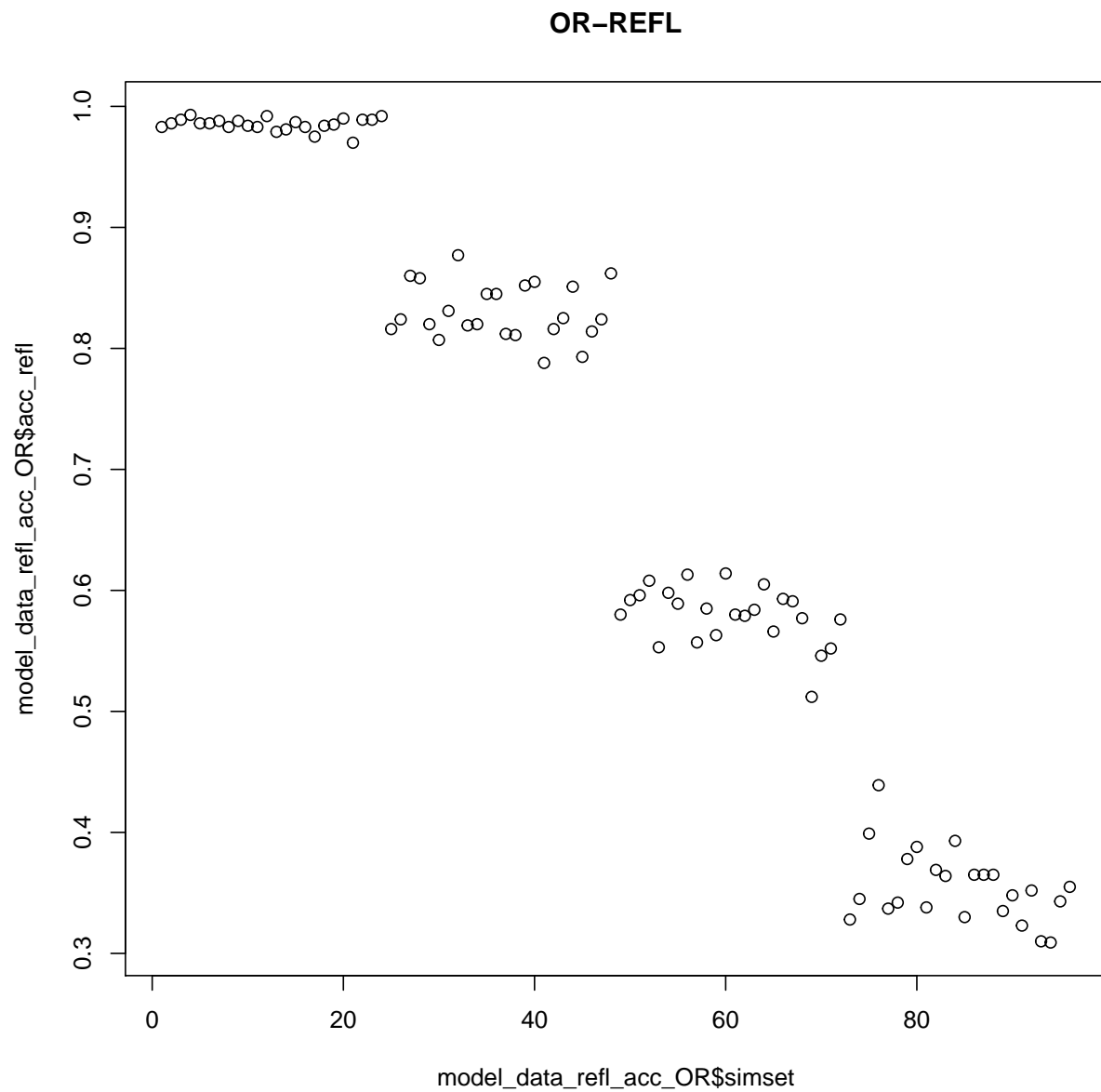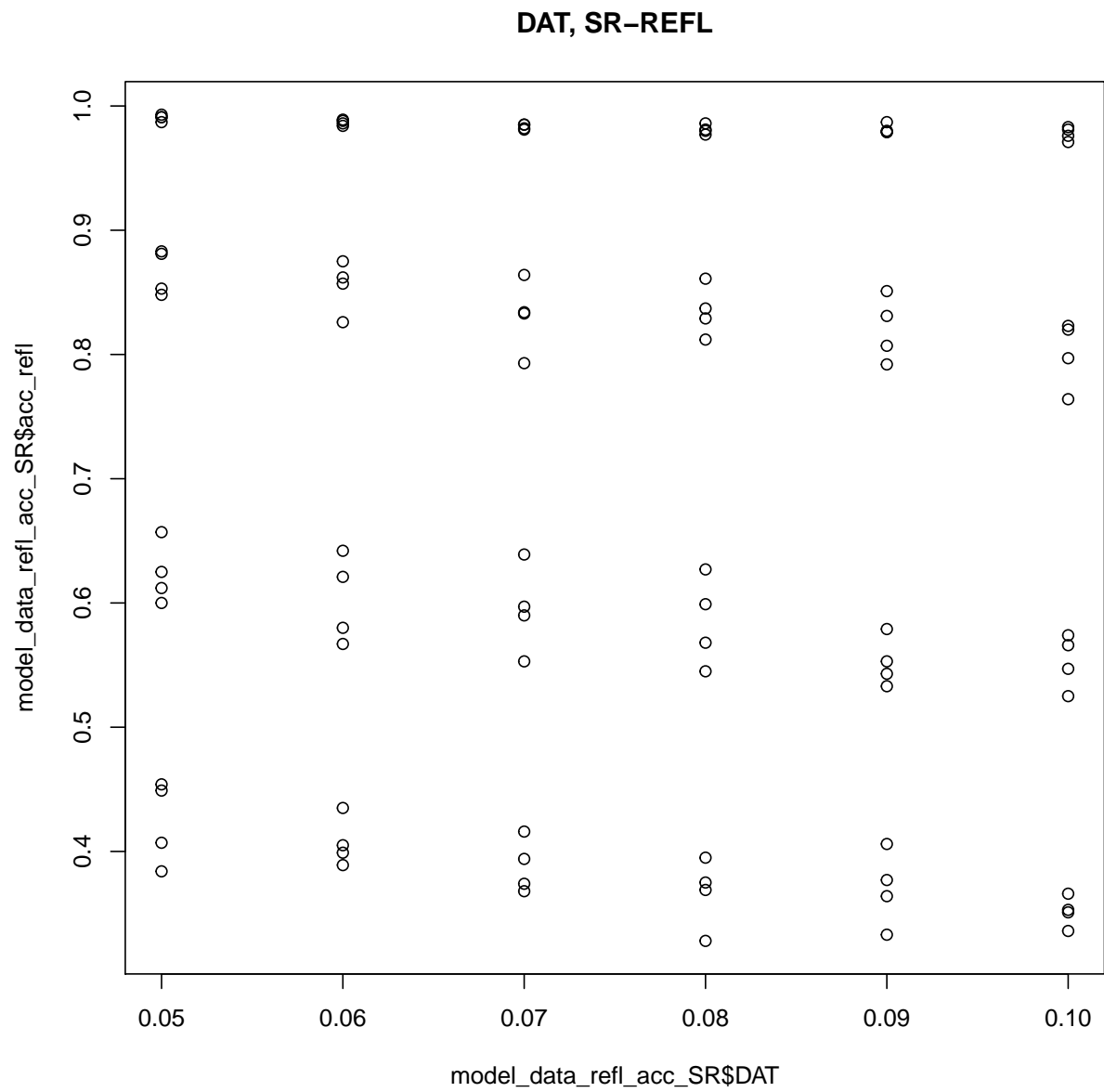
OR−REFL


```
qplot(model_data_refl_acc_SR$GA, model_data_refl_acc_SR$acc_refl, main="GA, SR-REFL")
```
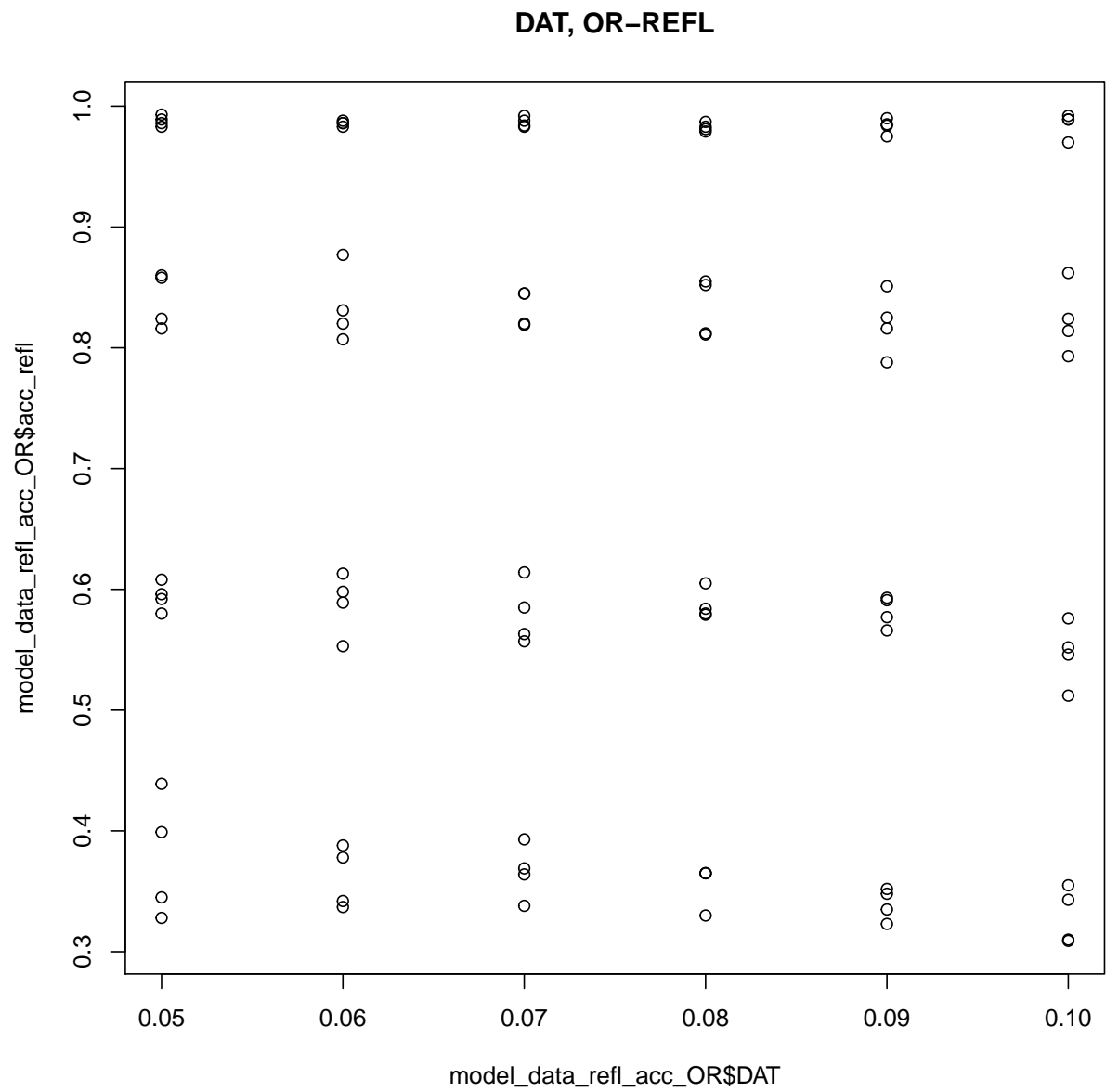
GA, SR−REFL

```
plot(model_data_refl_acc_OR$GA, model_data_refl_acc_OR$acc_refl, main="GA, OR−REFL")
```
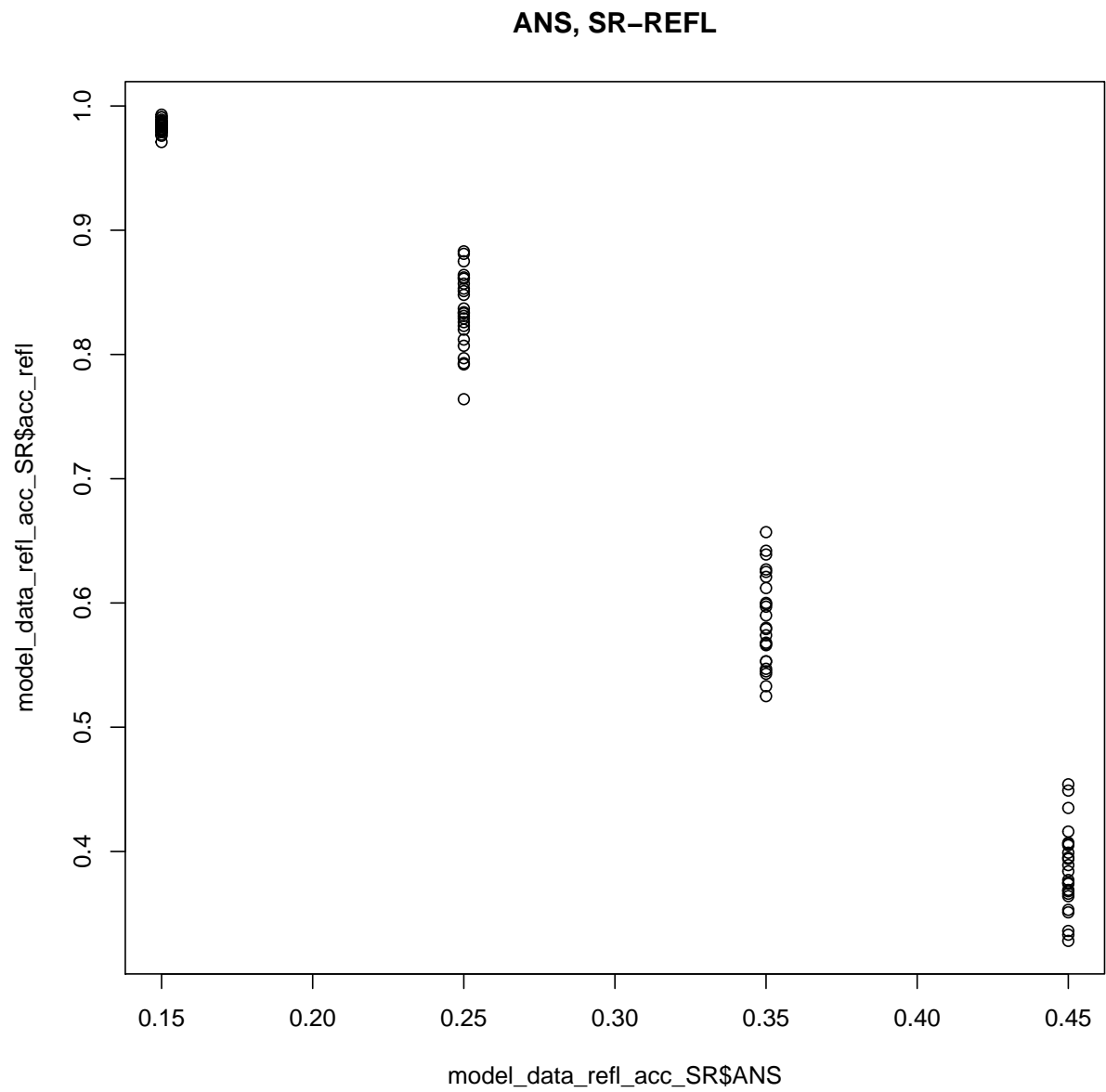
GA, OR–REFL


```
plot(model_data_refl_acc_SR$DAT, model_data_refl_acc_SR$acc_refl, main="DAT, SR-REFL")
```

DAT, SR−REFL


```
plot(model_data_refl_acc_OR$DAT, model_data_refl_acc_OR$acc_refl, main="DAT, OR-REFL")
```
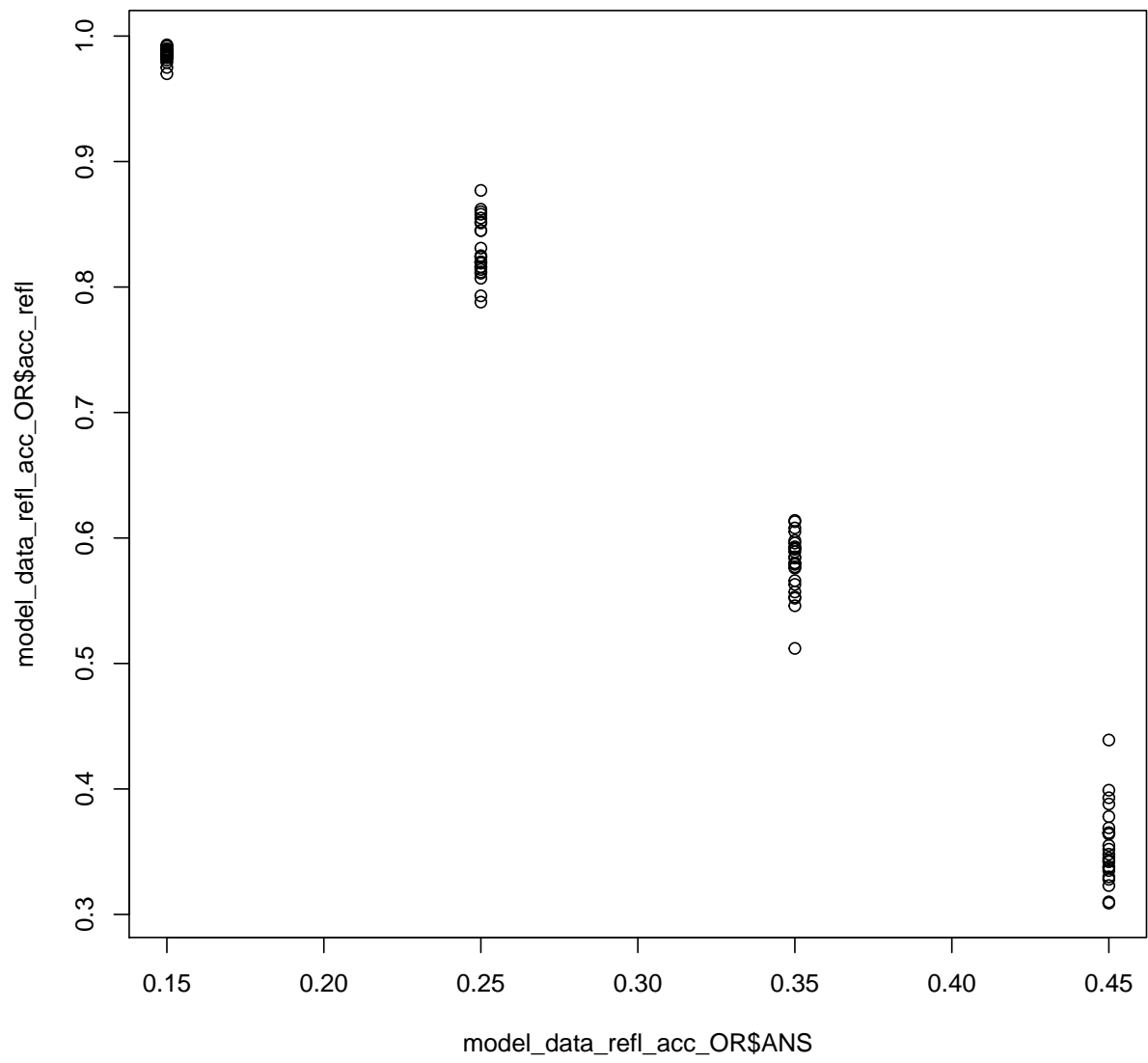
DAT, OR–REFL


```
plot(model_data_refl_acc_SR$ANS, model_data_refl_acc_SR$acc_refl, main="ANS, SR-REFL")
```
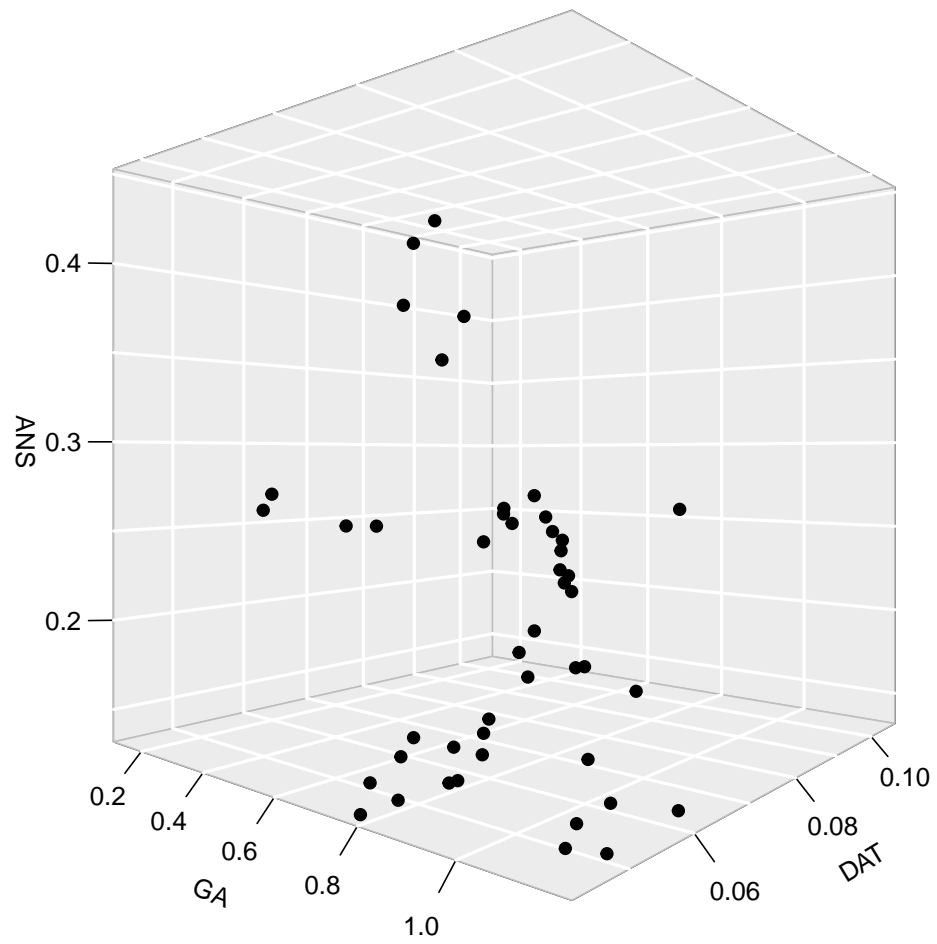
ANS, SR−REFL


```r
plot(model_data_refl_acc_OR$ANS, model_data_refl_acc_OR$acc_refl, main="ANS, OR−REFL")
```
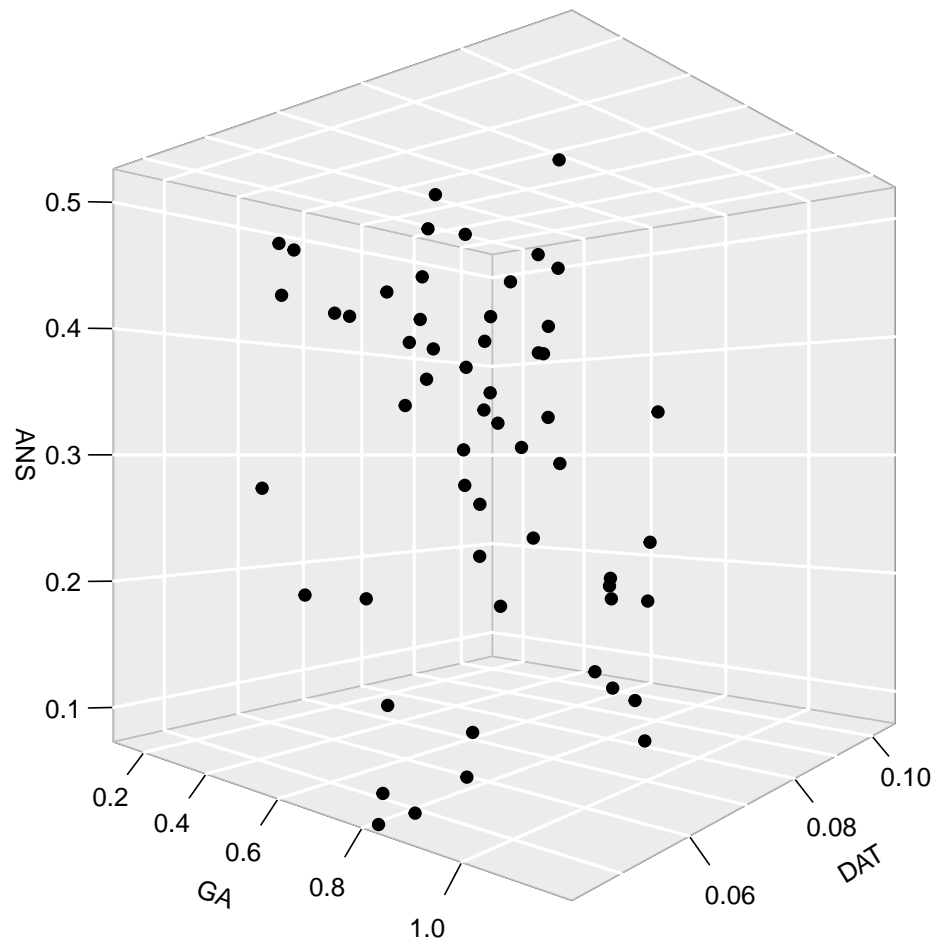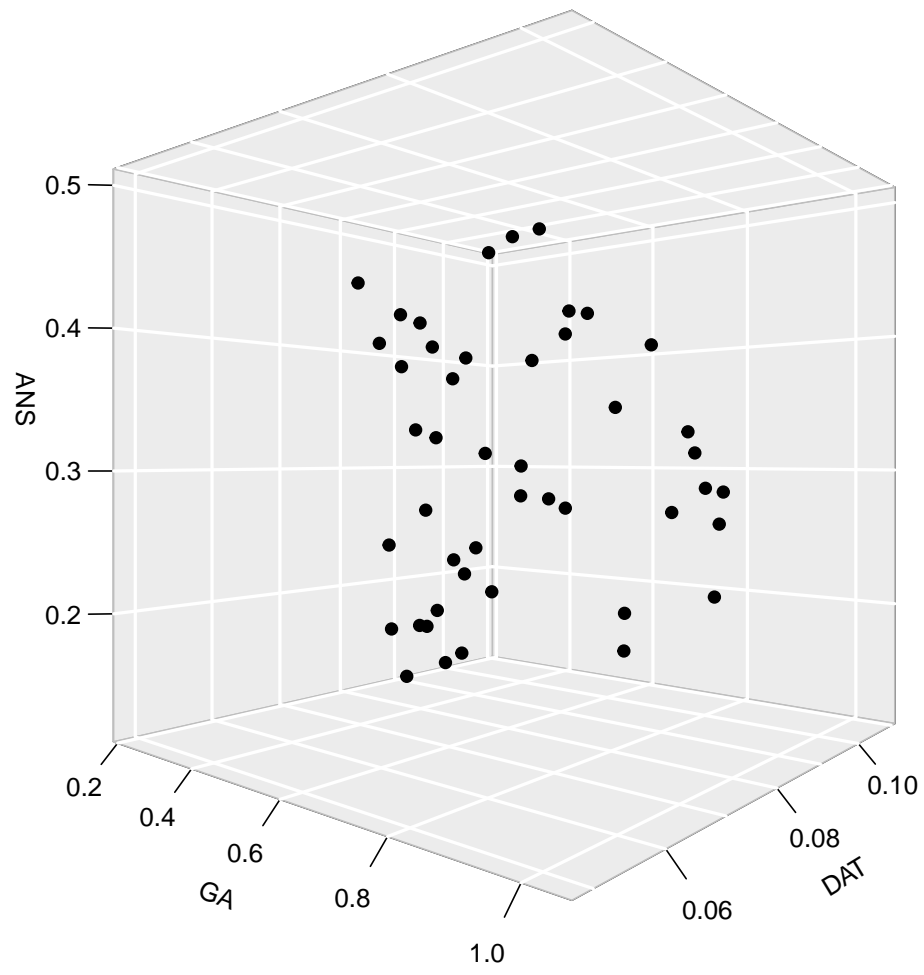
ANS, OR–REFL

Controls, SR

IWA, SR

Controls, OR

IWA, OR

Subject relatives   **Subject relatives**   **Subject relatives**
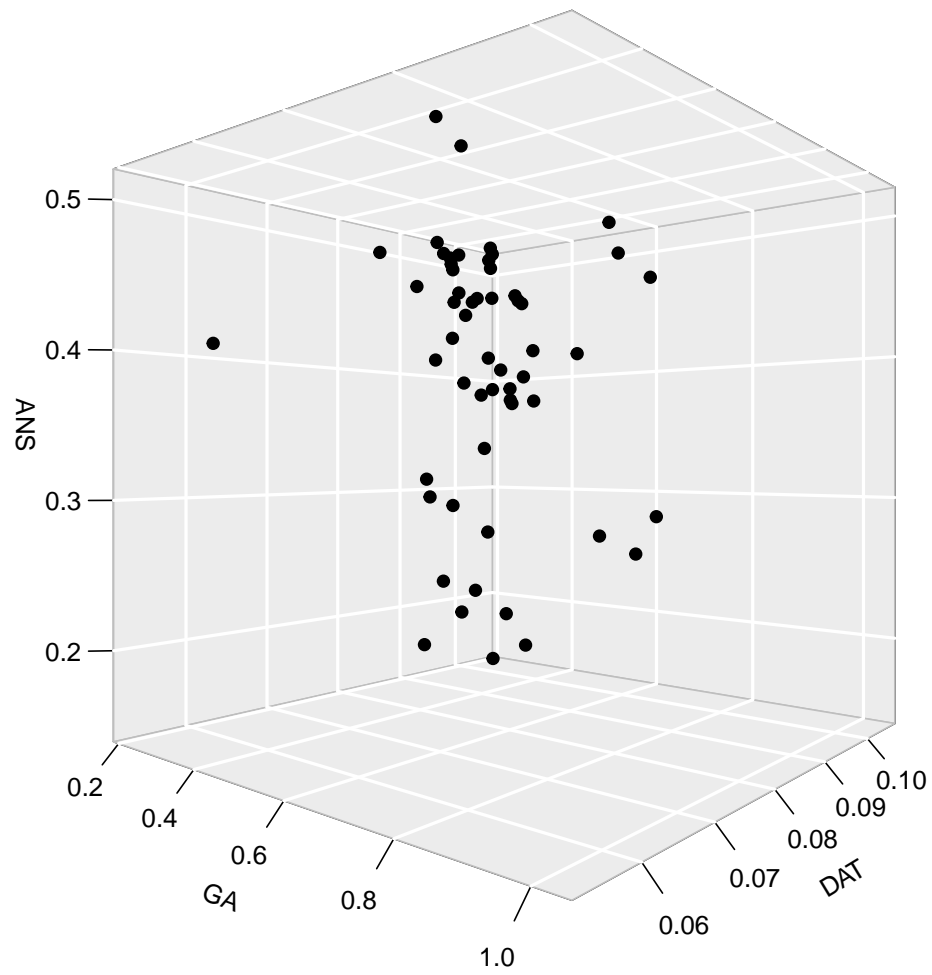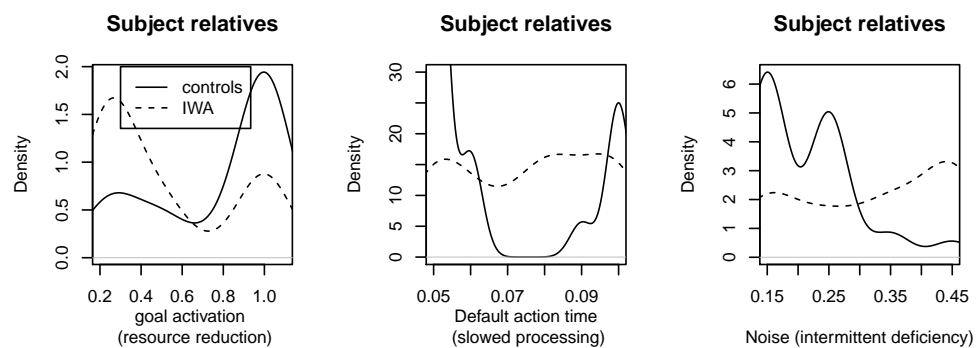
```
op<-par(
    mar=c(2,4.5,4,2),
    mfrow=c(1,3),pty="s",
    oma = c(2,10,2,2)
    )

#plot(density(results_SR_iwa$GA),xlab="GA",main="IWA, SR")
#plot(density(results_SR_iwa$DAT),xlab="DAT",main="IWA, SR")
#plot(density(results_SR_iwa$ANS),xlab="ANS",main="IWA, SR")

#op<-par(mfrow=c(2,3),pty="s")
plot(density(results_OR_controls$GA),xlab="goal activation \n (resource reduction)",main="Object relativ
lines(density(results_OR_iwa$GA),xlab="GA",lty=2)
legend(x=.3,y=3,legend=c("controls","patients"),
```

```
        lty=1:2)

plot(density(results_OR_controls$DAT),xlab="Default action time \n (slowed processing)",main="Object rel
lines(density(results_OR_iwa$DAT),xlab="DAT",lty=2)

plot(density(results_OR_controls$ANS),xlab="Noise (intermittent deficiency)",main="Object relatives",xl
lines(density(results_OR_iwa$ANS),xlab="ANS",lty=2)
```
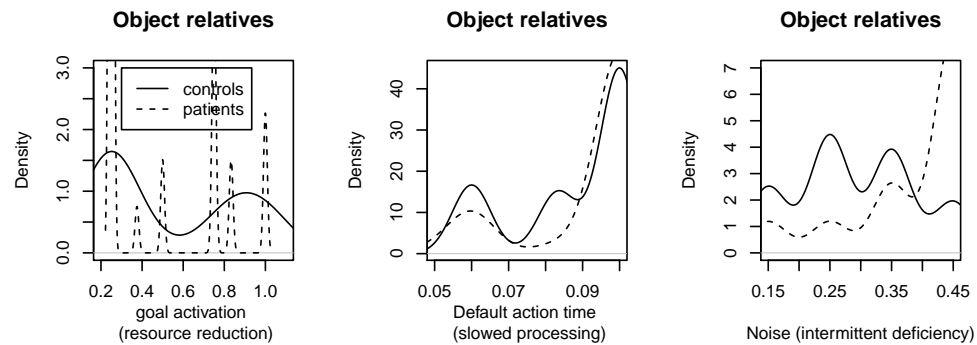


```
#plot(density(results_OR_iwa$GA),xlab="GA",main="IWA, OR")
#plot(density(results_OR_iwa$DAT),xlab="DAT",main="IWA, OR")
#plot(density(results_OR_iwa$ANS),xlab="ANS",main="IWA, OR")
```

Hierarchical clustering

[pmaetzig, 2017-05-28] I just copied the steps from the previous ICCM analysis file for now, may need to change methods in hierarchical clustering.

I don't see much of a difference between the two groups in the clustering.

In SRs I chose centroid as the method because that gives the best discrimination ability between controls and iwa.

```
results_SR_controls$type <- "control"
results_SR_iwa$type <- "iwa"
results_OR_controls$type <- "control"
results_OR_iwa$type <- "iwa"
```

Hierarchical clustering, SR controls
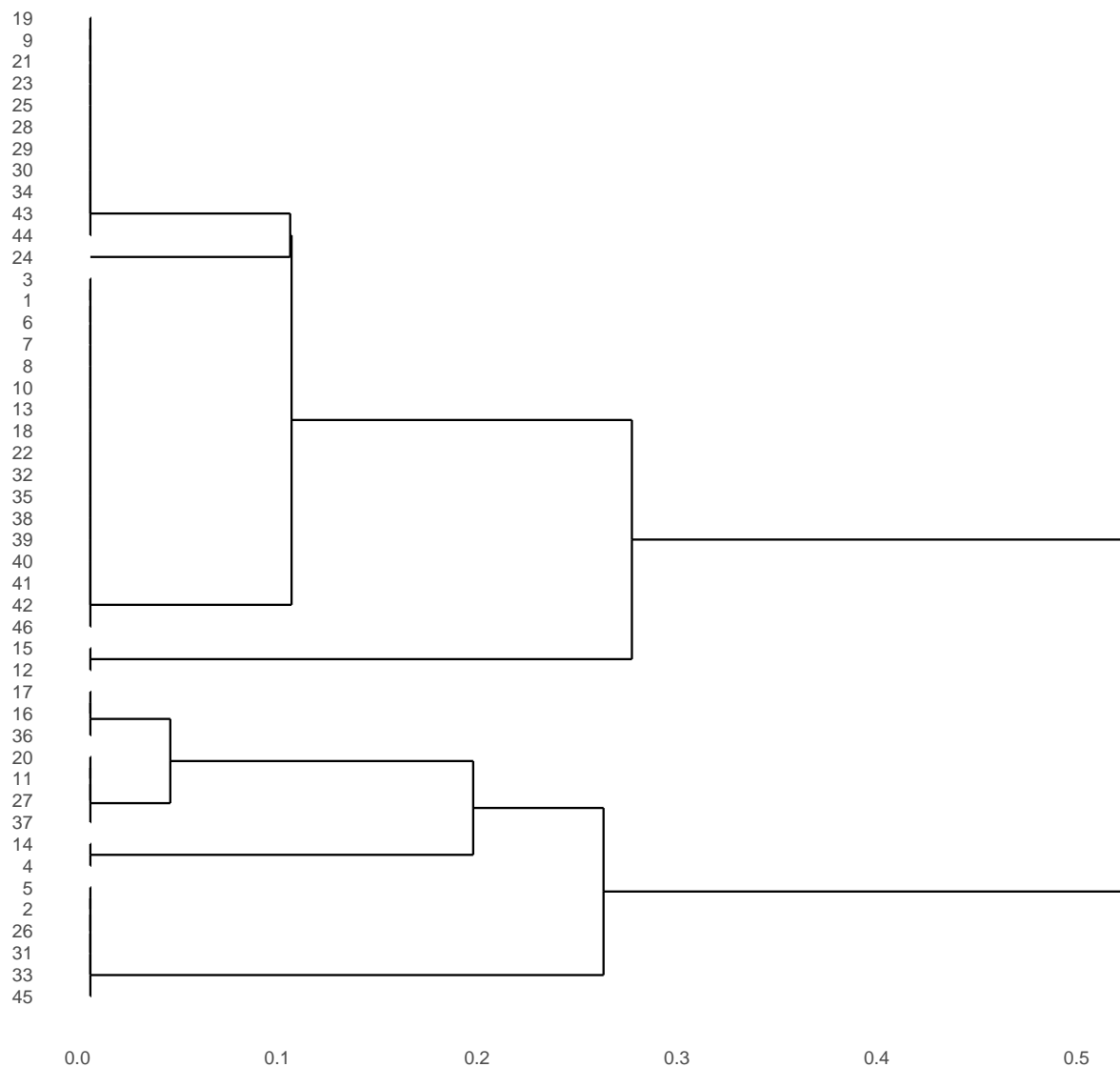
Hierarchical clustering, SR iwa
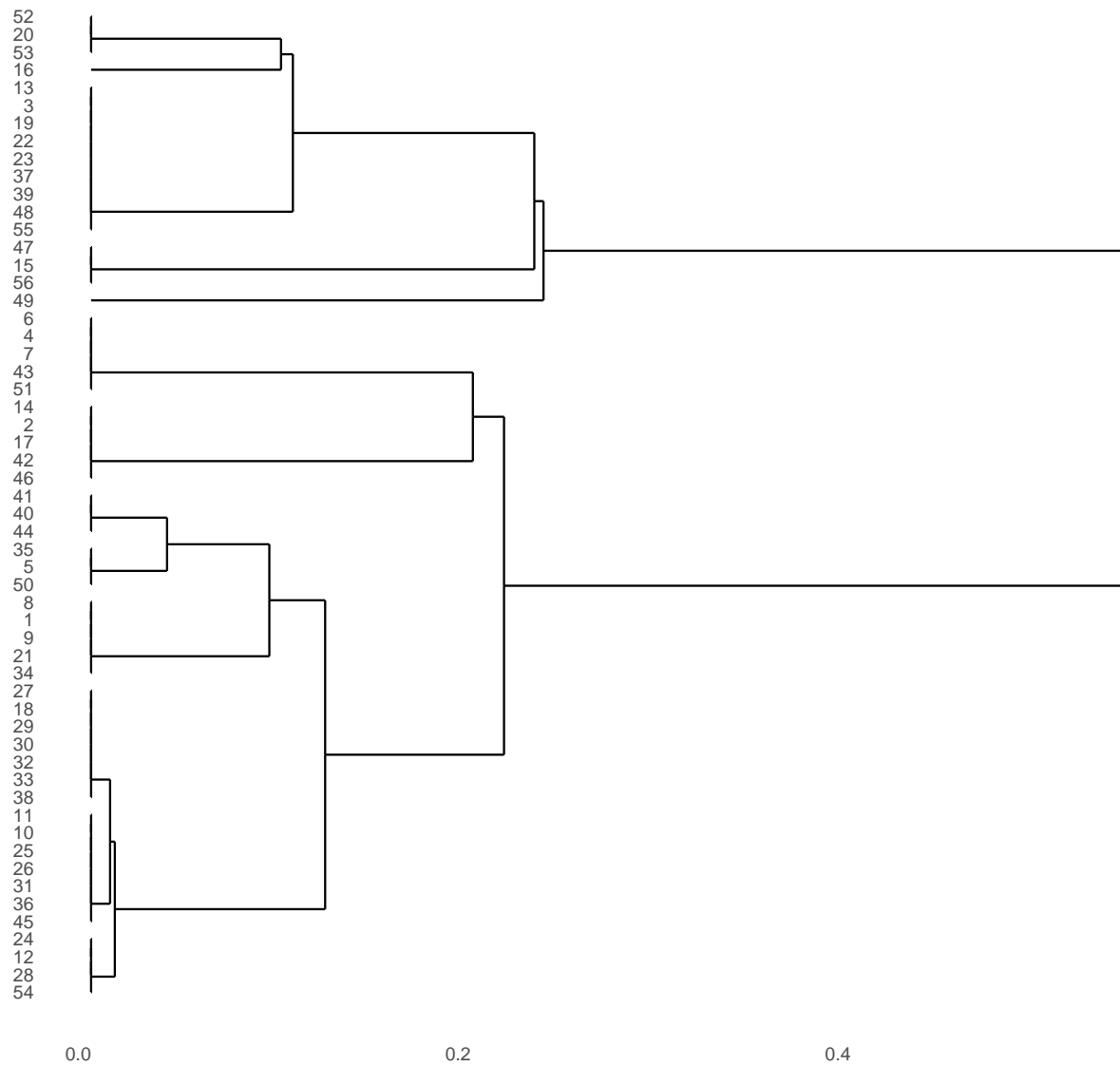
For ORs I chose complete as method because it gives the best discriminative performance between controls and iwa.
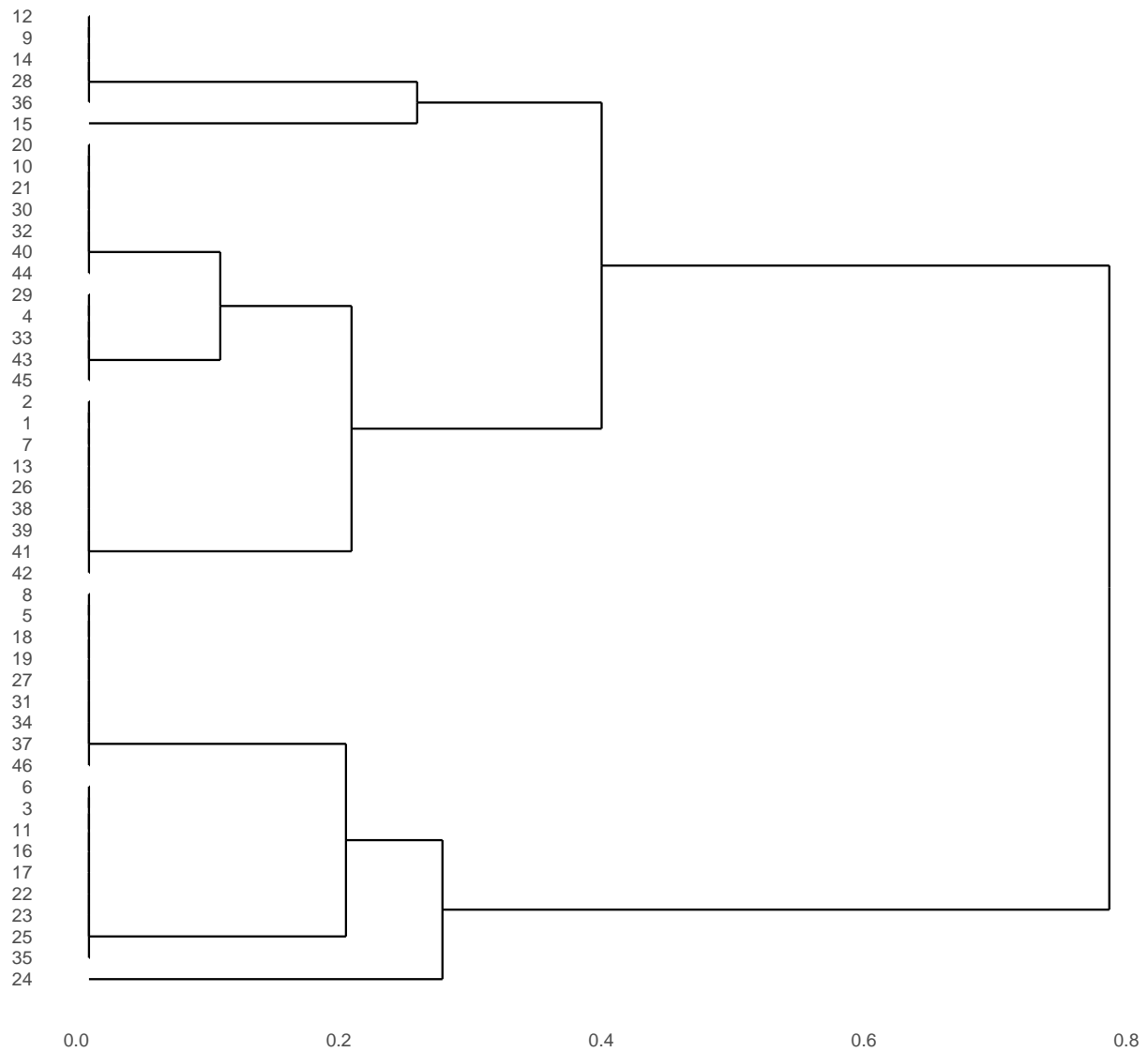
Hierarchical clustering, OR controls

Hierarchical clustering, OR iwa
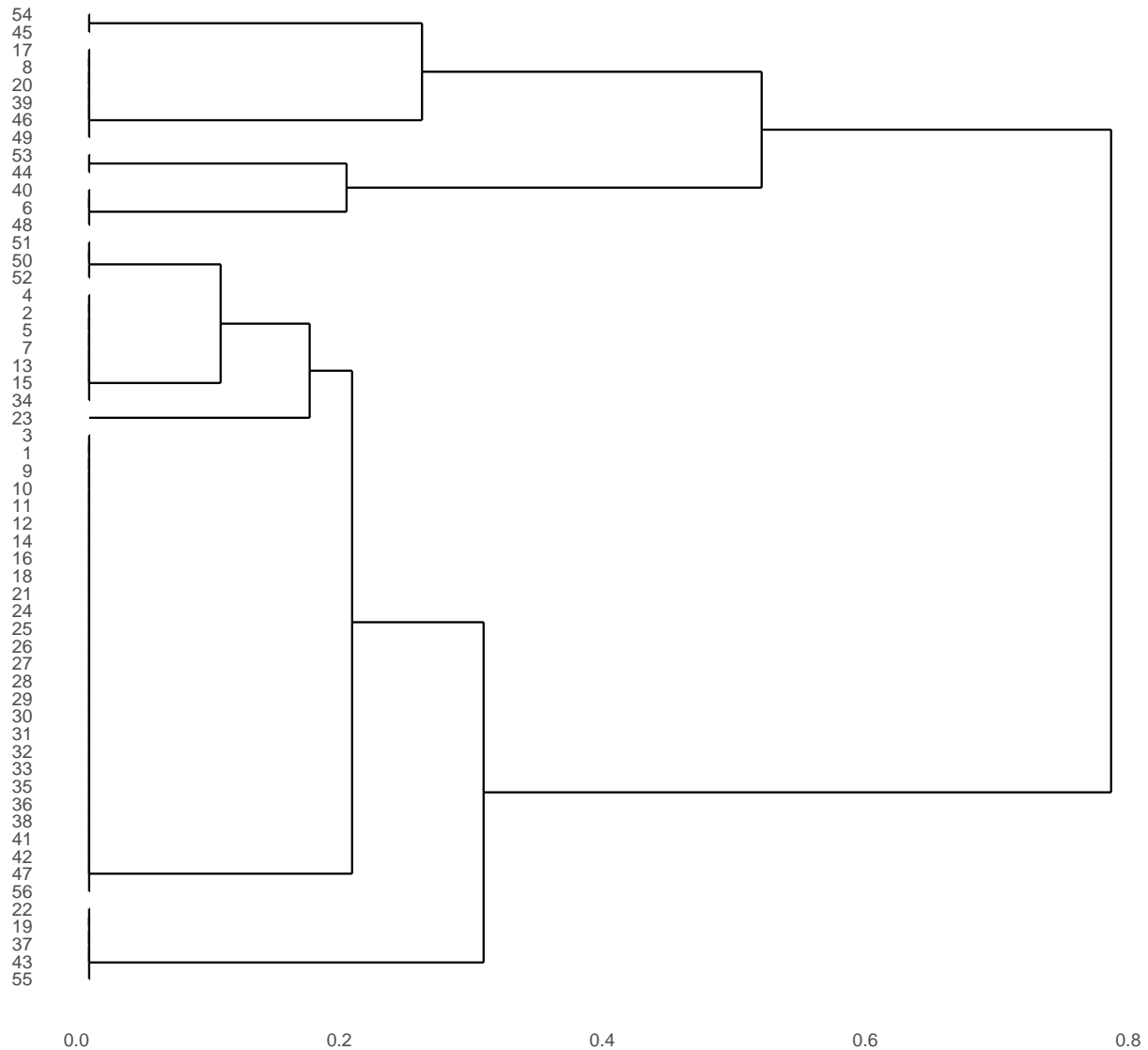
Discrimination ability using hierarchical clustering

The discrimination ability is OK for controls, but below 50% for iwa, suggesting that the three parameters don't give us much ability to discriminate controls and iwa.

```
results_OR<-rbind(results_OR_controls,results_OR_iwa)
results_SR<-rbind(results_SR_controls,results_SR_iwa)

results_SR$subject_full <- paste(results_SR$type, results_SR$subj, sep="-")
results_SR$subject_full <- as.character(results_SR$subject_full)
clusters_SR <- hclust(dist(results_SR[, 1:3]),method="centroid")
plot(clusters_SR, labels=results_SR$subject_full)
```

Cluster Dendrogram
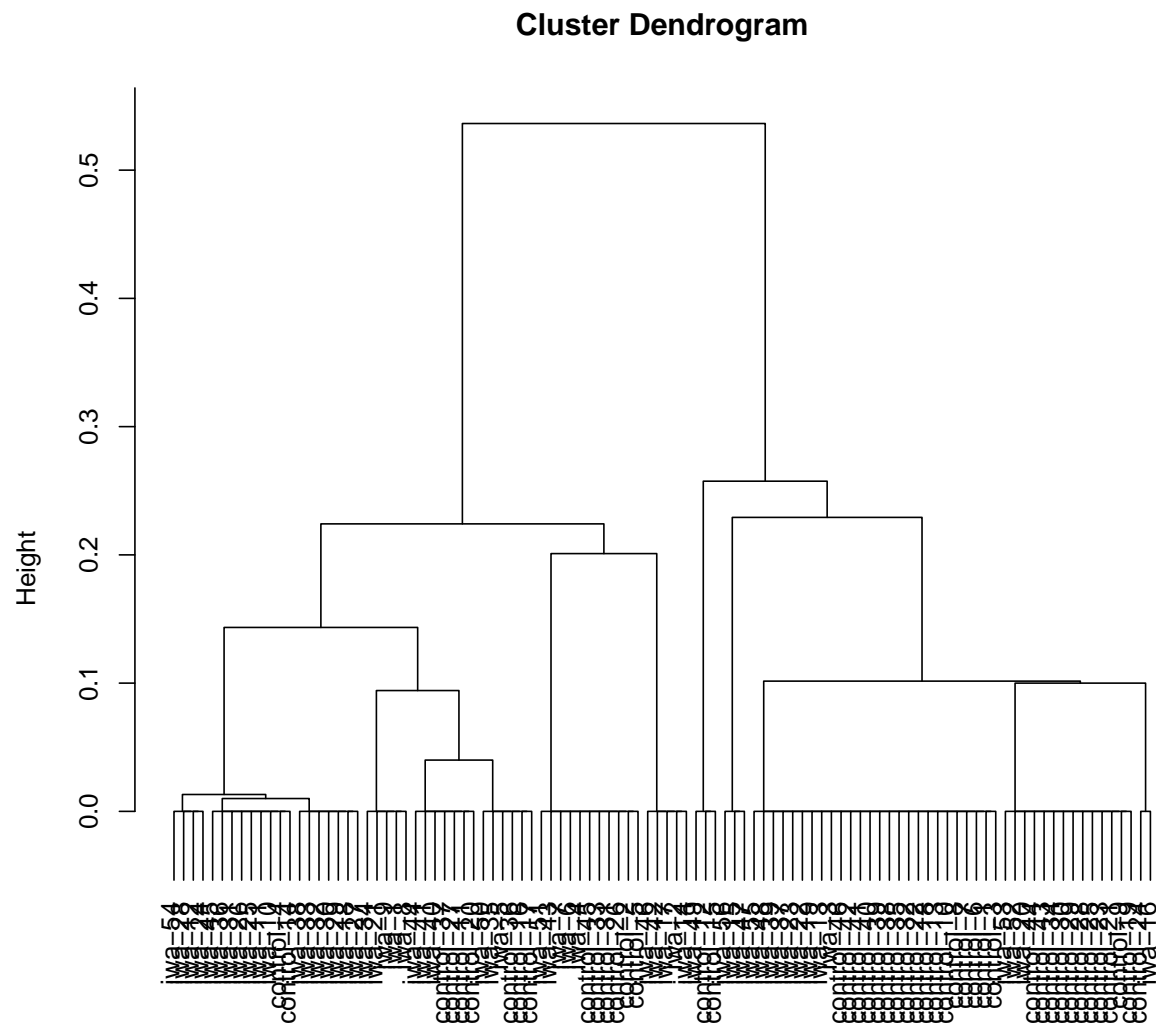
dist(results_SR[, 1:3])
hclust (*, "centroid")

```
clusterCut_SR <- cutree(clusters_SR, 2)
## not able to identify aphasics well
cluster_table_SR <- table(clusterCut_SR, results_SR$type)
print(cluster_table_SR)

##
## clusterCut_SR control iwa
##             1      31  17
##             2      15  39
# SR, controls
(cluster_table_SR[1,1]/(cluster_table_SR[1,1] + cluster_table_SR[2,1]))

## [1] 0.673913
```

```
# SR, iwa
(cluster_table_SR[2,2]/(cluster_table_SR[1,2] + cluster_table_SR[2,2]))
```

```
## [1] 0.6964286
```

```
results_OR$subject_full <- paste(results_OR$type, results_OR$subj, sep="-")
results_OR$subject_full <- as.character(results_OR$subject_full)
clusters_OR <- hclust(dist(results_OR[, 1:3]),method="centroid")
plot(clusters_OR, labels=results_OR$subject_full)
```

Cluster Dendrogram

dist(results_OR[, 1:3])
hclust (*, "centroid")

```
clusterCut_OR <- cutree(clusters_OR, 2)
## so-so discrimination ability for aphasics:
cluster_table_OR <- table(clusterCut_OR, results_OR$type)
print(cluster_table_OR)
```

```
## 
## clusterCut_OR control iwa
##              1     27  45
##              2     19  11
```

```
# OR, controls
(cluster_table_OR[1,1]/(cluster_table_OR[1,1] + cluster_table_OR[2,1]))
```
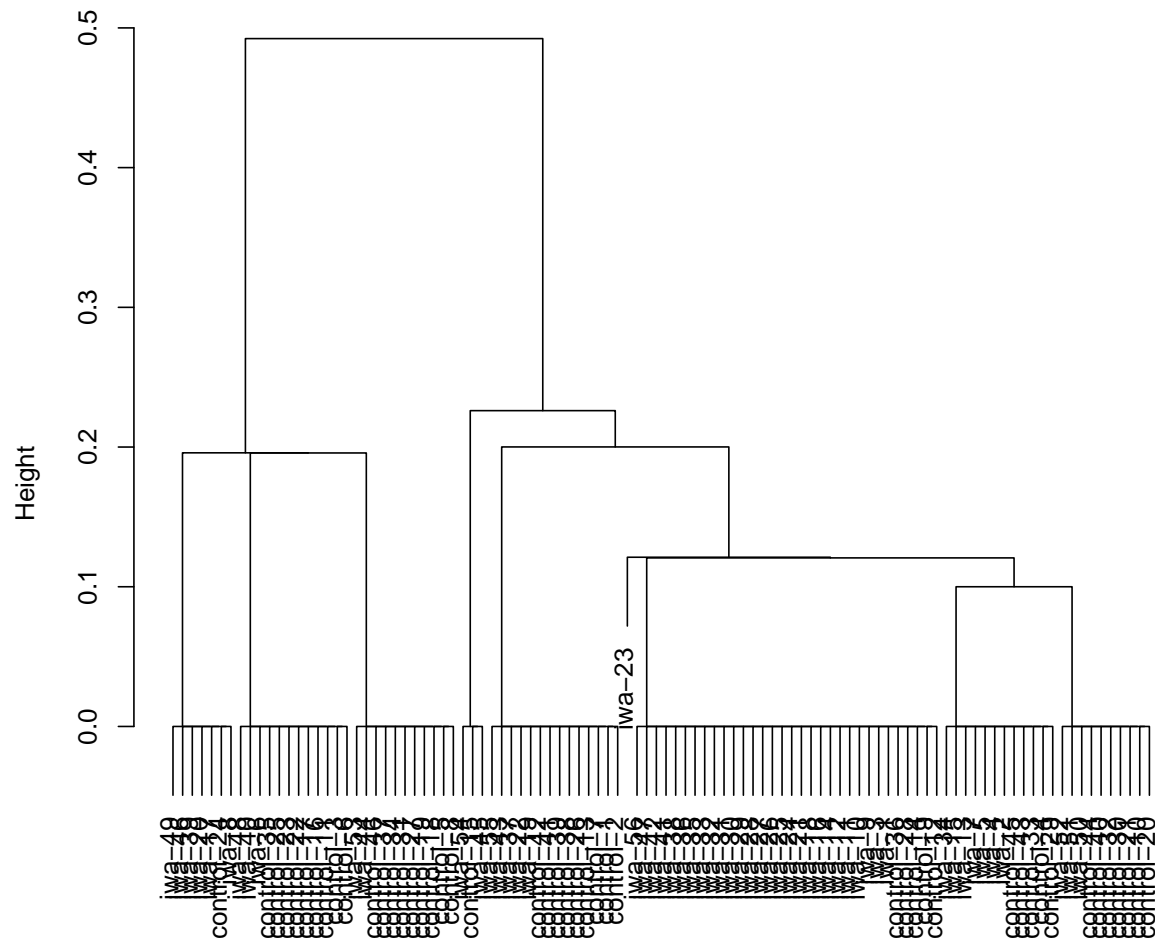
```
## [1] 0.5869565
```

```
# OR, iwa
(cluster_table_OR[2,2]/(cluster_table_OR[1,2] + cluster_table_OR[2,2]))
```

```
## [1] 0.1964286
```

Discrimination ability using k-means

I'm using the `factoextra` package for an automatised workflow for k-means and their implementation of the gap statistic (cf. Hastie, Tibshirani & Friedman, 2008 2nd ed, p. 519f.). Using the `eclust` function with method k-means and a k equal to NULL results in estimation of optimal cluster size using the gap statistic.

CODE CHUNKS TAKEN OUT BECAUSE NOT NEEDED

Distributions of subjects

```
results_SR$RCType<-"SR"
results_OR$RCType<-"OR"
results_all<-rbind(results_SR,results_OR)

results_all$DATnormal<-ifelse(results_all$DAT > .06,
      "elevated","normal")
xtabs(~DATnormal+type+RCType,results_all)
```

```
## , , RCType = OR
## 
##           type
## DATnormal  control iwa
##    elevated      36  46
##    normal        10  10
## 
## , , RCType = SR
## 
##           type
## DATnormal  control iwa
##    elevated      11  36
##    normal        35  20
```

```
results_all$ANSnormal<-ifelse(results_all$ANS > .2,
      "elevated","normal")
xtabs(~ANSnormal+type+RCType,results_all)
```

```
## , , RCType = OR
## 
```

```
##           type
## ANSnormal   control iwa
##    elevated      37  51
##    normal         9   5
##
## , , RCType = SR
##
##           type
## ANSnormal   control iwa
##    elevated      23  42
##    normal        23  14
```

```r
results_all$GAnormal<-ifelse(results_all$GA < .8,
      "elevated","normal")
xtabs(~GAnormal+type+RCType,results_all)
```

```
## , , RCType = OR
##
##           type
## GAnormal    control iwa
##    elevated      28  51
##    normal        18   5
##
## , , RCType = SR
##
##           type
## GAnormal    control iwa
##    elevated      17  40
##    normal        29  16
```

```r
## two impairments:
results_all$DATGAnormal<-ifelse(results_all$DAT > .06 & results_all$GA < .8,
      "elevated","normal")
xtabs(~DATGAnormal+type+RCType,results_all)
```

```
## , , RCType = OR
##
##             type
## DATGAnormal control iwa
##    elevated      27  44
##    normal        19  12
##
## , , RCType = SR
##
##             type
## DATGAnormal control iwa
##    elevated      11  36
##    normal        35  20
```

```r
results_all$ANSGAnormal<-ifelse(results_all$ANS > .2 & results_all$GA < .8,
      "elevated","normal")
xtabs(~ANSGAnormal+type+RCType,results_all)
```

```
## , , RCType = OR
##
##             type
```

```
## ANSGAnormal control iwa
##     elevated      19  46
##     normal        27  10
##
## , , RCType = SR
##
##           type
## ANSGAnormal control iwa
##     elevated      11  35
##     normal        35  21
```

```r
results_all$DATANSnormal<-ifelse(results_all$DAT > .06 & results_all$ANS > .2,
        "elevated","normal")
xtabs(~DATANSnormal+type+RCType,results_all)
```

```
## , , RCType = OR
##
##             type
## DATANSnormal control iwa
##     elevated       27  41
##     normal         19  15
##
## , , RCType = SR
##
##             type
## DATANSnormal control iwa
##     elevated        5  31
##     normal         41  25
## all three
```

```r
results_all$GADATANSnormal<-ifelse(results_all$GA < 0.8 & results_all$DAT > .06 & results_all$ANS > .2,
        "elevated","normal")
xtabs(~GADATANSnormal+type+RCType,results_all)
```

```
## , , RCType = OR
##
##               type
## GADATANSnormal control iwa
##     elevated        18  39
##     normal          28  17
##
## , , RCType = SR
##
##               type
## GADATANSnormal control iwa
##     elevated         5  31
##     normal          41  25
```