Data analyses for PLoS One paper Vasishth et al 2013

Shravan Vasishth

Date of compilation: June 15, 2021

Note: This compilation needs to be cleaned up a bit. I (SV) will do this soon. But the basic analyses should be reproducible.

```
> library(plyr)
```

1 Gibson and Wu's data

```
  subj item    type pos correct   rt
1    1   13 obj-ext   0       - 3301
2    1   13 obj-ext   1       - 7013
3    1   13 obj-ext   2       - 3941
4    1   13 obj-ext   3       - 1615
5    1   13 obj-ext   4       -  437
6    1   13 obj-ext   5       -  510
```

```
[1] 37
```

```
    item
subj  1  2  3  4  5  6  7  8  9 10 11 13 14 15 16
  1  15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
  2  15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
  3  15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
  4  15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
  5  15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
  6  15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
  7  15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
  8  15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
  9  15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
  11 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
  12 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
  14 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
  15 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
  16 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
```

```
 17 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 18 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 19 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 20 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 21 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 22 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 23 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 24 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 26 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 27 15  0 17  0  0  0  0 13 16  0  0 13  0 11 13
 28 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 29 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 30 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 31 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 32 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 33 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 34 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 35 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 36 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 37 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 38 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 39 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13
 40 15 15 17 17 12 12 12 13 16 14 13 13 13 11 13

[1] 37
```

Next we do some pre-processing to get ready for analysis:

```
[1] 547   8

[1] 547   8

         subj-ext
obj-ext         0
subj-ext        1
```

```
> hnoun$cond2<-factor(ifelse(hnoun$type=="obj-ext","object relative","subject relative"),lev
> tiff("boxplotsboxcox.tiff",res=300,width=17.35,#height=23.35,
+      height=17.35,
+      units="cm",compression="lzw",bg="white")
> par( mfrow=c(2,2) )
> boxplot(rt~cond2,hnoun,ylab="reading time (ms)")
> boxplot(log(rt)~cond2,hnoun,ylab="log reading time (log ms)")
> boxplot(-1000/rt~cond2,hnoun,ylab="negative reciprocal reading time (-1/s)")
> library(MASS)
> boxcox(rt~type*subj,data=hnoun)
> dev.off()
```

```
X11cairo
       2
```

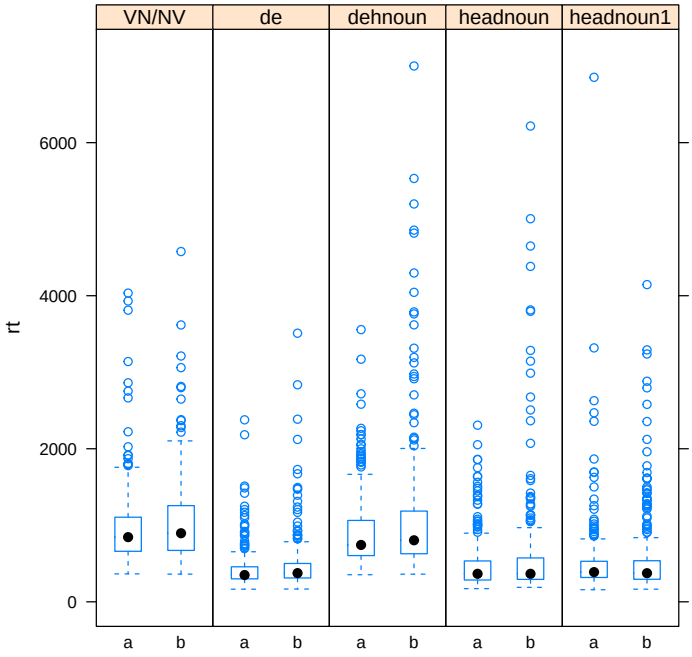
High resolution image:

```
> bitmap("fig1.tiff", height = 4, width = 4, units = 'in', type="tifflzw", res=300)
> par( mfrow=c(2,2) )
> boxplot(rt~cond2,hnoun,ylab="reading time (ms)")
> boxplot(log(rt)~cond2,hnoun,ylab="log reading time (log ms)")
> boxplot(-1000/rt~cond2,hnoun,ylab="negative reciprocal reading time (-1/s)")
> library(MASS)
> boxcox(rt~type*subj,data=hnoun)
> dev.off()

X11cairo
       2
```

The Box-Cox transform suggests using the inverse for the head noun and the region after:

```
> cond<-factor(ifelse(critdata$type=="obj-ext","a","b"))
> critdata$cond<-cond
> ## all regions:
> lattice::bwplot(rt~cond|region,data=critdata,layout=c(5,1))
> par( mfrow=c(3,3) )
> library(MASS)
> #boxcox(rt~type*subj,data=critdata[critdata$region=="de1", ])
>
> boxcox(rt~type*subj,data=critdata[critdata$region=="de", ])
> boxcox(rt~cond*subj,data=critdata[critdata$region=="headnoun", ])
> boxcox(rt~type*subj,data=critdata[critdata$region=="headnoun1", ])
> ## transform:
> critdata$rrt <- -1000/critdata$rt
> means.rrt<-round(with(critdata,tapply(rrt,IND=list(region,type),mean)),digits=3)
> means.rt<-round(with(critdata,tapply(rt,IND=list(region,type),mean)),digits=0)
> library(xtable)
> xtable(cbind(means.rt,means.rrt))

% latex table generated in R 4.0.4 by xtable 1.8-4 package
% Tue Jun 15 21:04:58 2021
\begin{table}[ht]
\centering
\begin{tabular}{rrrrr}
  \hline
 & obj-ext & subj-ext & obj-ext & subj-ext \\
  \hline
VN/NV & 984.00 & 1050.00 & -1.20 & -1.18 \\
```

```
  de & 430.00 & 485.00 & -2.78 & -2.62 \\
  dehnoun & 917.00 & 1095.00 & -1.32 & -1.24 \\
  headnoun & 487.00 & 611.00 & -2.72 & -2.63 \\
  headnoun1 & 520.00 & 564.00 & -2.56 & -2.60 \\
   \hline
\end{tabular}
\end{table}
```




```
> (gw.VN.rrt <- lmer(rrt~so+(1+so|subj)+(1+so|item),subset(critdata,region=="VN/NV")))

Linear mixed model fit by REML ['lmerMod']
Formula: rrt ~ so + (1 + so | subj) + (1 + so | item)
   Data: subset(critdata, region == "VN/NV")
REML criterion at convergence: 500.53
Random effects:
 Groups   Name        Std.Dev. Corr
 subj     (Intercept) 0.3267
          so          0.1328   -0.80
 item     (Intercept) 0.1042
          so          0.0398   -1.00
 Residual             0.3379
Number of obs: 547, groups:  subj, 37; item, 15
```

```
Fixed Effects:
(Intercept)           so
     -1.190       -0.014
optimizer (nloptwrap) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings
```

```
> qqPlot(residuals(gw.VN.rrt))
```

```
6066 2050
 442  150
```

```
> (gw.de.rrt <- lmer(rrt~so+(1+so|subj)+(1+so|item),subset(critdata,region=="de")))
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: rrt ~ so + (1 + so | subj) + (1 + so | item)
   Data: subset(critdata, region == "de")
REML criterion at convergence: 1371.9
Random effects:
 Groups   Name        Std.Dev. Corr
 subj     (Intercept) 0.5526
          so          0.0293   1.00
 item     (Intercept) 0.1825
          so          0.1215   -0.40
 Residual             0.7727
Number of obs: 547, groups:  subj, 37; item, 15
Fixed Effects:
(Intercept)           so
     -2.699       -0.148
optimizer (nloptwrap) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings
```

```
> qqPlot(residuals(gw.de.rrt))
```

```
5927 3768
 432  275
```

```
> (gw.hnoun.rrt <- lmer(rrt~so+(1+so|subj)+(1+so|item),subset(critdata,region=="headnoun")))
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: rrt ~ so + (1 + so | subj) + (1 + so | item)
   Data: subset(critdata, region == "headnoun")
REML criterion at convergence: 1595.6
Random effects:
 Groups   Name        Std.Dev. Corr
 subj     (Intercept) 0.609
          so          0.231    -0.51
 item     (Intercept) 0.332
          so          0.096    1.00
 Residual             0.944
```

```
Number of obs: 547, groups:  subj, 37; item, 15
Fixed Effects:
(Intercept)           so
    -2.6715      -0.0776
optimizer (nloptwrap) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings
```

```
> qqPlot(residuals(gw.hnoun.rrt))
```

```
18261 59661
  133   435
```

```
> (gw.dehnoun.rrt <- lmer(rrt~so+(1+so|subj)+(1+so|item),subset(critdata,region=="dehnoun"))
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: rrt ~ so + (1 + so | subj) + (1 + so | item)
   Data: subset(critdata, region == "dehnoun")
REML criterion at convergence: 685.2
Random effects:
 Groups   Name        Std.Dev. Corr
 subj     (Intercept) 0.2820
          so          0.0160   -1.00
 item     (Intercept) 0.1454
          so          0.0554   0.65
 Residual             0.4096
Number of obs: 547, groups:  subj, 37; item, 15
Fixed Effects:
(Intercept)           so
    -1.2784      -0.0737
optimizer (nloptwrap) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings
```

```
> qqPlot(residuals(gw.dehnoun.rrt))
```

```
1826 5928
 133  432
```

```
> (gw.hnoun1.rrt <- lmer(rrt~so+(1+so|subj)+(1+so|item),subset(critdata,region=="headnoun1")
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: rrt ~ so + (1 + so | subj) + (1 + so | item)
   Data: subset(critdata, region == "headnoun1")
REML criterion at convergence: 1539.3
Random effects:
 Groups   Name        Std.Dev. Corr
 subj     (Intercept) 0.4931
          so          0.1653   -1.00
 item     (Intercept) 0.3988
          so          0.0439   -1.00
```

```
 Residual             0.9048
Number of obs: 547, groups:  subj, 37; item, 15
Fixed Effects:
(Intercept)           so
    -2.5836       0.0423
optimizer (nloptwrap) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings

> qqPlot(residuals(gw.hnoun1.rrt))

5967 6251
 435  455

> ## get coefs, SEs, t-values:
> gw.VN<-c(fixef(gw.VN.rrt)[2],
+       sqrt(vcov(gw.VN.rrt))[2,2],
+       fixef(gw.VN.rrt)[2]/sqrt(vcov(gw.VN.rrt))[2,2])
> gw.de<-c(fixef(gw.de.rrt)[2],
+       sqrt(vcov(gw.de.rrt))[2,2],
+       fixef(gw.de.rrt)[2]/sqrt(vcov(gw.de.rrt))[2,2])
> gw.hn<-c(fixef(gw.hnoun.rrt)[2],
+       sqrt(vcov(gw.hnoun.rrt))[2,2],
+       fixef(gw.hnoun.rrt)[2]/sqrt(vcov(gw.hnoun.rrt))[2,2])
> gw.hn1<-c(fixef(gw.hnoun1.rrt)[2],
+       sqrt(vcov(gw.hnoun1.rrt))[2,2],
+       fixef(gw.hnoun1.rrt)[2]/sqrt(vcov(gw.hnoun1.rrt))[2,2])
> gwresults<-rbind(gw.VN,gw.de,gw.hn,gw.hn1)
> rownames(gwresults)<-c("VN/NV","de","head noun","head noun+1")
> colnames(gwresults)<-c("coef","SE","t-value")
> xtable(round(gwresults,digits=2))

% latex table generated in R 4.0.4 by xtable 1.8-4 package
% Tue Jun 15 21:04:59 2021
\begin{table}[ht]
\centering
\begin{tabular}{rrrr}
  \hline
 & coef & SE & t-value \\
  \hline
VN/NV & -0.01 & 0.04 & -0.37 \\
  de & -0.15 & 0.07 & -2.01 \\
  head noun & -0.08 & 0.09 & -0.84 \\
  head noun+1 & 0.04 & 0.08 & 0.51 \\
   \hline
\end{tabular}
\end{table}
```

We have predictions for the head noun and the word after that, but with

reciprocal RT these are not borne out, cf the published paper's results based on raw RTs (also see below).

```
> (gw.hn.rt <- lmer(rt~so+(1+so|subj)+(1+so|item),subset(critdata,region=="headnoun")))

Linear mixed model fit by REML ['lmerMod']
Formula: rt ~ so + (1 + so | subj) + (1 + so | item)
   Data: subset(critdata, region == "headnoun")
REML criterion at convergence: 8480.1
Random effects:
 Groups   Name        Std.Dev. Corr
 subj     (Intercept) 160
          so          195      -1.00
 item     (Intercept) 154
          so          142      -1.00
 Residual             544
Number of obs: 547, groups:  subj, 37; item, 15
Fixed Effects:
(Intercept)           so
        547         -120
optimizer (nloptwrap) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings

> (gw.hn.rrt <- lmer(-1000/rt~so+(1+so|subj)+(1+so|item),subset(critdata,region=="headnoun")

Linear mixed model fit by REML ['lmerMod']
Formula: -1000/rt ~ so + (1 + so | subj) + (1 + so | item)
   Data: subset(critdata, region == "headnoun")
REML criterion at convergence: 1595.6
Random effects:
 Groups   Name        Std.Dev. Corr
 subj     (Intercept) 0.609
          so          0.231    -0.51
 item     (Intercept) 0.332
          so          0.096    1.00
 Residual             0.944
Number of obs: 547, groups:  subj, 37; item, 15
Fixed Effects:
(Intercept)           so
    -2.6715      -0.0776
optimizer (nloptwrap) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings

> (gw.hn.lrt <- lmer(log(rt)~so+(1+so|subj)+(1+so|item),subset(critdata,region=="headnoun"))

Linear mixed model fit by REML ['lmerMod']
Formula: log(rt) ~ so + (1 + so | subj) + (1 + so | item)
   Data: subset(critdata, region == "headnoun")
REML criterion at convergence: 911.38
```

```
Random effects:
 Groups   Name        Std.Dev. Corr
 subj     (Intercept) 0.244817
          so          0.119070 -1.00
 item     (Intercept) 0.182034
          so          0.000448 1.00
 Residual             0.514326
Number of obs: 547, groups:  subj, 37; item, 15
Fixed Effects:
(Intercept)           so
     6.0618      -0.0725
optimizer (nloptwrap) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings

>

> #tiff("residuals.tiff",res=300,width=7.35,#height=23.35,
> #     height=17.35,
> #     units="cm",compression="lzw",bg="white")
> #
>
> bitmap("fig2.tiff", height = 4, width = 4, units = 'in', type="tifflzw", res=600)
> op<-par(mfrow=c(1,3),pty="s")
> par(cex.lab=1.3)
> qqPlot(residuals(gw.hn.rt),
+        ylab="raw reading time (ms)",
+        envelope=F)

4781 2402
  35   18

> qqPlot(residuals(gw.hn.lrt),ylab="log reading times",envelope=F)

 2402 14901
   18   109

> qqPlot(residuals(gw.hn.rrt),ylab="negative reciprocal reading times",envelope=F)

18261 59661
  133   435

> dev.off()

X11cairo
       2
```

On this raw reading time scale, the differences in rt are about 178 ms at the head noun (OR advantage):

```
> means<-with(critdata,tapply(rt,IND=list(region,type),mean))
```

However, standard deviation is not similar:

```
> sds<-with(critdata,tapply(rt,IND=list(region,type),sd))
```

At the head noun, the ratio of variances is:

```
> round(sds[4,2]/sds[4,1],digits=1)

[1] 2.2
```

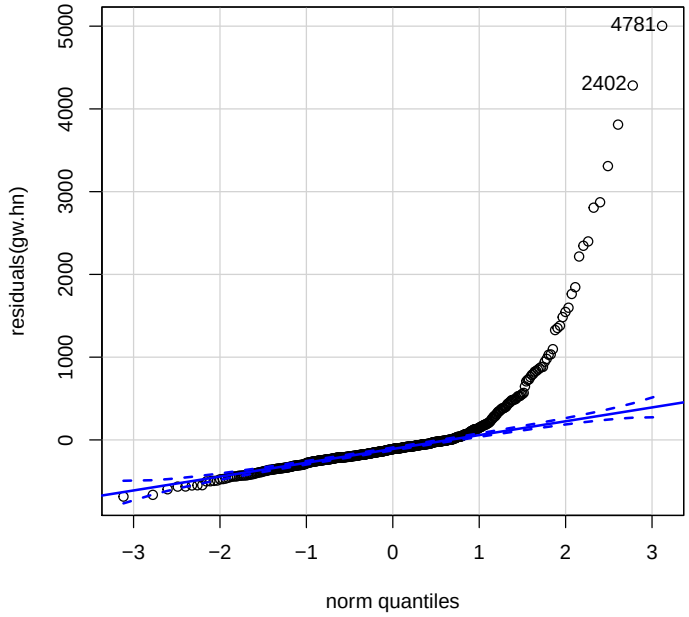
Note that Gibson and Wu fit raw reading times, and got significant effects (OR advantage). Here is an lmer fit analogous (but not identical) to what they did:

```
> ##head noun:
> (gw.hn <- lmer(rt~so+(1|item)+(1|subj),subset(critdata,region=="headnoun")))

Linear mixed model fit by REML ['lmerMod']
Formula: rt ~ so + (1 | item) + (1 | subj)
   Data: subset(critdata, region == "headnoun")
REML criterion at convergence: 8499.5
Random effects:
 Groups   Name        Std.Dev.
 subj     (Intercept) 147
 item     (Intercept) 150
 Residual             560
Number of obs: 547, groups:  subj, 37; item, 15
Fixed Effects:
(Intercept)           so
        548         -120
```

The model estimates that ORs are about 120 ms easier to process than SRs at the head noun.

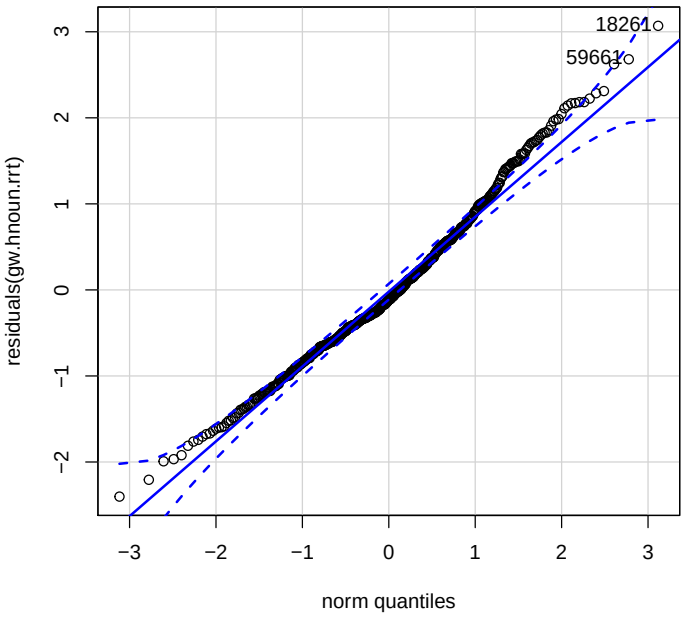
However, statistical significance here is a consequence of the normality assumption (of residuals) not being satisfied; I think that, more precisely, it's the equal variance assumption that's an issue (SR variance is much higher due to those extreme values).

```
> qqPlot(residuals(gw.hn))

4781 2402
  35   18
```



Compare with the reciprocal rt's residuals:

```
> qqPlot(residuals(gw.hnoun.rrt))

18261 59661
  133   435
```



Plotting for paper:

```
> se <- function(x)
+   {
+           y <- x[!is.na(x)] # remove the missing values
+         sqrt(var(as.vector(y))/length(y))
+ }
> ##############Remove between subject variance for SE################
> # (a) Aggregate to Subject x Condition means
>
> library(reshape)
> data.rs <- melt(critdata, id=c("type","region", "subj"), measure=c("rt"),na.rm=TRUE)
> data.id <- data.frame(cast(data.rs, subj + type + region ~ ., function(x) c(rt=mean(x), N=
> # (b) Remove between-subject variance
> ##
> (GM <- mean(tapply(data.id$rt, data.id$subj, mean)))

[1] 716.46

> data.id <- ddply(data.id, .(subj), transform, rt.w = rt - mean(rt) + GM)
> # (c) Compute condition means and error bars:  +/- 2 SE of means after removal of between-
> temp<-melt(data.id, id.var=c("subj","type","region"), measure.var="rt.w")
> (M.id.w <- cast(temp,type+region  ~ .,
+                 function(x) c(M=mean(x), SE=sd(x)/sqrt(length(x)), N=length(x) ) ) )
```

```
       type    region       M     SE  N
1   obj-ext     VN/NV  988.55 27.768 37
2   obj-ext        de  434.78 18.680 37
3   obj-ext   dehnoun  924.96 20.746 37
4   obj-ext  headnoun  490.18 22.305 37
5   obj-ext headnoun1  518.91 34.298 37
6  subj-ext     VN/NV 1058.08 36.187 37
7  subj-ext        de  482.02 23.896 37
8  subj-ext   dehnoun 1095.40 38.137 37
9  subj-ext  headnoun  613.38 31.490 37
10 subj-ext headnoun1  558.29 27.286 37
```

No actual plotting needed here.

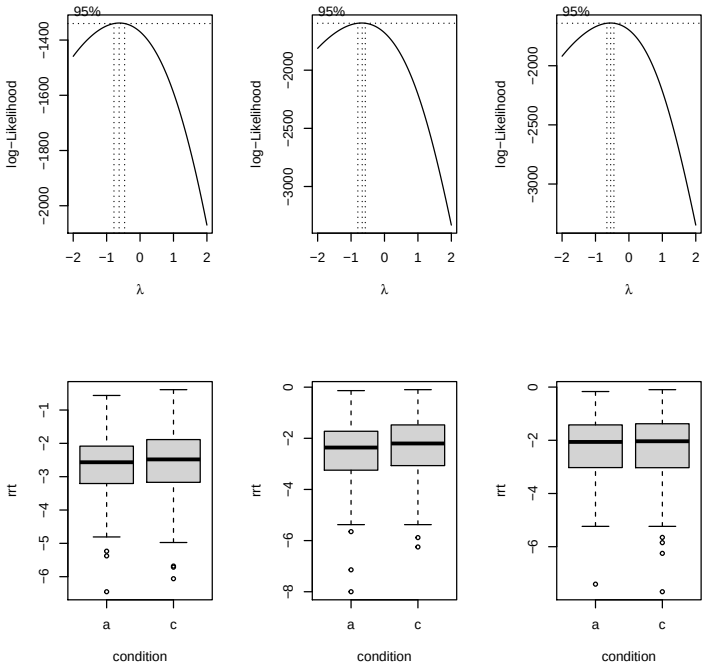
```
> ## to be used later:
> gwcritdata<-critdata
```

2 Experiment 1: Kuo study

```
> ## Design:
> #condition a: SRC    V | OBJ | DE | SBJ (head) | N1 | N2 | (N3)
> #condition c: ORC    SBJ | V | DE | OBJ (head) | N1 | N2 | (N3)
>
> critdata<-read.table("expt1critdata.txt",header=T)
```

a=subj rel, c=obj rel.

```
> par( mfrow=c(2,3) )
> boxcox(rt~condition*subj,data=critdata[critdata$region=="de", ])
> boxcox(rt~condition*subj,data=critdata[critdata$region=="headnoun", ])
> boxcox(rt~condition*subj,data=critdata[critdata$region=="headnoun1", ])
> ## transform:
> critdata$rrt <- -1000/critdata$rt
> critdata.orig<-critdata
> boxplot(rrt~condition,subset(critdata,region=="de"))
> boxplot(rrt~condition,subset(critdata,region=="headnoun"))
> boxplot(rrt~condition,subset(critdata,region=="headnoun1"))
> boxplot(rt~condition,subset(critdata,region=="de"))
> boxplot(rt~condition,subset(critdata,region=="headnoun"))
> boxplot(rt~condition,subset(critdata,region=="headnoun1"))
> ## critical regions:
> headnoun <- subset(critdata,region=="headnoun")
> headnoun1 <- subset(critdata,region=="headnoun1")
```



Reviewer asks for full dataset:

```
> ## load full dataset for PLoS One paper plot:
> critdata<-read.table("expt1fulldata.txt",header=T)
> critdata<-critdata[,c(1,3,4,8,11)]
> critdata<-subset(critdata,rcpos>0)
> summary(critdata)

      subj           item        condition
 Min.   : 1.0   Min.   : 1.00   Length:3000
 1st Qu.:15.8   1st Qu.: 5.75   Class :character
 Median :30.5   Median :10.50   Mode  :character
 Mean   :30.5   Mean   :10.50
 3rd Qu.:45.2   3rd Qu.:15.25
 Max.   :60.0   Max.   :20.00
       rt            rcpos
 Min.   :  125   Min.   :1
 1st Qu.:  322   1st Qu.:2
 Median :  429   Median :3
 Mean   :  578   Mean   :3
 3rd Qu.:  586   3rd Qu.:4
 Max.   :10260   Max.   :5

> region<-ifelse(critdata$rcpos==1,"V/N",
```

```
+             ifelse(critdata$rcpos==2,"N/V",
+               ifelse(critdata$rcpos==3,"de",
+                      ifelse(critdata$rcpos==4,"head noun",
+                             ifelse(critdata$rcpos==5,"head noun+1",NA)))))
> critdata$region<-region
> critdata$region<-factor(critdata$region,levels=c("V/N","N/V","de","head noun","head noun+1
> critdata$cond<-factor(ifelse(critdata$condition=="a",
+                        "subject relative","object relative"),levels=c("subject relative","o
> library(reshape)
> data.rs <- melt(critdata, id=c("cond","region", "subj"), measure=c("rt"),na.rm=TRUE)
> data.id <- data.frame(cast(data.rs, subj + cond + region ~ ., function(x) c(rt=mean(x), N=
> # (b) Remove between-subject variance
> ##
> (GM <- mean(tapply(data.id$rt, data.id$subj, mean)))

[1] 577.63

> data.id <- ddply(data.id, .(subj), transform, rt.w = rt - mean(rt) + GM)
> # (c) Compute condition means and error bars:  +/- 2 SE of means after removal of between-
> temp<-melt(data.id, id.var=c("subj","cond","region"), measure.var="rt.w")
> (M.id.w <- cast(temp,cond+region  ~ .,
+                 function(x) c(M=mean(x), SE=sd(x)/sqrt(length(x)), N=length(x) ) ) )

               cond      region      M     SE  N
1  subject relative         V/N 445.86 24.331 60
2  subject relative         N/V 531.86 19.321 60
3  subject relative          de 428.26 24.670 60
4  subject relative   head noun 631.91 31.581 60
5  subject relative head noun+1 689.74 33.786 60
6   object relative         V/N 486.49 27.758 60
7   object relative         N/V 505.98 22.058 60
8   object relative          de 472.67 24.886 60
9   object relative   head noun 797.69 61.873 60
10  object relative head noun+1 785.83 53.077 60

> byregion.plot<-function(data,
+                         mytitle,k=1,
+                         x.lab="region",
+                         y.lab="reading time [msec]"){
+   ggplot(data,aes(x=region,y=M,
+                         group=cond)) +
+      geom_point(shape=21,size=k*3) +
+      geom_line(aes(linetype=cond),size=k) +
+      geom_errorbar(aes(ymin=M-2*SE,
+                         ymax=M+2*SE),
+                     width=.1,size=k)+
+      xlab(x.lab)+
```

```
+      ylab(y.lab)+
+      labs(title=mytitle) +
+      theme_bw()
+ }
> ## plot:
> #tiff("expt1.tiff",res=300,width=17.35,
> #     height=15.35,
> #     units="cm",compression="lzw",bg
> #="white")
>
> bitmap("fig3.tiff", height = 4, width = 7, units = 'in', type="tifflzw", res=600)
> (plot.regions<-byregion.plot(M.id.w,
+               mytitle="Experiment 1",k=.5,
+               x.lab="region",y.lab="reading time [msec]")
+ )
> dev.off()

X11cairo
       2
```

By region analyses demanded by reviewer: we focus on de, head noun, head noun+1.

```
> critdata$so<-ifelse(critdata$condition=="a",-0.5,0.5)
> ## we stay with negative reciprocal for consistency:
> with(subset(critdata,region=="de"),boxcox(rt~condition*subj))
> with(subset(critdata,region=="head noun"),boxcox(rt~condition*subj))
> with(subset(critdata,region=="head noun+1"),boxcox(rt~condition*subj))
> critdata$rrt<- -1000/critdata$rt
> (kuo.de.rrt <- lmer(rrt~so+(1+so|subj)+(1+so|item),subset(critdata,region=="de")))

Linear mixed model fit by REML ['lmerMod']
Formula: rrt ~ so + (1 + so | subj) + (1 + so | item)
   Data: subset(critdata, region == "de")
REML criterion at convergence: 1293.5
Random effects:
 Groups   Name        Std.Dev. Corr
 subj     (Intercept) 0.7172
          so          0.0895   1.00
 item     (Intercept) 0.0830
          so          0.1290   -0.33
 Residual             0.6124
Number of obs: 600, groups:  subj, 60; item, 20
Fixed Effects:
(Intercept)           so
    -2.6182       0.0823
optimizer (nloptwrap) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings
```

```
> qqPlot(residuals(gw.de.rrt))

5927 3768
 432  275

> (kuo.hnoun.rrt <- lmer(rt~so+(1+so|subj)+(1+so|item),subset(critdata,region=="head noun"))

Linear mixed model fit by REML ['lmerMod']
Formula: rt ~ so + (1 + so | subj) + (1 + so | item)
   Data: subset(critdata, region == "head noun")
REML criterion at convergence: 9576.6
Random effects:
 Groups   Name        Std.Dev. Corr
 subj     (Intercept) 534
          so          237      1.00
 item     (Intercept) 200
          so          161      1.00
 Residual             633
Number of obs: 600, groups:  subj, 60; item, 20
Fixed Effects:
(Intercept)           so
        715          166
optimizer (nloptwrap) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings

> hist(residuals(kuo.hnoun.rrt))
> qqPlot(residuals(kuo.hnoun.rrt))

2189 1880
 174  150

> qqPlot(residuals(gw.hnoun.rrt))

18261 59661
  133   435

> (kuo.hnoun1.rrt <- lmer(rrt~so+(1+so|subj)+(1+so|item),subset(critdata,region=="head noun+

Linear mixed model fit by REML ['lmerMod']
Formula: rrt ~ so + (1 + so | subj) + (1 + so | item)
   Data: subset(critdata, region == "head noun+1")
REML criterion at convergence: 1614.9
Random effects:
 Groups   Name        Std.Dev. Corr
 subj     (Intercept) 0.891
          so          0.108    1.00
 item     (Intercept) 0.206
          so          0.100    -1.00
 Residual             0.799
```

```
Number of obs: 600, groups:  subj, 60; item, 20
Fixed Effects:
(Intercept)           so
    -2.2292       0.0555
optimizer (nloptwrap) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings
```

```
> qqPlot(residuals(gw.hnoun1.rrt))
```

```
5967 6251
 435  455
```

Tabulating results:

```
> ## get coefs, SEs, t-values:
> kuo.de<-c(fixef(kuo.de.rrt)[2],
+       sqrt(vcov(kuo.de.rrt))[2,2],
+       fixef(kuo.de.rrt)[2]/sqrt(vcov(kuo.de.rrt))[2,2])
> kuo.hnoun<-c(fixef(kuo.hnoun.rrt)[2],
+       sqrt(vcov(kuo.hnoun.rrt))[2,2],
+       fixef(kuo.hnoun.rrt)[2]/sqrt(vcov(kuo.hnoun.rrt))[2,2])
> kuo.hnoun1<-c(fixef(kuo.hnoun1.rrt)[2],
+       sqrt(vcov(kuo.hnoun1.rrt))[2,2],
+       fixef(kuo.hnoun1.rrt)[2]/sqrt(vcov(kuo.hnoun1.rrt))[2,2])
> kuoresults<-rbind(kuo.de,kuo.hnoun,kuo.hnoun1)
> rownames(kuoresults)<-c("de","head noun","head noun+1")
> colnames(kuoresults)<-c("coef.","SE","t-value")
> xtable(round(kuoresults,digits=2))
```

```
% latex table generated in R 4.0.4 by xtable 1.8-4 package
% Tue Jun 15 21:05:00 2021
\begin{table}[ht]
\centering
\begin{tabular}{rrrr}
  \hline
 & coef. & SE & t-value \\
  \hline
de & 0.08 & 0.06 & 1.40 \\
  head noun & 165.78 & 70.01 & 2.37 \\
  head noun+1 & 0.06 & 0.07 & 0.79 \\
   \hline
\end{tabular}
\end{table}
```

3 Experiment 2: Qiang Li

```
> expt2.allregions<-read.table("expt2allregions.txt",header=T)
> head(expt2.allregions)
```

```
    subj item condition   rt rcpos
58 1.dat   22         b  995     1
59 1.dat   22         b  904     2
60 1.dat   22         b  859     3
61 1.dat   22         b 2150     4
62 1.dat   22         b 1491     5
95 1.dat   18         b  950     1
```

```
> region<-factor(ifelse(expt2.allregions$rcpos==1,"V/N",
+        ifelse(expt2.allregions$rcpos==2,"N/V",
+               ifelse(expt2.allregions$rcpos==3,"de",
+                      ifelse(expt2.allregions$rcpos==4,"head noun",ifelse(expt2.allregions$
> region<-factor(region,levels=c("V/N","N/V","de","head noun","head noun+1"))
> unique(region)
```

```
[1] V/N         N/V         de          head noun   head noun+1
Levels: V/N N/V de head noun head noun+1
```

```
> expt2.allregions$region<-region
```

Plotting:

```
> critdata<-subset(expt2.allregions,condition%in%c("a","b"))
> critdata$condition<-factor(critdata$condition)
> RCType<-factor(ifelse(critdata$condition=="a","subject relative","object relative"),levels
> critdata$RCType<-RCType
> library(reshape)
> data.rs <- melt(critdata, id=c("region","RCType","subj"), measure=c("rt"),na.rm=TRUE)
> data.id <- data.frame(cast(data.rs, subj + RCType + region ~ ., function(x) c(rt=mean(x),
> # (b) Remove between-subject variance
> ##
> (GM <- mean(tapply(data.id$rt, data.id$subj, mean)))
```

```
[1] 631.91
```

```
> data.id <- ddply(data.id, .(subj), transform, rt.w = rt - mean(rt) + GM)
> # (c) Compute condition means and error bars:  +/- 2 SE of means after removal of between-
> temp<-melt(data.id, id.var=c("subj","RCType","region"), measure.var="rt.w")
> (M.id.w <- cast(temp,RCType+region  ~ .,
+                 function(x) c(M=mean(x), SE=sd(x)/sqrt(length(x)), N=length(x) ) ) )
```

```
             RCType      region      M     SE  N
1  subject relative         V/N 575.89 23.216 61
2  subject relative         N/V 583.89 19.754 61
3  subject relative          de 471.93 17.337 61
4  subject relative   head noun 608.15 20.878 61
5  subject relative head noun+1 791.38 37.315 61
```

```
6   object relative         V/N 537.37 18.805 61
7   object relative         N/V 628.26 22.209 61
8   object relative          de 492.00 15.557 61
9   object relative   head noun 691.13 35.912 61
10  object relative head noun+1 939.14 50.793 61
```

```
> #tiff("expt2.tiff",res=300,width=17.35,
> #     height=15.35,
> #     units="cm",compression="lzw",bg
> #="white")
>
> bitmap("fig4.tiff", height = 4, width = 7, units = 'in', type="tifflzw", res=600)
> byregion.plot<-function(data,
+                         mytitle,k=.5,
+                         x.lab="region",
+                         y.lab="reading time [msec]"){
+   ggplot(data,aes(x=region,y=M,
+                         group=RCType)) +
+      geom_point(shape=21,size=k*3) +
+      geom_line(aes(linetype=RCType),size=k) +
+      geom_errorbar(aes(ymin=M-2*SE,
+                         ymax=M+2*SE),
+                     width=.1,size=k)+
+      xlab(x.lab)+
+      ylab(y.lab)+
+      labs(title=mytitle) +
+      theme_bw()
+ }
> (plot.regions<-byregion.plot(M.id.w,
+               mytitle="Experiment 2",k=.5,
+               x.lab="region",y.lab="reading time [msec]")
+ )
> dev.off()
```

```
X11cairo
       2
```

```
> # condition a: Subj-modifying SRC
> # condition b: Subj-modifying ORC
> # condition c: Obj-modifying  SRC
> # condition d: Obj-modifying  ORC
> ## We focus on a,b for consistency with the
> ## earlier studies
>
> #a=subj rel, c=obj rel.
>
> critdata<-read.table("expt2critdata.txt",header=T)
```

```
> par( mfrow=c(2,3) )
> boxcox(rt~condition*subj,data=critdata[critdata$region=="de", ])
> boxcox(rt~condition*subj,data=critdata[critdata$region=="headnoun", ])
> boxcox(rt~condition*subj,data=critdata[critdata$region=="headnoun1", ])
> critdata$rrt <- -1000/critdata$rt
> headnoun <- subset(critdata,region=="headnoun")
> headnoun1 <- subset(critdata,region=="headnoun1")
> library(MASS)
> with(headnoun,boxcox(rt~condition*subj))
```

lmer models:

```
> with(critdata,tapply(rt,IND=list(region,condition),mean))

               a      b
de        471.93 492.00
headnoun  608.15 691.13
headnoun1 791.38 939.14

> (m1<-lmer(rt~condition+(1|subj)+(1|item),headnoun))

Linear mixed model fit by REML ['lmerMod']
Formula: rt ~ condition + (1 | subj) + (1 | item)
   Data: headnoun
REML criterion at convergence: 11392
Random effects:
 Groups   Name        Std.Dev.
 subj     (Intercept) 218.0
 item     (Intercept)  95.6
 Residual             557.1
Number of obs: 732, groups:  subj, 61; item, 24
Fixed Effects:
(Intercept)   conditionb
      608.4         82.6

> qqPlot(residuals(m1))

28965  4031
  278    38

> (m2<-lmer(rrt~so+(1+so|subj)+(1|item),headnoun))

Linear mixed model fit by REML ['lmerMod']
Formula: rrt ~ so + (1 + so | subj) + (1 | item)
   Data: headnoun
REML criterion at convergence: 1782.3
Random effects:
 Groups   Name        Std.Dev. Corr
```

```
 subj     (Intercept) 0.5054
          so          0.0669   1.00
 item     (Intercept) 0.1609
 Residual             0.7429
Number of obs: 732, groups:  subj, 61; item, 24
Fixed Effects:
(Intercept)           so
    -2.1357       0.0683

> qqPlot(residuals(m2))

25541 58311
  247   559

> (m3<-lmer(rrt~so+(1+so|subj)+(1|item),headnoun1))

Linear mixed model fit by REML ['lmerMod']
Formula: rrt ~ so + (1 + so | subj) + (1 | item)
   Data: headnoun1
REML criterion at convergence: 1638.6
Random effects:
 Groups   Name        Std.Dev. Corr
 subj     (Intercept) 0.4423
          so          0.0593   0.19
 item     (Intercept) 0.3305
 Residual             0.6603
Number of obs: 732, groups:  subj, 61; item, 24
Fixed Effects:
(Intercept)           so
     -1.756        0.229

> qqPlot(residuals(m3))

25079 47734
  241   459

>
```

By region analyses as demanded by reviewer:

```
> (qiang.de.rrt<-lmer(rrt~so+(1+so|subj)+(1+so|item),subset(critdata,region=="de")))

Linear mixed model fit by REML ['lmerMod']
Formula: rrt ~ so + (1 + so | subj) + (1 + so | item)
   Data: subset(critdata, region == "de")
REML criterion at convergence: 1497.5
Random effects:
 Groups   Name        Std.Dev. Corr
```

```
 subj     (Intercept) 0.4000
          so          0.0544   1.00
 item     (Intercept) 0.1267
          so          0.1526   -0.86
 Residual             0.6110
Number of obs: 732, groups:  subj, 61; item, 24
Fixed Effects:
(Intercept)           so
    -2.3909       0.0745
optimizer (nloptwrap) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings

> qqPlot(residuals(qiang.de.rrt))

36373 25077
  350   241

> (qiang.hnoun.rrt<-lmer(rrt~so+(1+so|subj)+(1+so|item),subset(critdata,region=="headnoun"))

Linear mixed model fit by REML ['lmerMod']
Formula: rrt ~ so + (1 + so | subj) + (1 + so | item)
   Data: subset(critdata, region == "headnoun")
REML criterion at convergence: 1782.2
Random effects:
 Groups   Name        Std.Dev. Corr
 subj     (Intercept) 0.5049
          so          0.0672   1.00
 item     (Intercept) 0.1610
          so          0.0323   -1.00
 Residual             0.7427
Number of obs: 732, groups:  subj, 61; item, 24
Fixed Effects:
(Intercept)           so
    -2.1356       0.0683
optimizer (nloptwrap) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings

> qqPlot(residuals(qiang.hnoun.rrt))

25541 58311
  247   559

> (qiang.hnoun1.rrt<-lmer(rrt~so+(1+so|subj)+(1+so|item),subset(critdata,region=="headnoun1"

Linear mixed model fit by REML ['lmerMod']
Formula: rrt ~ so + (1 + so | subj) + (1 + so | item)
   Data: subset(critdata, region == "headnoun1")
REML criterion at convergence: 1637.3
Random effects:
```

```
 Groups   Name        Std.Dev. Corr
 subj     (Intercept) 0.443
          so          0.074    0.15
 item     (Intercept) 0.331
          so          0.144    0.21
 Residual             0.656
Number of obs: 732, groups:  subj, 61; item, 24
Fixed Effects:
(Intercept)           so
     -1.756        0.229

> qqPlot(residuals(qiang.hnoun1.rrt))

25079 47734
  241   459

> #bwplot(rrt~cond,subset(critdata,region=="headnoun1" & rrt>-4))
>
> ## assemble table:
> qiang.de<-c(fixef(qiang.de.rrt)[2],
+       sqrt(vcov(qiang.de.rrt))[2,2],
+       fixef(qiang.de.rrt)[2]/sqrt(vcov(qiang.de.rrt))[2,2])
> qiang.hnoun<-c(fixef(qiang.hnoun.rrt)[2],
+       sqrt(vcov(qiang.hnoun.rrt))[2,2],
+       fixef(qiang.hnoun.rrt)[2]/sqrt(vcov(qiang.hnoun.rrt))[2,2])
> qiang.hnoun1<-c(fixef(qiang.hnoun1.rrt)[2],
+       sqrt(vcov(qiang.hnoun1.rrt))[2,2],
+       fixef(qiang.hnoun1.rrt)[2]/sqrt(vcov(qiang.hnoun1.rrt))[2,2])
> qiangresults<-rbind(qiang.de,qiang.hnoun,qiang.hnoun1)
> rownames(qiangresults)<-c("de","head noun","head noun+1")
> colnames(qiangresults)<-c("coef.","SE","t-value")
> xtable(qiangresults)

% latex table generated in R 4.0.4 by xtable 1.8-4 package
% Tue Jun 15 21:05:02 2021
\begin{table}[ht]
\centering
\begin{tabular}{rrrr}
  \hline
 & coef. & SE & t-value \\
  \hline
de & 0.07 & 0.06 & 1.35 \\
  head noun & 0.07 & 0.06 & 1.22 \\
  head noun+1 & 0.23 & 0.06 & 3.97 \\
   \hline
\end{tabular}
\end{table}
```

```
> ##############Remove between subject variance for SE################
> # (a) Aggregate to Subject x Condition means
>
> data.rs <- melt(critdata, id=c("condition","region", "subj"), measure=c("rt"),na.rm=TRUE)
> data.id <- data.frame(cast(data.rs, subj + condition + region ~ ., function(x) c(rt=mean(x
> # (b) Remove between-subject variance
> ##
> (GM <- mean(tapply(data.id$rt, data.id$subj, mean)))

[1] 665.62

> data.id <- ddply(data.id, .(subj), transform, rt.w = rt - mean(rt) + GM)
> # (c) Compute condition means and error bars:  +/- 2 SE of means after removal of between-
> temp<-melt(data.id, id.var=c("subj","condition","region"), measure.var="rt.w")
> (M.id.w <- cast(temp,condition+region  ~ .,
+                 function(x) c(M=mean(x), SE=sd(x)/sqrt(length(x)), N=length(x) ) ) )

  condition    region      M     SE  N
1         a        de 471.93 23.335 61
2         a  headnoun 608.15 19.881 61
3         a headnoun1 791.38 34.033 61
4         b        de 492.00 20.754 61
5         b  headnoun 691.13 33.132 61
6         b headnoun1 939.14 46.386 61
```

4 Experiment 3: Gibson and Wu replication

Now we look at our "exact" replication of Gibson and Wu:

```
> gwrerun<-read.table("gwrerun.txt",header=F)
> colnames(gwrerun) <- c("machine","subj","item","condition","pos","word","correct","rt")
> gwrerun$subj<-paste(gwrerun$machine,gwrerun$subj,sep="")
> questions.gwrerun <- subset(gwrerun,correct%in%c(0,1))
> #with(questions.gwrerun,tapply(as.integer(as.character(correct)),condition,mean))
>
> #lmer(I(-1/rt)~ condition+(1|subj)+(1|item),questions.gwrerun)
>
>
> #with(questions.gwrerun,tapply(rt,condition,mean))
>
> questions.lmer.gwrerun<- glmer(factor(correct)~condition+(1|item)+(1|item),questions.gwrer
```

Plotting:

```
> ## isolate relevant columns:
> gwdata<-gwrerun[,c(2,3,4,5,8)]
> #xtabs(~subj+condition,gwdata)
```

```
>
> region<-factor(ifelse(gwdata$pos==5,"V/N",
+                 ifelse(gwdata$pos==6,"N/V",
+                 ifelse(gwdata$pos==7,"de",
+                        ifelse(gwdata$pos==8,"head noun",
+                        ifelse(gwdata$pos==9,"head noun+1",-1))))),
+                 levels=c("V/N","N/V","de","head noun","head noun+1"))
> gwdata$region<-region
> gwdata<-subset(gwdata,region!="-1")
> gwdata$region<-factor(gwdata$region)
> critdata<-gwdata
> critdata$subj<-factor(critdata$subj)
> with(subset(critdata,region=="de"),boxcox(rt~condition*subj))
> with(subset(critdata,region=="head noun"),boxcox(rt~condition*subj))
> with(subset(critdata,region=="head noun+1"),boxcox(rt~condition*subj))
> critdata$rrt<- -1000/critdata$rt
> critdata$cond <- factor(ifelse(critdata$cond=="subj-ext","subject relative","object relati
> ## rt plot
> data.rs <- melt(critdata, id=c("cond","region", "subj"), measure=c("rt"),na.rm=TRUE)
> data.id <- data.frame(cast(data.rs, subj + cond + region ~ ., function(x) c(rt=mean(x), N=
> # (b) Remove between-subject variance
> ##
> (GM <- mean(tapply(data.id$rt, data.id$subj, mean)))

[1] 492.31

> data.id <- ddply(data.id, .(subj), transform, rt.w = rt - mean(rt) + GM)
> # (c) Compute condition means and error bars:  +/- 2 SE of means after removal of between-
> temp<-melt(data.id, id.var=c("subj","cond","region"), measure.var="rt.w")
> (M.id.w <- cast(temp,cond+region  ~ .,
+                 function(x) c(M=mean(x), SE=sd(x)/sqrt(length(x)), N=length(x) ) ) )

               cond      region      M     SE  N
1  subject relative         V/N 496.43 21.150 40
2  subject relative         N/V 563.19 26.806 40
3  subject relative          de 479.97 25.724 40
4  subject relative   head noun 557.81 44.064 40
5  subject relative head noun+1 534.22 21.598 40
6   object relative         V/N 516.72 25.454 40
7   object relative         N/V 458.01 12.378 40
8   object relative          de 384.93 13.454 40
9   object relative   head noun 442.30 16.056 40
10  object relative head noun+1 489.50 31.283 40

> byregion.plot<-function(data,
+                         mytitle,k=1,
+                         x.lab="region",
```

```
+                          y.lab="reading time [msec]"){
+   ggplot(data,aes(x=region,y=M,
+                          group=cond)) +
+      geom_point(shape=21,size=k*3) +
+      geom_line(aes(linetype=cond),size=k) +
+      geom_errorbar(aes(ymin=M-2*SE,
+                          ymax=M+2*SE),
+                      width=.1,size=k)+
+      xlab(x.lab)+
+      ylab(y.lab)+
+      labs(title=mytitle) +
+      theme_bw()
+ }
> ## plot:
> #tiff("expt3.tiff",res=300,width=17.35,
> #     height=15.35,
> #     units="cm",compression="lzw",bg
> #="white")
>
> bitmap("fig5.tiff", height = 4, width = 7, units = 'in', type="tifflzw", res=600)
> (plot.regions<-byregion.plot(M.id.w,
+                mytitle="Experiment 3",k=0.5,
+                x.lab="region",y.lab="reading time [msec]")
+ )
> dev.off()

X11cairo
       2
```

The reviewer demands a fuller analysis:

```
> so<-ifelse(critdata$condition=="subj-ext",-0.5,0.5)
> critdata$so<-so
> ## a surprising OR advantage, cannot be attributed to
> ## storage cost:
> (gwrerun.de.rrt <- lmer(rrt~so+(1+so|subj)+(1+so|item),subset(critdata,region=="de")))

Linear mixed model fit by REML ['lmerMod']
Formula: rrt ~ so + (1 + so | subj) + (1 + so | item)
   Data: subset(critdata, region == "de")
REML criterion at convergence: 1419.3
Random effects:
 Groups   Name        Std.Dev. Corr
 subj     (Intercept) 0.452950
          so          0.000769 -1.00
 item     (Intercept) 0.125248
          so          0.162093 -0.19
```

```
 Residual             0.733794
Number of obs: 596, groups:  subj, 40; item, 15
Fixed Effects:
(Intercept)           so
     -2.737       -0.221
optimizer (nloptwrap) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings
```

```
> qqPlot(residuals(gwrerun.de.rrt))
```

```
8218 6333
 596  459
```

```
> lattice::bwplot(rrt~cond,subset(critdata,region=="de"))
> row<-which(subset(critdata,region=="de")$rrt< -5.5)
> ## item 1 is unusual but doesn't affect result:
> subset(critdata,region=="de")[row,]
```

```
     subj item condition pos  rt region    rrt            cond
8218  2m9    1   obj-ext   7 167     de -5.988 object relative
      so
8218 0.5
```

```
> (gwrerun.hnoun.rrt <- lmer(rrt~so+(1+so|subj)+(1+so|item),subset(critdata,region=="head no
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: rrt ~ so + (1 + so | subj) + (1 + so | item)
   Data: subset(critdata, region == "head noun")
REML criterion at convergence: 1553.5
Random effects:
 Groups   Name        Std.Dev. Corr
 subj     (Intercept) 0.607
          so          0.267    0.27
 item     (Intercept) 0.261
          so          0.261    -0.48
 Residual             0.794
Number of obs: 595, groups:  subj, 40; item, 15
Fixed Effects:
(Intercept)           so
     -2.738       -0.149
```

```
> qqPlot(residuals(gwrerun.hnoun.rrt))
```

```
6358 1749
 460  126
```

```
> row<-which(subset(critdata,region=="head noun")$rrt< -5.5)
> ## item 3 in the object extracted condition is unusual:
> subset(critdata,region=="head noun")[row,]
```

```
     subj item condition pos  rt    region    rrt
1733 1m17    3   obj-ext   8 172 head noun -5.814
                cond  so
1733 object relative 0.5

> (gwrerun.hnoun1.rrt <- lmer(rrt~so+(1+so|subj)+(1+so|item),subset(critdata,region=="head n

Linear mixed model fit by REML ['lmerMod']
Formula: rrt ~ so + (1 + so | subj) + (1 + so | item)
   Data: subset(critdata, region == "head noun+1")
REML criterion at convergence: 1617.1
Random effects:
 Groups   Name        Std.Dev. Corr
 subj     (Intercept) 0.558
          so          0.246    -0.19
 item     (Intercept) 0.390
          so          0.204    -0.15
 Residual             0.842
Number of obs: 595, groups:  subj, 40; item, 15
Fixed Effects:
(Intercept)           so
    -2.6547      -0.0562

> qqPlot(residuals(gwrerun.hnoun1.rrt))

1467 8220
 106  595

> ## one data point is unusual on ORs, but does not affect result:
> lattice::bwplot(rrt~cond,subset(critdata,region=="head noun+1" & rrt> -6))
> ## get coefs, SEs, t-values:
> gwrerun.de<-c(fixef(gwrerun.de.rrt)[2],
+       sqrt(vcov(gwrerun.de.rrt))[2,2],
+       fixef(gwrerun.de.rrt)[2]/sqrt(vcov(gwrerun.de.rrt))[2,2])
> gwrerun.hnoun<-c(fixef(gwrerun.hnoun.rrt)[2],
+       sqrt(vcov(gwrerun.hnoun.rrt))[2,2],
+       fixef(gwrerun.hnoun.rrt)[2]/sqrt(vcov(gwrerun.hnoun.rrt))[2,2])
> gwrerun.hnoun1<-c(fixef(gwrerun.hnoun1.rrt)[2],
+       sqrt(vcov(gwrerun.hnoun1.rrt))[2,2],
+       fixef(gwrerun.hnoun1.rrt)[2]/sqrt(vcov(gwrerun.hnoun1.rrt))[2,2])
> gwrerunresults<-rbind(gwrerun.de,gwrerun.hnoun,gwrerun.hnoun1)
> rownames(gwrerunresults)<-c("de","head noun","head noun+1")
> colnames(gwrerunresults)<-c("coef.","SE","t-value")
> xtable(gwrerunresults)

% latex table generated in R 4.0.4 by xtable 1.8-4 package
% Tue Jun 15 21:05:03 2021
```

```
\begin{table}[ht]
\centering
\begin{tabular}{rrrr}
  \hline
 & coef. & SE & t-value \\
  \hline
de & -0.22 & 0.07 & -3.02 \\
  head noun & -0.15 & 0.10 & -1.44 \\
  head noun+1 & -0.06 & 0.10 & -0.59 \\
   \hline
\end{tabular}
\end{table}
```

Combined analysis:

```
> ## rerun:
> gwrerun<-critdata
> ## original data:
> gwcritdata<-gwcritdata[,c(1,2,3,4,7)]
> gwrerun<-gwrerun[,1:5]
> head(gwrerun)

   subj item condition pos   rt
6   1m1   15   obj-ext   5  566
7   1m1   15   obj-ext   6 1041
8   1m1   15   obj-ext   7  733
9   1m1   15   obj-ext   8  832
10  1m1   15   obj-ext   9 1859
17  1m1    8  subj-ext   5  515

> head(gwcritdata)

   subj item     type pos   rt
7     1   13  obj-ext   6 1140
20    1    6 subj-ext   6 1197
32    1    5  obj-ext   6  756
44    1    9  obj-ext   6  643
60    1   14 subj-ext   6  860
73    1    4 subj-ext   6  868

> colnames(gwcritdata)[3]<-"condition"
> gwcritdata$expt<-factor("gw")
> gwrerun$expt<-factor("gwrerun")
> gwall<-rbind(gwcritdata,gwrerun)
> head(gwall)

   subj item condition pos   rt expt
7     1   13   obj-ext   6 1140   gw
```

```
20    1    6  subj-ext   6 1197   gw
32    1    5   obj-ext   6  756   gw
44    1    9   obj-ext   6  643   gw
60    1   14  subj-ext   6  860   gw
73    1    4  subj-ext   6  868   gw
```

```
> length(unique(gwall$subj))
```

```
[1] 77
```

```
> #xtabs(~subj+condition,gwall)
```

```
> critdata<-gwall
> region<-factor(ifelse(critdata$pos==5,"V/N",
+                 ifelse(critdata$pos==6,"N/V",
+                 ifelse(critdata$pos==7,"de",
+                        ifelse(critdata$pos==8,"head noun",
+                        ifelse(critdata$pos==9,"head noun+1",-1))))),
+                 levels=c("V/N","N/V","de","head noun","head noun+1"))
> critdata$region<-region
> critdata<-subset(critdata,region!="-1")
> critdata$region<-factor(critdata$region)
> critdata$subj<-factor(critdata$subj)
> with(subset(critdata,region=="de"),boxcox(rt~condition*subj))
> with(subset(critdata,region=="head noun"),boxcox(rt~condition*subj))
> with(subset(critdata,region=="head noun+1"),boxcox(rt~condition*subj))
> ## for consistency, we use negative reciprocal:
> critdata$rrt<- -1000/critdata$rt
> critdata$type <- factor(ifelse(critdata$condition=="subj-ext","subject relative","object r
> data.rs <- melt(critdata, id=c("type","region", "subj"), measure=c("rt"),na.rm=TRUE)
> data.id <- data.frame(cast(data.rs, subj + type + region ~ ., function(x) c(rt=mean(x), N=
> # (b) Remove between-subject variance
> ##
> (GM <- mean(tapply(data.id$rt, data.id$subj, mean)))
```

```
[1] 592.26
```

```
> data.id <- ddply(data.id, .(subj), transform, rt.w = rt - mean(rt) + GM)
> # (c) Compute condition means and error bars:  +/- 2 SE of means after removal of between-
> temp<-melt(data.id, id.var=c("subj","type","region"), measure.var="rt.w")
> (M.id.w <- cast(temp,type+region  ~ .,
+                 function(x) c(M=mean(x), SE=sd(x)/sqrt(length(x)), N=length(x) ) ) )
```

```
              type      region      M     SE  N
1  subject relative         V/N 596.39 21.150 40
2  subject relative         N/V 801.00 27.791 77
3  subject relative          de 480.96 20.660 77
```

```
4  subject relative   head noun 700.32 28.832 77
5  subject relative head noun+1 545.79 19.726 77
6   object relative         V/N 616.67 25.454 40
7   object relative         N/V 712.95 23.618 77
8   object relative          de 408.89 14.039 77
9   object relative   head noun 569.77 12.369 77
10  object relative head noun+1 503.63 24.681 77
```

```
> byregion.plot<-function(data,
+                         mytitle,k=1,
+                         x.lab="region",
+                         y.lab="reading time [msec]"){
+   ggplot(data,aes(x=region,y=M,
+                         group=type)) +
+      geom_point(shape=21,size=k*3) +
+      geom_line(aes(linetype=type),size=k) +
+      geom_errorbar(aes(ymin=M-2*SE,
+                         ymax=M+2*SE),
+                     width=.1,size=k)+
+      xlab(x.lab)+
+      ylab(y.lab)+
+      labs(title=mytitle) +
+      theme_bw()
+ }
> tiff("expt3a.tiff",res=300,width=17.35,
+      height=15.35,
+      units="cm",compression="lzw",bg="white")
> bitmap("fig6.tiff", height = 4, width = 7, units = 'in', type="tifflzw", res=600)
> ## plot:
> (plot.regions<-byregion.plot(M.id.w,
+               mytitle="Combined data (Gibson and Wu expt. \n and replication)",k=0.5,
+               x.lab="region",y.lab="reading time [msec]")
+ )
> dev.off()
```

```
X11cairo
       2
```

```
> critdata$so<-ifelse(critdata$condition=="subj-ext",-0.5,0.5)
> (gwall.de.rrt <- lmer(rrt~so+(1+so|subj)+(1+so|item),subset(critdata,region=="de")))
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: rrt ~ so + (1 + so | subj) + (1 + so | item)
   Data: subset(critdata, region == "de")
REML criterion at convergence: 2787.3
Random effects:
 Groups   Name        Std.Dev. Corr
```

```
 subj     (Intercept) 0.4999
          so          0.0297   1.00
 item     (Intercept) 0.1462
          so          0.1188   -0.08
 Residual             0.7553
Number of obs: 1143, groups:  subj, 77; item, 15
Fixed Effects:
(Intercept)           so
     -2.719       -0.187
optimizer (nloptwrap) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings
```

```
> qqPlot(residuals(gwall.de.rrt))
```

```
8218 6333
1143 1006
```

```
> (gwall.hnoun.rrt <- lmer(rrt~so+(1+so|subj)+(1+so|item),subset(critdata,region=="head noun
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: rrt ~ so + (1 + so | subj) + (1 + so | item)
   Data: subset(critdata, region == "head noun")
REML criterion at convergence: 4889.1
Random effects:
 Groups   Name        Std.Dev. Corr
 subj     (Intercept) 0.6439
          so          0.0115   0.98
 item     (Intercept) 0.2367
          so          0.0253   1.00
 Residual             0.9658
Number of obs: 1689, groups:  subj, 77; item, 15
Fixed Effects:
(Intercept)           so
     -2.358       -0.102
optimizer (nloptwrap) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings
```

```
> qqPlot(residuals(gwall.hnoun.rrt))
```

```
68921 28801
 1049   757
```

```
> (gwall.hnoun1.rrt <- lmer(rrt~so+(1+so|subj)+(1+so|item),subset(critdata,region=="head nou
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: rrt ~ so + (1 + so | subj) + (1 + so | item)
   Data: subset(critdata, region == "head noun+1")
REML criterion at convergence: 3136.2
Random effects:
```

```
 Groups   Name        Std.Dev. Corr
 subj     (Intercept) 0.5231
          so          0.0994   -0.97
 item     (Intercept) 0.3970
          so          0.0118   -1.00
 Residual             0.8804
Number of obs: 1142, groups:  subj, 77; item, 15
Fixed Effects:
(Intercept)           so
   -2.61994     -0.00867
optimizer (nloptwrap) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings
```

```
> qqPlot(residuals(gwall.hnoun1.rrt))
```

```
14671  5967
  653   435
```

```
> ## sign of fit above changes if we remove the data points that have rrt> -6, but result do
> ##not change:
> lattice::bwplot(rrt~condition,subset(critdata,region=="head noun+1" & rrt> -6))
> ## get coefs, SEs, t-values:
>
> gwall.de<-c(fixef(gwall.de.rrt)[2],
+       sqrt(vcov(gwall.de.rrt))[2,2],
+       fixef(gwall.de.rrt)[2]/sqrt(vcov(gwall.de.rrt))[2,2])
> gwall.hn<-c(fixef(gwall.hnoun.rrt)[2],
+       sqrt(vcov(gwall.hnoun.rrt))[2,2],
+       fixef(gwall.hnoun.rrt)[2]/sqrt(vcov(gwall.hnoun.rrt))[2,2])
> gwall.hn1<-c(fixef(gwall.hnoun1.rrt)[2],
+       sqrt(vcov(gwall.hnoun1.rrt))[2,2],
+       fixef(gwall.hnoun1.rrt)[2]/sqrt(vcov(gwall.hnoun1.rrt))[2,2])
> gwallresults<-rbind(gwall.de,gwall.hn,gwall.hn1)
> rownames(gwallresults)<-c("de","head noun","head noun+1")
> colnames(gwallresults)<-c("coef","SE","t-value")
> xtable(round(gwallresults,digits=2))
```

```
% latex table generated in R 4.0.4 by xtable 1.8-4 package
% Tue Jun 15 21:05:04 2021
\begin{table}[ht]
\centering
\begin{tabular}{rrrr}
  \hline
 & coef & SE & t-value \\
  \hline
de & -0.19 & 0.05 & -3.45 \\
  head noun & -0.10 & 0.05 & -2.15 \\
  head noun+1 & -0.01 & 0.05 & -0.16 \\
```

```
   \hline
\end{tabular}
\end{table}
```