

Reproducible analysis for: Individual flexibility in group foraging behaviour of reef manta rays (Mobula alfredi)

Raphaël Royauté

2024-02-22

Table of contents

Rationale	1
Packages and data import	2
What explains group vs. solo foraging? (TODO)	3
Null	3
What explains group size? (TODO)	4
Null	4
Model fitting	4
AIC model comparison	4
What explains variation in group leadership? (TODO)	4
Individual differences in group leadership compared between sexes, age and injury status	9
Male - female differences	9
Juvenile - adult differences	14
Injured - non-injured differences	19
Are these differences statistically significant?	23

Rationale

This is a reproducible script for all statistical models and their outputs presented in our article.

Packages and data import

Make sure to have these packages installed before running the code in this report

```
library(tidyverse); library(here); library(easystats);  
library(kableExtra); library(lme4); library(effects)  
library(marginaleffects); library(ggeffects); library(rptR)  
library(gtsummary); library(ggthemes); library(patchwork)  
library(tidybayes)
```

Import complete dataset

```
df.total = read.csv(here("data/data-clean/Manta Data_Annie.csv"),  
                    header=TRUE, sep=",", na.strings="NA", dec=".",  
                    strip.white=TRUE)  
df.total[1:10,] %>%  
  select(1:10) %>%  
  kable(digits = 2)
```

npid	id	day	year	nyear	site	clips	sex	size	maturity
P204	204	185	2016	year3	1	10	2	3	2
P204	204	185	2016	year3	1	10	2	3	2
P225	225	259	2016	year3	1	24	1	3	1
P281	281	259	2016	year3	1	24	2	3	2
P281	281	259	2016	year3	1	24	2	3	2
P302	302	259	2016	year3	1	24	1	3	1
P413	413	174	2016	year3	1	5	2	3	2
P423	423	185	2016	year3	1	10	1	4	2
P423	423	185	2016	year3	1	10	1	4	2
P423	423	185	2016	year3	1	10	1	4	2

Import group foraging dataset

```
df.group = read.csv(here("data/data-clean/group_rp.csv"),  
                    header=TRUE, sep=",", na.strings="NA", dec=".",  
                    strip.white=TRUE)  
df.group$id = as.factor(df.group$id)  
df.group$site = as.factor(df.group$site)  
df.group$sex_f = as.factor(ifelse(df.group$sex=="1", "F", "M"))  
df.group$size = as.factor(df.group$size)
```

```

df.group$maturity = as.factor(df.group$maturity)
df.group$plankton = as.factor(df.group$plankton)
df.group$shark_bite = as.factor(df.group$shark_bite)
df.group$anthropogenic = as.factor(df.group$anthropogenic)
df.group$Id = df.group$id

# Transform Time columns centered around 12:00pm expressed in hours
df.group$time_cen = hm(df.group$time)
df.group$time_cen = as.numeric(df.group$time_cen-hours(12))/3600

# subset with only individuals within groups
df.group = df.group[df.group$group==1,]
df.group=df.group[complete.cases(df.group$position),]
df.group[1:10,] %>%
  select(1:10) %>%
  kable(digits = 2)

```

npid	id	obs_nb	rep_tot	day	year	nyear	site	clips	sex
P19	19	1	2	155	2016	year3	1	46	1
P19	19	2	2	155	2016	year3	1	46	1
P33	33	1	17	153	2016	year3	1	16	2
P33	33	2	17	155	2016	year3	1	46	2
P33	33	3	17	184	2016	year3	1	25	2
P33	33	4	17	184	2016	year3	1	25	2
P33	33	5	17	184	2016	year3	1	25	2
P33	33	6	17	184	2016	year3	1	25	2
P33	33	7	17	214	2016	year3	1	8	2
P33	33	8	17	215	2016	year3	1	1	2

What explains group vs. solo foraging? (TODO)

TODO ### Models tested

Null

Abiotic Site, current, time to high tide Biotic (external) Plankton, number of mantas (scaled)
 Abiotic + Biotic (ext) Site, current, time to high tide, plankton, number of mantas (scaled)

What explains group size? (TODO)

TODO ### Models tested

Null

Abiotic Site, current, time to high tide Biotic (external) Plankton, number of mantas (scaled)
Abiotic + Biotic (ext) Site, current, time to high tide, plankton, number of mantas (scaled for control)

Model fitting

```
# glmm.group.size.null = glmer(group_size ~ 1 + (1|id),  
#                               family=poisson,  
#                               df)  
#  
# glmm.group.size.abio = glmer(group_size ~ 1 + (1|id),  
#                               family=poisson,  
#                               df)
```

AIC model comparison

What explains variation in group leadership? (TODO)

We fit a binomial GLMM on leadership probability with the following covariates:

- Sex
- Age class
- Shark injury status
- Anthropogenic injury status
- Individual ID (random effect)

```
glmm.lead=glmer(leader ~  
                 sex_f + maturity +  
                 shark_bite + anthropogenic + (1|Id),  
                 family = "binomial",  
                 df.group)  
  
summary(glmm.lead)
```

```

Generalized linear mixed model fit by maximum likelihood (Laplace
Approximation) [glmerMod]
Family: binomial ( logit )
Formula: leader ~ sex_f + maturity + shark_bite + anthropogenic + (1 | Id)
Data: df.group

```

AIC	BIC	logLik	deviance	df.resid
2152.9	2185.7	-1070.4	2140.9	1754

Scaled residuals:

Min	1Q	Median	3Q	Max
-1.318	-0.650	-0.566	1.217	2.302

Random effects:

Groups	Name	Variance	Std.Dev.
Id	(Intercept)	0.2715	0.521

Number of obs: 1760, groups: Id, 322

Fixed effects:

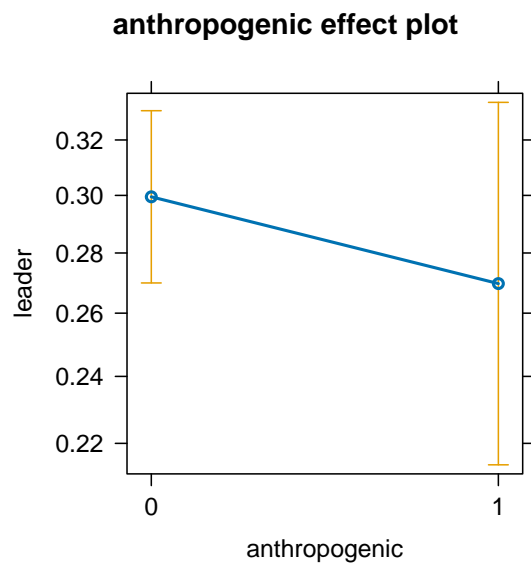
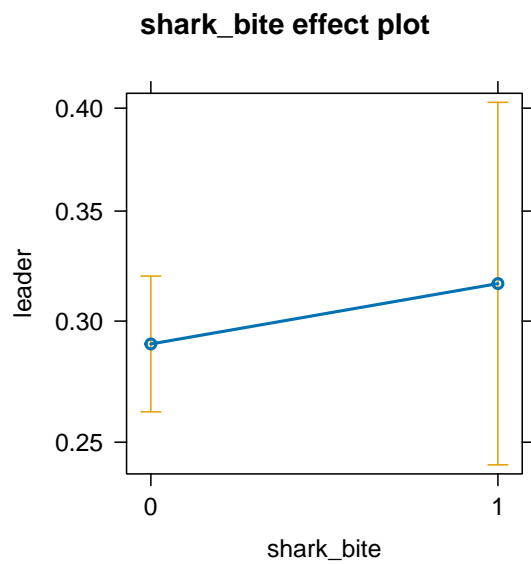
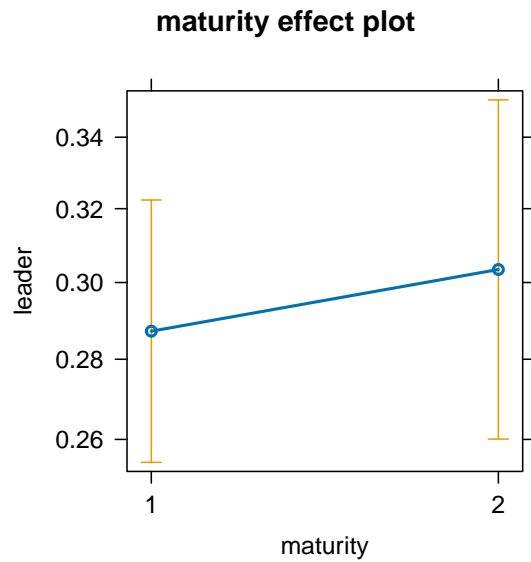
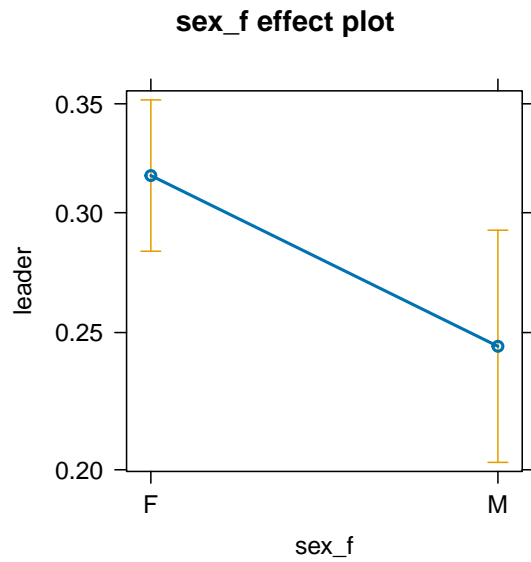
	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.78194	0.09655	-8.099	5.55e-16 ***
sex_fM	-0.35794	0.14851	-2.410	0.0159 *
maturity2	0.07813	0.13718	0.570	0.5690
shark_bite1	0.12525	0.20364	0.615	0.5385
anthropogenic1	-0.14638	0.17133	-0.854	0.3929

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:

	(Intr)	sex_fM	mtrty2	shrk_1
sex_fM		-0.310		
maturity2	-0.377	-0.282		
shark_bite1	-0.302	-0.024	0.051	
anthropgnc1	-0.305	0.007	-0.078	0.102

```
plot(allEffects(glmm.lead))
```



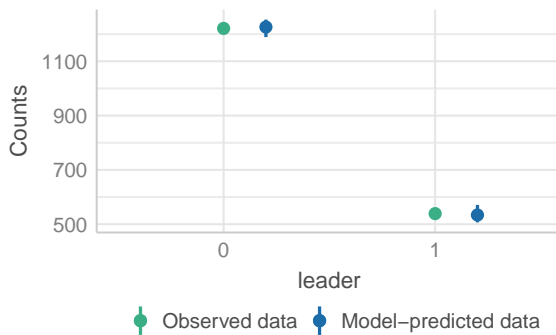
```
# Save model file
saveRDS(glmm.lead, file = here("outputs/mods/glmm.lead.rds"))
```

Inspect the model

```
check_model(glmm.lead)
```

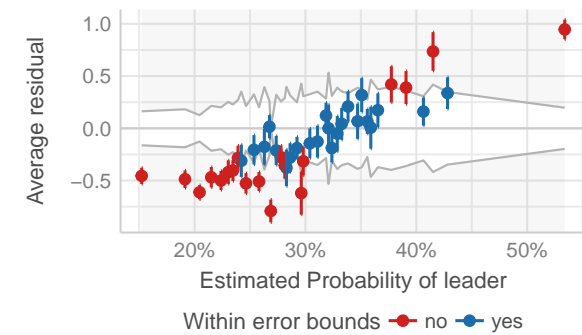
Posterior Predictive Check

Model-predicted intervals should include observed data point:



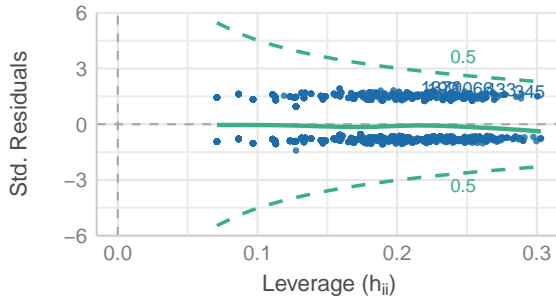
Binned Residuals

Points should be within error bounds



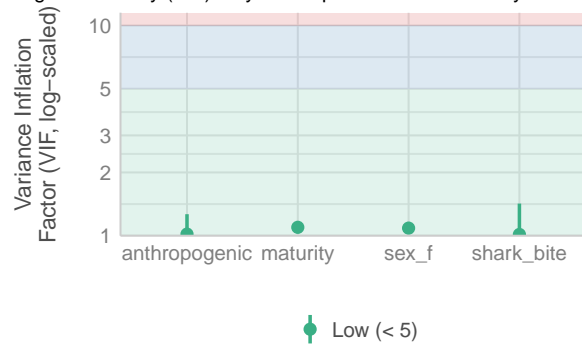
Influential Observations

Points should be inside the contour lines



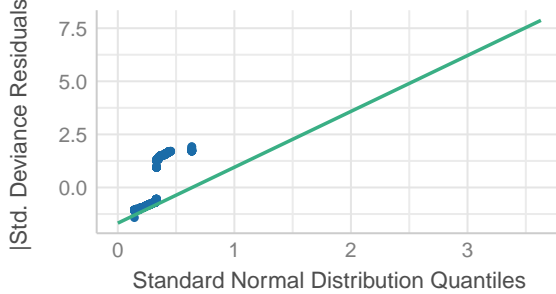
Collinearity

High collinearity (VIF) may inflate parameter uncertainty



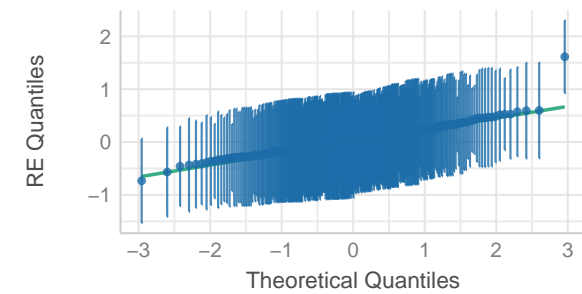
Normality of Residuals

Dots should fall along the line



Normality of Random Effects (Id)

Dots should be plotted along the line



Check R2 and repeatability

```
r2.lead = r2_nakagawa(glmm.lead, ci = T)
icc.lead = icc(glmm.lead, ci = T)
r2.lead; icc.lead

# Save files to avoid reloading
saveRDS(r2.lead, file = here("outputs/mods/r2.lead.rds"))
saveRDS(icc.lead, file = here("outputs/mods/icc.lead.rds"))
```

Get model summary in table format

```
tbl.glmm.lead = glmm.lead %>%
  tbl_regression(
    intercept = T,
    exponentiate = TRUE,
    pvalue_fun = ~ style_pvalue(.x, digits = 2)) %>%
  add_global_p() %>%
  bold_p(t = .05) %>%
  bold_labels() %>%
  italicize_levels() %>%
  add_nevent(location = "level")
tbl.glmm.lead
```

Characteristic	Event N	OR	95% CI	p-value
(Intercept)	539	0.46	0.38, 0.55	<0.001
sex_f				0.016
<i>F</i>	395	—	—	
<i>M</i>	144	0.70	0.52, 0.94	
maturity				0.57
<i>1</i>	344	—	—	
<i>2</i>	195	1.08	0.83, 1.41	
shark_bite				0.54
<i>0</i>	473	—	—	
<i>1</i>	66	1.13	0.76, 1.69	
anthropogenic				0.39
<i>0</i>	435	—	—	
<i>1</i>	104	0.86	0.62, 1.21	

We can now export the table in both word and html format

```
tbl.glmm.lead %>%
  as_gt() %>%
  gt::gtsave(filename = here("outputs/tables/tbl.glmm.lead.html"))

tbl.glmm.lead %>%
  as_gt() %>%
  gt::gtsave(filename = here("outputs/tables/tbl.glmm.lead.docx"))
```


Individual differences in group leadership compared between sexes, age and injury status

We can further investigate whether the magnitude of individual differences varies among meaningful biological categories such as sex, age or injury status. To do so, we fit one model for each subset of the data and compute the distribution of the different variance component through bootstrapping. We can then test whether among-, within-individual or fixed effect variance differs between the different subset by computing the difference between the distribution of each subset considered: ΔV . The overlap of the 95 % confidence interval of this distribution with 0 indicates whether the differences are statistically significant. The scale on which to make the comparison can be either on the latent scale with the logit link in a binomial glmm or on the original data scale. Here we estimated variances on the latent scale but calculated the difference in terms of % difference from a reference level (sex: females, maturity: juveniles, injury: not injured).

Male - female differences

Given that females are more frequently seen leading the group, it seems plausible that females are more consistent in their leadership preferences. As a result, we would expect less among-individual differences in leadership if females are highly biased toward leading the group.

Fit glmer models

```
glmm.lead.f = glmer(leader ~ maturity + shark_bite +  
  anthropogenic + (1|id),  
  family = "binomial",  
  subset(df.group, sex_f == "F"))  
glmm.lead.m = glmer(leader ~ maturity + shark_bite +  
  anthropogenic + (1|id),  
  family = "binomial",  
  subset(df.group, sex_f == "M"))  
  
# Save model file  
saveRDS(glmm.lead.f, file = here("outputs/mods/glmm.lead.f.rds"))  
saveRDS(glmm.lead.m, file = here("outputs/mods/glmm.lead.m.rds"))
```

The model does not converge properly with the male dataset. This comes from the fact that there are few adult males in the population. Because our question is now related to variance component, we keep the fixed effect structure the same regardless for consistency.

Fit rptR models

Adjusted repeatability

```
rpt.R.f = rpt(leader ~ maturity + shark_bite +
              anthropogenic + (1|id),
              grname = "id",
              datatype = "Binary",
              parallel = T,
              data = subset(df.group, sex_f == "F"))
rpt.R.m = rpt(leader ~ maturity + shark_bite +
              anthropogenic + (1|id),
              grname = "id",
              datatype = "Binary",
              parallel = T,
              data = subset(df.group, sex_f == "M"))

saveRDS(rpt.R.f, here("outputs/mods/rpt.R.f.rds"))
saveRDS(rpt.R.m, here("outputs/mods/rpt.R.m.rds"))
```

All variance components

```
rpt.V.f = rpt(leader ~ maturity + shark_bite +
              anthropogenic + (1|id),
              grname = c("id", "Fixed", "Residual"),
              datatype = c("Binary"),
              parallel = T,
              data = subset(df.group, sex_f == "F"),
              ratio = FALSE)
rpt.V.m = rpt(leader ~ maturity + shark_bite +
              anthropogenic + (1|id),
              grname = c("id", "Fixed", "Residual"),
              datatype = "Binary",
              parallel = T,
              data = subset(df.group, sex_f == "M"),
              ratio = FALSE)

saveRDS(rpt.V.f, here("outputs/mods/rpt.V.f.rds"))
saveRDS(rpt.V.m, here("outputs/mods/rpt.V.m.rds"))
```

All variance ratios

```

rpt.r2.f = rpt(leader ~ maturity + shark_bite +
               anthropogenic + (1|id),
               grname = c("id", "Fixed", "Residual"),
               datatype = c("Binary"),
               parallel = T,
               data = subset(df.group, sex_f == "F"),
               ratio = T)
rpt.r2.m = rpt(leader ~ maturity + shark_bite +
               anthropogenic + (1|id),
               grname = c("id", "Fixed", "Residual"),
               datatype = "Binary",
               parallel = T,
               data = subset(df.group, sex_f == "M"),
               ratio = T)
saveRDS(rpt.r2.f, here("outputs/mods/rpt.r2.f.rds"))
saveRDS(rpt.r2.m, here("outputs/mods/rpt.r2.m.rds"))

```

Plot the model estimates

Store all variance components

```

Vi_f = rpt.V.f$R_boot_link$id
Vi_m = rpt.V.m$R_boot_link$id
Vfe_f = rpt.V.f$R_boot_link$Fixed
Vfe_m = rpt.V.m$R_boot_link$Fixed
VR_f = rpt.V.f$R_boot_link$Residual
VR_m = rpt.V.m$R_boot_link$Residual
R_f = rpt.R.f$R_boot_link$id
R_m = rpt.R.m$R_boot_link$id

```

Put into formatted table

```

# Variance ratio difference figure
r2_Vi_f = rpt.r2.f$R_boot_link$id
r2_Vi_m = rpt.r2.m$R_boot_link$id
r2_Vfe_f = rpt.r2.f$R_boot_link$Fixed
r2_Vfe_m = rpt.r2.m$R_boot_link$Fixed
r2_VR_f = rpt.r2.f$R_boot_link$Residual
r2_VR_m = rpt.r2.m$R_boot_link$Residual

df.sex = data.frame(r2_Vi = c(r2_Vi_f, r2_Vi_m),

```

```

        r2_Vfe = c(r2_Vfe_f, r2_Vfe_m),
        r2_VR = c(r2_VR_f, r2_VR_m),
        Sex = c(rep("F", length(r2_Vi_f)),
                 rep("M", length(r2_Vi_m)))) %>%
pivot_longer(cols = r2_Vi:r2_VR,
              names_to = "v.compo",
              values_to = "var")

# Store effect sizes
df.sex.2 = data.frame(delta_r2_Vi = r2_Vi_f - r2_Vi_m,
                      delta_r2_Vfe = r2_Vfe_f - r2_Vfe_m,
                      delta_r2_VR = r2_VR_f - r2_VR_m)

```

Plot

```

p1 = df.sex %>%
  group_by(v.compo, Sex) %>%
  summarise(var = mean(var)) %>%
  ggplot(aes(y = var, x = Sex, fill = v.compo)) +
  geom_bar(position = "fill",
            stat = "identity", width = .2) +
  scale_fill_wsj(labels = c(
    bquote(V[fe]),
    bquote(V[i]),
    bquote(V[R])))) +
  ylab("Variance explained") +
  theme_bw(18) +
  labs(fill = "") +
  theme(legend.position = c(.9, .5))

delta.vi.r = df.sex.2 %>%
  ggplot(aes(x = delta_r2_Vi * 100)) +
  stat_halfeye(alpha = .6) +
  geom_vline(xintercept = 0,
             linetype = "dashed") +
  xlab(bquote(Delta[V[i]])) +
  ylab("Density") +
  theme_bw(18) +
  ggtitle("Difference in variance explained (%)")

delta.vfe.r = df.sex.2 %>%
  ggplot(aes(x = delta_r2_Vfe * 100)) +

```

```

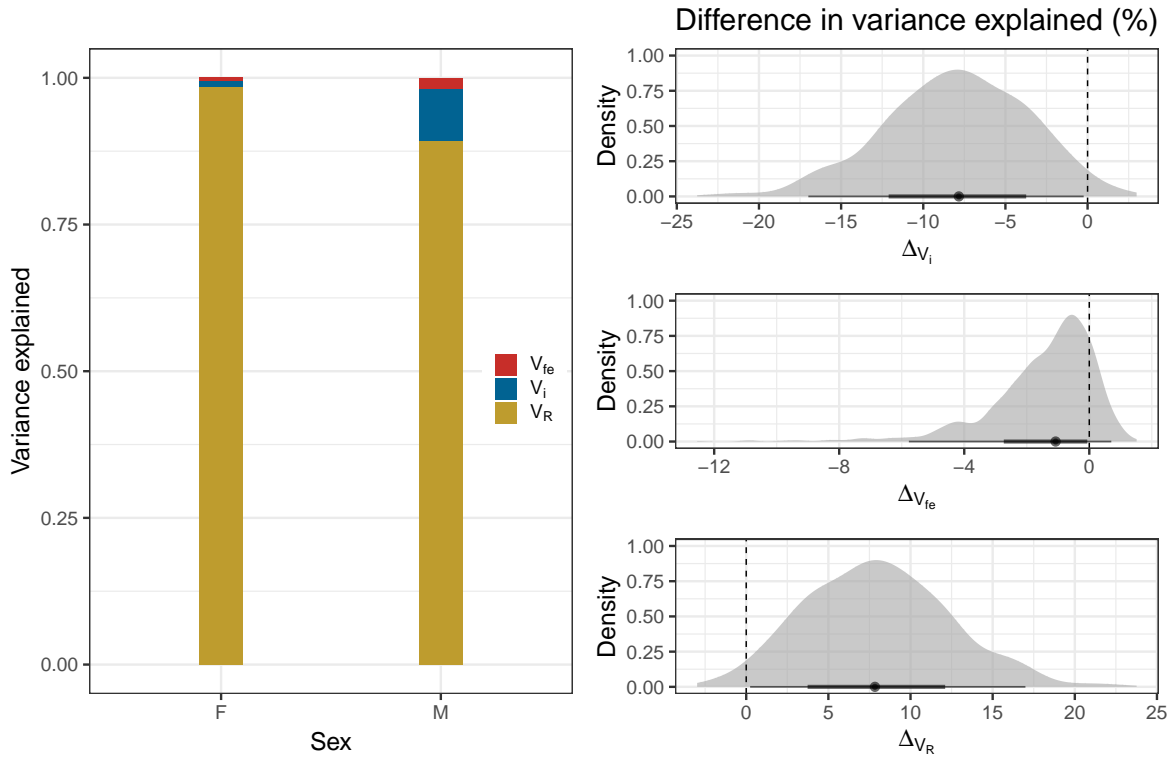
stat_halfeye(alpha = .6) +
geom_vline(xintercept = 0,
            linetype = "dashed") +
xlab(bquote(Delta[V[fe]])) +
ylab("Density") +
theme_bw(18)

delta.vr.r = df.sex.2 %>%
  ggplot(aes(x = delta_r2_VR * 100)) +
  stat_halfeye(alpha = .6) +
  geom_vline(xintercept = 0,
            linetype = "dashed") +
  xlab(bquote(Delta[V[R]])) +
  ylab("Density") +
  theme_bw(18)

delta.v.r = delta.vi.r / delta.vfe.r / delta.vr.r

delta.v.r = p1 + (delta.vi.r / delta.vfe.r / delta.vr.r)
delta.v.r

```



```
ggsave(filename = "outputs/figs/delta.v.r.sex.jpeg", delta.v.r,
        width = 12, height = 8)
ggsave(filename = "outputs/figs/delta.v.r.sex.pdf", delta.v.r,
        width = 12, height = 8)
```

Juvenile - adult differences

We hypothesized that adult individuals have more experience and are therefore more likely to lead the foraging group. Similarly to females, we expect less among-individual variation in this class compared to juveniles.

Fit glmer models

```
glmm.lead.j = glmer(leader ~ sex_f + shark_bite + anthropogenic + (1|id),
                    family = "binomial",
                    subset(df.group, maturity == "1"))
glmm.lead.a = glmer(leader ~ sex_f + shark_bite + anthropogenic + (1|id),
                    family = "binomial",
```

```
subset(df.group, maturity == "2"))

# Save model file
saveRDS(glm.lead.j, file = here("outputs/mods/glm.lead.j.rds"))
saveRDS(glm.lead.a, file = here("outputs/mods/glm.lead.a.rds"))
```

Fit rptR models

Adjusted repeatability

```
rpt.R.j = rpt(leader ~ sex_f + shark_bite + anthropogenic + (1|id),
  grname = "id",
  datatype = "Binary",
  parallel = T,
  data = subset(df.group, maturity == "1"))
rpt.R.a = rpt(leader ~ sex_f + shark_bite + anthropogenic + (1|id),
  grname = "id",
  datatype = "Binary",
  parallel = T,
  data = subset(df.group, maturity == "2"))

saveRDS(rpt.R.j, here("outputs/mods/rpt.R.j.rds"))
saveRDS(rpt.R.a, here("outputs/mods/rpt.R.a.rds"))
```

All variance components

```
rpt.V.j = rpt(leader ~ sex_f + shark_bite + anthropogenic + (1|id),
  grname = c("id", "Fixed", "Residual"),
  datatype = c("Binary"),
  parallel = T,
  data = subset(df.group, maturity == "1"),
  ratio = FALSE)
rpt.V.a = rpt(leader ~ sex_f + shark_bite + anthropogenic + (1|id),
  grname = c("id", "Fixed", "Residual"),
  datatype = "Binary",
  parallel = T,
  data = subset(df.group, maturity == "2"),
  ratio = FALSE)

saveRDS(rpt.V.j, here("outputs/mods/rpt.V.j.rds"))
saveRDS(rpt.V.a, here("outputs/mods/rpt.V.a.rds"))
```

All variance ratios

```
rpt.r2.j = rpt(leader ~ sex_f + shark_bite + anthropogenic + (1|id),
  grname = c("id", "Fixed", "Residual"),
  datatype = c("Binary"),
  parallel = T,
  data = subset(df.group, maturity == "1"),
  ratio = T)
rpt.r2.a = rpt(leader ~ sex_f + shark_bite + anthropogenic + (1|id),
  grname = c("id", "Fixed", "Residual"),
  datatype = "Binary",
  data = subset(df.group, maturity == "2"),
  ratio = T)
saveRDS(rpt.r2.j, here("outputs/mods/rpt.r2.j.rds"))
saveRDS(rpt.r2.a, here("outputs/mods/rpt.r2.a.rds"))
```

Plot the model estimates

Store all variance components

```
Vi_j = rpt.V.j$R_boot_link$id
Vi_a = rpt.V.a$R_boot_link$id
Vfe_j = rpt.V.j$R_boot_link$Fixed
Vfe_a = rpt.V.a$R_boot_link$Fixed
VR_j = rpt.V.j$R_boot_link$Residual
VR_a = rpt.V.a$R_boot_link$Residual
R_j = rpt.R.j$R_boot_link$id
R_a = rpt.R.a$R_boot_link$id
```

Put into formatted table

```
# Variance ratio difference figure
r2_Vi_j = rpt.r2.j$R_boot_link$id
r2_Vi_a = rpt.r2.a$R_boot_link$id
r2_Vfe_j = rpt.r2.j$R_boot_link$Fixed
r2_Vfe_a = rpt.r2.a$R_boot_link$Fixed
r2_VR_j = rpt.r2.j$R_boot_link$Residual
r2_VR_a = rpt.r2.a$R_boot_link$Residual

df.mat = data.frame(r2_Vi = c(r2_Vi_j, r2_Vi_a),
  r2_Vfe = c(r2_Vfe_j, r2_Vfe_a),
  r2_VR = c(r2_VR_j, r2_VR_a),
```



```

      Maturity = c(rep("Juveniles", length(r2_Vi_j)),
                    rep("Adults", length(r2_Vi_a)))) %>%
pivot_longer(cols = r2_Vi:r2_VR,
              names_to = "v.compo",
              values_to = "var")

# Store effect sizes
df.mat.2 = data.frame(delta_r2_Vi = r2_Vi_j - r2_Vi_a,
                      delta_r2_Vfe = r2_Vfe_j - r2_Vfe_a,
                      delta_r2_VR = r2_VR_j - r2_VR_a)

```

Plot

```

p1 = df.mat %>%
  group_by(v.compo, Maturity) %>%
  summarise(var = mean(var)) %>%
  ggplot(aes(y = var, x = Maturity, fill = v.compo)) +
  geom_bar(position = "fill",
            stat = "identity", width = .2) +
  scale_fill_wsj(labels = c(
    bquote(V[fe]),
    bquote(V[i]),
    bquote(V[R])))) +
  ylab("Variance explained") +
  theme_bw(18) +
  labs(fill = "") +
  theme(legend.position = c(.9, .5))

delta.vi.r = df.mat.2 %>%
  ggplot(aes(x = delta_r2_Vi * 100)) +
  stat_halfeye(alpha = .6) +
  geom_vline(xintercept = 0,
             linetype = "dashed") +
  xlab(bquote(Delta[V[i]])) +
  ylab("Density") +
  theme_bw(18) +
  ggtitle("Difference in variance explained (%)")

delta.vfe.r = df.mat.2 %>%
  ggplot(aes(x = delta_r2_Vfe * 100)) +
  stat_halfeye(alpha = .6) +
  geom_vline(xintercept = 0,

```

```

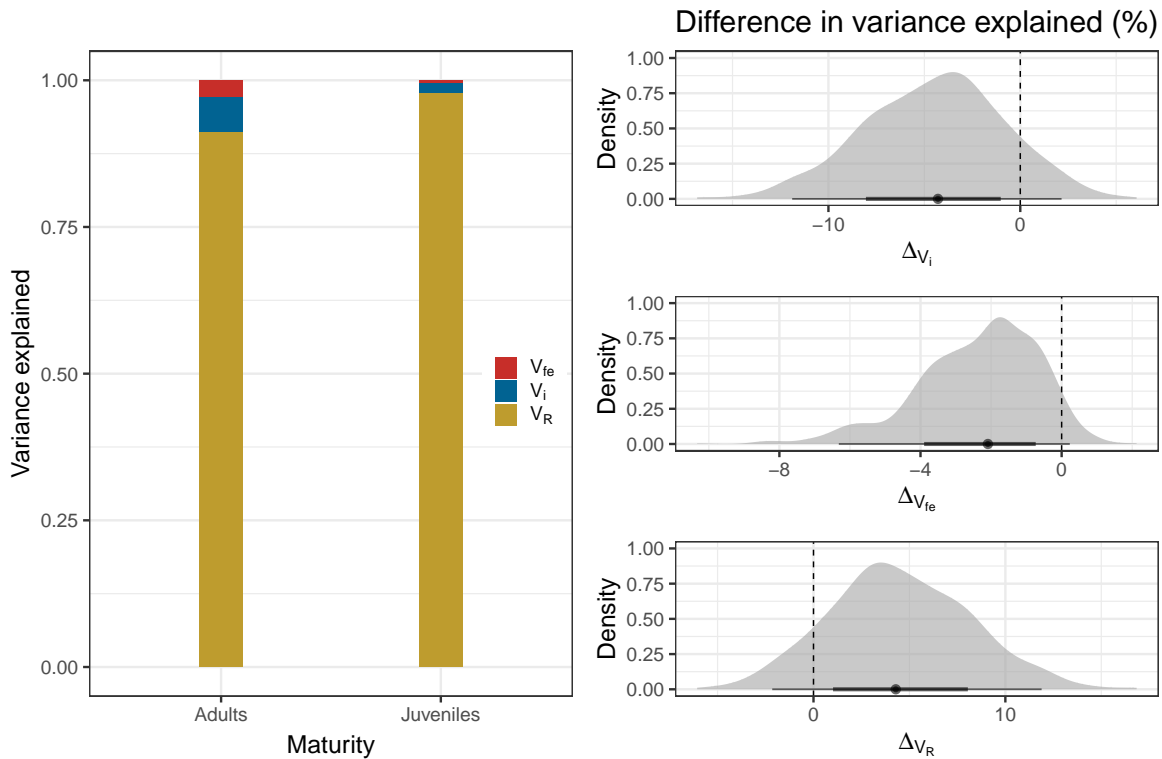
        linetype = "dashed") +
  xlab(bquote(Delta[V[fe]])) +
  ylab("Density") +
  theme_bw(18)

delta.vr.r = df.mat.2 %>%
  ggplot(aes(x = delta_r2_VR * 100)) +
  stat_halfeye(alpha = .6) +
  geom_vline(xintercept = 0,
            linetype = "dashed") +
  xlab(bquote(Delta[V[R]])) +
  ylab("Density") +
  theme_bw(18)

delta.v.r = delta.vi.r / delta.vfe.r / delta.vr.r

delta.v.r = p1 + (delta.vi.r / delta.vfe.r / delta.vr.r)
delta.v.r

```



```
ggsave(filename = "outputs/figs/delta.v.r.maturity.jpeg", delta.v.r,
        width = 12, height = 8)
ggsave(filename = "outputs/figs/delta.v.r.maturity.pdf", delta.v.r,
        width = 12, height = 8)
```

Injured - non-injured differences

Here we're considering combining both anthropogenic and shark bite injury with the hypothesis that injured individuals have less chance of being at the front of the group and therefore should be more consistent in their positioning in the group.

Fit glmer models

```
glmm.lead.ni = glmer(leader ~ sex_f + maturity + (1|id),
                     family = "binomial",
                     subset(df.group, injury == "0"))
glmm.lead.i = glmer(leader ~ sex_f + maturity + (1|id),
                    family = "binomial",
                    subset(df.group, injury == "1"))

# Save model file
saveRDS(glmm.lead.ni, file = here("outputs/mods/glmm.lead.ni.rds"))
saveRDS(glmm.lead.i, file = here("outputs/mods/glmm.lead.i.rds"))
```

Fit rptR models

Adjusted repeatability

```
rpt.R.ni = rpt(leader ~ sex_f + maturity + (1|id),
               grname = "id",
               datatype = "Binary",
               parallel = T,
               data = subset(df.group, injury == "0"))
rpt.R.i = rpt(leader ~ sex_f + maturity + (1|id),
              grname = "id",
              datatype = "Binary",
              parallel = T,
              data = subset(df.group, injury == "1"))
```

```
saveRDS(rpt.R.ni, here("outputs/mods/rpt.R.ni.rds"))
saveRDS(rpt.R.i, here("outputs/mods/rpt.R.i.rds"))
```

All variance components

```
rpt.V.ni = rpt(leader ~ sex_f + maturity + (1|id),
               grname = c("id", "Fixed", "Residual"),
               datatype = c("Binary"),
               parallel = T,
               data = subset(df.group, injury == "0"),
               ratio = FALSE)
rpt.V.i = rpt(leader ~ sex_f + maturity + (1|id),
              grname = c("id", "Fixed", "Residual"),
              datatype = "Binary",
              parallel = T,
              data = subset(df.group, injury == "1"),
              ratio = FALSE)
saveRDS(rpt.V.ni, here("outputs/mods/rpt.V.ni.rds"))
saveRDS(rpt.V.i, here("outputs/mods/rpt.V.i.rds"))
```

All variance ratios

```
rpt.r2.ni = rpt(leader ~ sex_f + maturity + (1|id),
                grname = c("id", "Fixed", "Residual"),
                datatype = c("Binary"),
                parallel = T,
                data = subset(df.group, injury == "0"),
                ratio = T)
rpt.r2.i = rpt(leader ~ sex_f + maturity + (1|id),
               grname = c("id", "Fixed", "Residual"),
               datatype = "Binary",
               data = subset(df.group, injury == "1"),
               ratio = T)
saveRDS(rpt.r2.ni, here("outputs/mods/rpt.r2.ni.rds"))
saveRDS(rpt.r2.i, here("outputs/mods/rpt.r2.i.rds"))
```

Plot the model estimates

Store all variance components

```

Vi_ni = rpt.V.ni$R_boot_link$id
Vi_i = rpt.V.i$R_boot_link$id
Vfe_ni = rpt.V.ni$R_boot_link$Fixed
Vfe_i = rpt.V.i$R_boot_link$Fixed
VR_ni = rpt.V.ni$R_boot_link$Residual
VR_i = rpt.V.i$R_boot_link$Residual
R_ni = rpt.R.ni$R_boot_link$id
R_i = rpt.R.i$R_boot_link$id

```

Put into formatted table

```

# Variance ratio difference figure
r2_Vi_ni = rpt.r2.ni$R_boot_link$id
r2_Vi_i = rpt.r2.i$R_boot_link$id
r2_Vfe_ni = rpt.r2.ni$R_boot_link$Fixed
r2_Vfe_i = rpt.r2.i$R_boot_link$Fixed
r2_VR_ni = rpt.r2.ni$R_boot_link$Residual
r2_VR_i = rpt.r2.i$R_boot_link$Residual

df.inj = data.frame(r2_Vi = c(r2_Vi_ni, r2_Vi_i),
                    r2_Vfe = c(r2_Vfe_ni, r2_Vfe_i),
                    r2_VR = c(r2_VR_ni, r2_VR_i),
                    Injury = c(rep("None", length(r2_Vi_ni)),
                               rep("Injured", length(r2_Vi_i)))) %>%
  pivot_longer(cols = r2_Vi:r2_VR,
               names_to = "v.compo",
               values_to = "var")

# Store effect sizes
df.inj.2 = data.frame(delta_r2_Vi = r2_Vi_ni - r2_Vi_i,
                      delta_r2_Vfe = r2_Vfe_ni - r2_Vfe_i,
                      delta_r2_VR = r2_VR_ni - r2_VR_i)

```

Plot

```

p1 = df.inj %>%
  group_by(v.compo, Injury) %>%
  summarise(var = mean(var)) %>%
  ggplot(aes(y = var, x = Injury, fill = v.compo)) +
  geom_bar(position = "fill",
           stat = "identity", width = .2) +

```

```

scale_fill_ws(j(labels = c(
  bquote(V[fe]),
  bquote(V[i]),
  bquote(V[R])))) +
ylab("Variance explained") +
theme_bw(18) +
labs(fill = "") +
theme(legend.position = c(.9, .5))

delta.vi.r = df.inj.2 %>%
  ggplot(aes(x = delta_r2_Vi * 100)) +
  stat_halfeye(alpha = .6) +
  geom_vline(xintercept = 0,
             linetype = "dashed") +
  xlab(bquote(Delta[V[i]])) +
  ylab("Density") +
  theme_bw(18) +
  ggtitle("Difference in variance explained (%)")

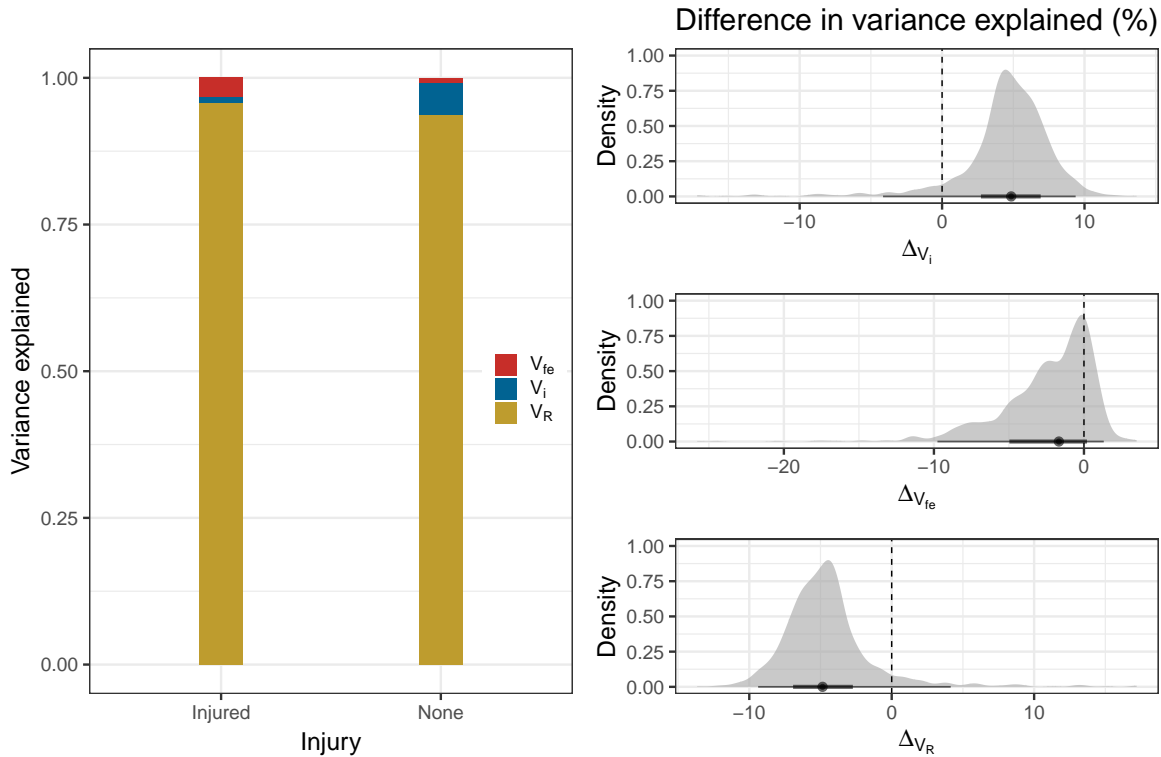
delta.vfe.r = df.inj.2 %>%
  ggplot(aes(x = delta_r2_Vfe * 100)) +
  stat_halfeye(alpha = .6) +
  geom_vline(xintercept = 0,
             linetype = "dashed") +
  xlab(bquote(Delta[V[fe]])) +
  ylab("Density") +
  theme_bw(18)

delta.vr.r = df.inj.2 %>%
  ggplot(aes(x = delta_r2_VR * 100)) +
  stat_halfeye(alpha = .6) +
  geom_vline(xintercept = 0,
             linetype = "dashed") +
  xlab(bquote(Delta[V[R]])) +
  ylab("Density") +
  theme_bw(18)

delta.v.r = delta.vi.r / delta.vfe.r / delta.vr.r

delta.v.r = p1 + (delta.vi.r / delta.vfe.r / delta.vr.r)
delta.v.r

```



```
ggsave(filename = "outputs/figs/delta.v.r.injury.jpeg", delta.v.r,
        width = 12, height = 8)
ggsave(filename = "outputs/figs/delta.v.r.injury.pdf", delta.v.r,
        width = 12, height = 8)
```

Are these differences statistically significant?

We can calculate the median and 95 % CI for the distribution of the difference between males and females. I'm using the `describe_posterior()` function from the `bayestestR` package here as it has nice option to summarize distributions. A positive value indicates that females are more variable for a given variance component expressed in % points relative to the total variance. The width of the 95 % CI indicates the precision of our estimates and the 'probability of direction' (pd) indicates the probability that the effect differs from 0. This metric is based on the proportion of values that are of the same sign as the median. pd values above 0.975 are roughly equivalent to a statistically-significant p-value with $\alpha = 0.05$ for a two-sided test.

```
tbl_delta_r2_sex = describe_posterior(df.sex.2) %>%
  data.frame() %>%
```

```

mutate("Variance_component" = as.factor(Parameter)) %>%
mutate(Variance_component = fct_recode(Variance_component,
    Delta_Vi = "delta_r2_Vi",
    Delta_Vfe = "delta_r2_Vfe",
    Delta_VR = "delta_r2_VR")) %>%
select(c(11, 2, 4 :6)) %>%
kable(digits = 2)

tbl_delta_r2_maturity = describe_posterior(df.mat.2) %>%
data.frame() %>%
mutate("Variance_component" = as.factor(Parameter)) %>%
mutate(Variance_component = fct_recode(Variance_component,
    Delta_Vi = "delta_r2_Vi",
    Delta_Vfe = "delta_r2_Vfe",
    Delta_VR = "delta_r2_VR")) %>%
select(c(11, 2, 4 :6)) %>%
kable(digits = 2)

tbl_delta_r2_injury = describe_posterior(df.inj.2) %>%
data.frame() %>%
mutate("Variance_component" = as.factor(Parameter)) %>%
mutate(Variance_component = fct_recode(Variance_component,
    Delta_Vi = "delta_r2_Vi",
    Delta_Vfe = "delta_r2_Vfe",
    Delta_VR = "delta_r2_VR")) %>%
select(c(11, 2, 4 :6)) %>%
kable(digits = 2)

tbl_delta_r2_sex; tbl_delta_r2_maturity; tbl_delta_r2_injury

```

	Variance_component	Median	CI_low	CI_high	pd
2	Delta_Vi	-0.08	-0.17	0.00	0.98
1	Delta_Vfe	-0.01	-0.06	0.01	0.85
3	Delta_VR	0.08	0.00	0.17	0.98

	Variance_component	Median	CI_low	CI_high	pd
2	Delta_Vi	-0.04	-0.12	0.02	0.89
1	Delta_Vfe	-0.02	-0.06	0.00	0.96
3	Delta_VR	0.04	-0.02	0.12	0.89

	Variance_component	Median	CI_low	CI_high	pd
2	Delta_Vi	0.05	-0.04	0.09	0.93
1	Delta_Vfe	-0.02	-0.10	0.01	0.78
3	Delta_VR	-0.05	-0.09	0.04	0.93