Repeated measures and the GLM

Professor Andy Field

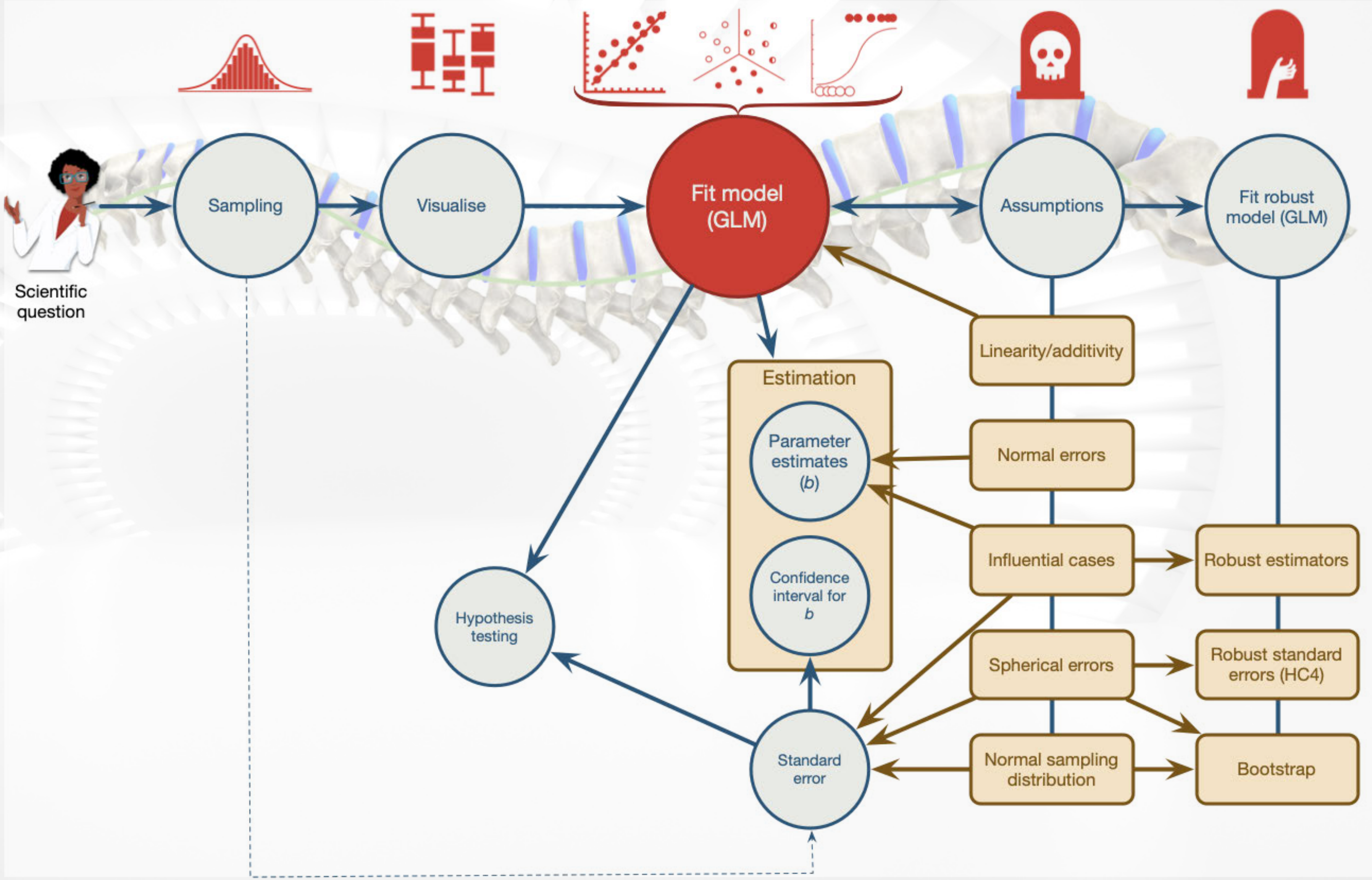
@profandyfield

www.youtube.com/user/ProfAndyField/

www.discoveringstatistics.com

www.milton-the-cat.rocks

www.discovr.rocks



Scientific question
Sampling
Visualise
Fit model (GLM)
Assumptions
Fit robust model (GLM)
Linearity/additivity
Estimation
Parameter estimates (b)
Normal errors
Influential cases
Robust estimators
Confidence interval for b
Hypothesis testing
Spherical errors
Robust standard errors (HC4)
Standard error
Normal sampling distribution
Bootstrap

ANDY FIELD

Experimental conditions

Sniff humans

Sniff aliens

Participants

Outcome

Vocalisations while sniffing

- Systematic variance: created by our manipulation
- Unsystematic variance: variance created by unknown factors

Benefits of repeated measures designs

- Sensitivity
  - Unsystematic variance is reduced
  - More sensitive to experimental effects
- Economy
  - Less participants are needed
  - But, be careful of fatigue

Can puppies sniff out aliens?

- Is there a link between video games and aggression?
  - Outcome = vocalisations during 1 min sniffing (vocalisations)
  - Predictor: type of entity being sniffed (entity)
    - Alien (not in humanoid form)
    - Human (control for alien vs human)
    - Mannequin (control for humanoid form)
    - Shapeshifter (alien in humanoid form)
  - dog_name indicates the name of the dog (*N* = 8)

The data

	dog_name	Alien	Human	Mannequin	Shapeshifter	Mean	Variance
	Milton	8	7	1	6	5.50	9.67
	Woofy	9	5	2	5	5.25	8.25
	Ramsey	6	2	3	8	4.75	7.58
	Mr. Snifficus III	5	3	1	9	4.50	11.67
	Willock	8	4	5	8	6.25	4.25
	The Venerable Dr. Waggy	7	5	6	7	6.25	0.92
	Lord Scenticle	10	2	7	2	5.25	15.58
	Professor Nose	12	6	8	1	6.75	20.92
Mean	—	8.12	4.25	4.12	5.75	—	—

The data in R

Table 2: Data for the sniffer dog example

	dog_name	entity	vocalisations
1	Milton	Alien	8
2	Milton	Human	7
3	Milton	Mannequin	1
4	Milton	Shapeshifter	6
5	Woofy	Alien	9
6	Woofy	Human	5
7	Woofy	Mannequin	2
8	Woofy	Shapeshifter	5
9	Ramsey	Alien	6
10	Ramsey	Human	2

Repeated measures and the linear model

- Same participants in all conditions
  - Scores across conditions correlate
  - Violates the assumption of independent residuals (think back to the lecture on bias)
- Need to adjust the model to estimate this dependency

$$\widehat{\text{vocalisations}}_{ij} = \hat{b}_{0j} + \hat{b}_{1j}\text{entity}_{ij} + e_{ij}$$
$$\hat{b}_{0j} = \hat{b}_0 + \hat{u}_{0j}$$
$$\hat{b}_{1j} = \hat{b}_1 + \hat{u}_{1j}$$

- Remember that entity would actually be split into 3 dummy/contrast variables

Approaches to repeated measures designs and the GLM

1. Assume sphericity
    - Estimate it
    - Correct for it (adjust the degrees of freedom)
2. Fit a different kind of model (a *multilevel growth model*)
    - Can model different kinds of dependency in errors

What is sphericity, ϵ ?

	Alien-Human	Alien-Mannequin	Alien-Shapeshifter	Human-Mannequin	Human-Shapeshifter	Mannequin-Shapeshifter
Milton	1	7	2	6	1	-5
Woofy	4	7	4	3	0	-3
Ramsey	4	3	-2	-1	-6	-5
Mr. Snifficus III	2	4	-4	2	-6	-8
Willock	4	3	0	-1	-4	-3
The Venerable Dr. Waggy	2	1	0	-1	-2	-1
Lord Scenticle	8	3	8	-5	0	5
Professor Nose	6	4	11	-2	5	7
Variance	5.27	4.29	25.70	11.55	14.29	26.55

The assumption of sphericity, ϵ

- What is it?
  - The differences between pairs of groups should have equal variances
- How is it estimated?
  - Greenhouse-Geisser estimate, $\hat{\epsilon}$
  - Huynh-Feldt estimate, $\tilde{\epsilon}$
  - If $\epsilon = 1$, sphericity is perfect
  - If $\epsilon < 1$, sphericity is violated (to some degree)

Estimating sphericity, ϵ

$$\hat{\epsilon} = \frac{k^2\left(\bar{D} - \overline{\text{cov}}_T\right)^2}{(k-1)\left(\sum \text{cov}_{ij}^2 - 2k\sum \overline{\text{cov}}_i^2 + k^2\text{cov}_T^2\right)}$$

$$\frac{1}{(k-1)} \le \hat{\epsilon} \le 1$$

What to do about sphericity?

- (R can) multiply *df* by these estimates to correct for the effect of sphericity
- Given that ϵ quantifies the deviation from perfect sphericity, we correct the *df* by the degree to which spehercity is violated
- *df* get smaller making it harder for the test statistic to be significant
- Routinely apply the G-G correction and forget about sphericity

Fitting the model

- The `afex::aov_4()` function
  - Specify the repeated measures with `(rm_predictors|id_var)`
  - Automatically sets contrasts
  - Built in interaction plot with `afex_plot()`
  - But … no parameter estimates, diagnostic plots, or robust methods

```
sniff_afx <- afex::aov_4(vocalisations ~
entity + (entity|dog_name), data = sniff_tib)
```

Fitting the model

- The `afex::aov_4()` function
  - Specify the repeated measures with `(rm_predictors|id_var)`
  - Automatically sets contrasts
  - Built in interaction plot with `afex_plot()`
  - But … no parameter estimates, diagnostic plots, or robust methods

```
sniff_afx <- afex::aov_4(vocalisations ~
entity + (entity|dog_name), data = sniff_tib)
```

- Use `lmer()` from the `lme4` package
  - A trickier option (we don't teach you it)
  - Manually set contrasts
  - Can get parameter estimates, diagnostic plots, and robust methods

Contrasts

If the dog training has been successful then we'd expect sniffer dogs to make more vocalisations when sniffing alien entities than non alien-entities.

- Contrast 1: {alien, shapeshifter} vs. {human, mannequin}

We have two 'chunks' in contrast 1 that would then need to be decomposed:

- Contrast 2: {alien} vs. {shapeshifter}
- Contrast 3: {human} vs. {mannequin}

Using the rules for contrast coding we'd get the codes in Table 4:

Table 4: Contrast coding for the entity variable

Group	Contrast 1	Contrast 2	Contrast 3
Alien	1/2	1/2	0
Human	-1/2	0	1/2
Mannequin	-1/2	0	-1/2
Shapeshifter	1/2	-1/2	0

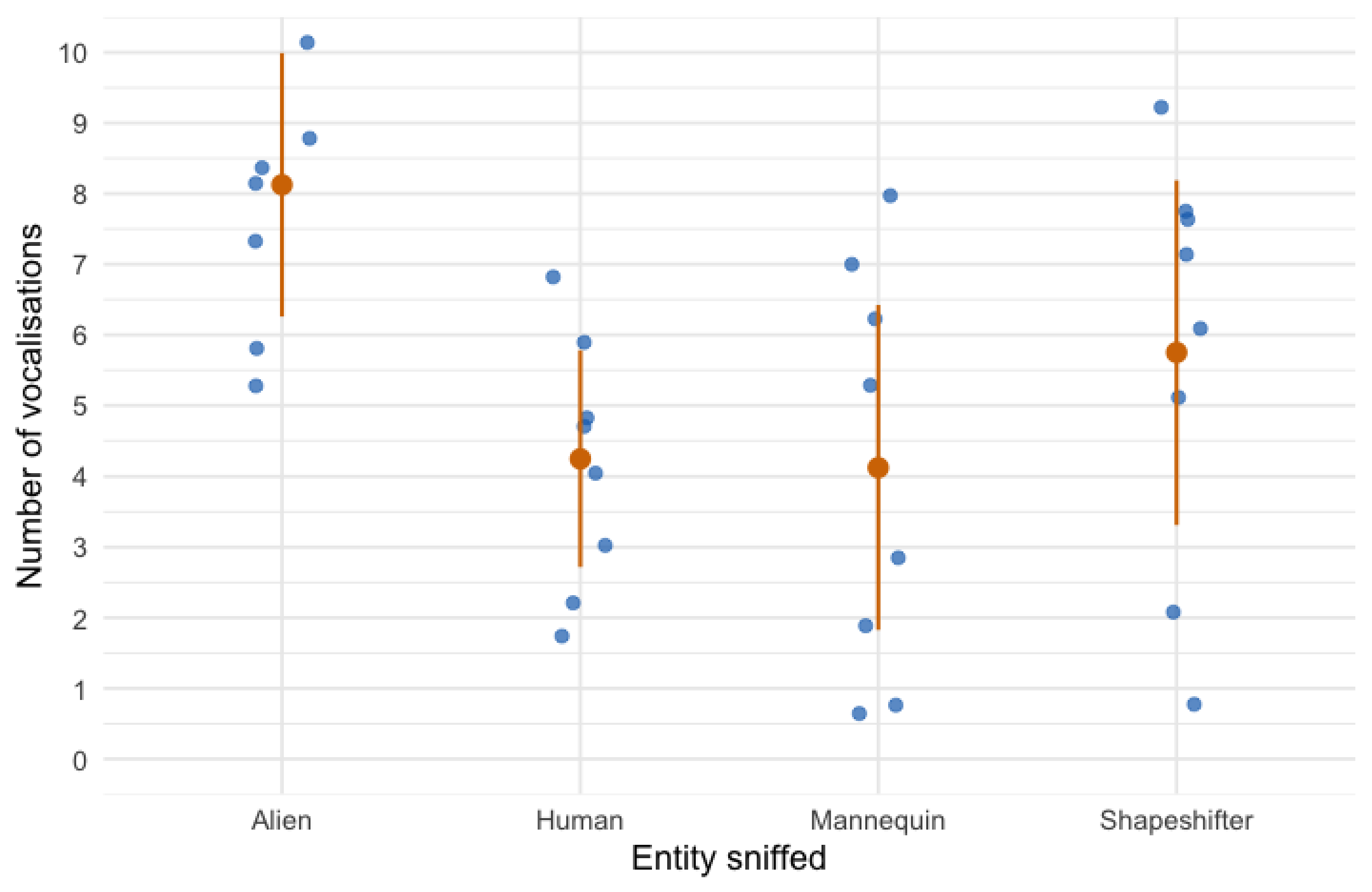
Overall model summary

```
sniff_afx <- afex::aov_4(vocalisations ~ entity + (entity|dog_name), data = sniff_tib)
sniff_afx
```

```
## Anova Table (Type 3 tests)
##
## Response: vocalisations
##   Effect           df   MSE      F  ges p.value
## 1 entity 1.60, 11.19 13.71 3.79 + .327    .063
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
##
## Sphericity correction method: GG
```

✏ The entity sniffed did not have a significant effect on the number of vocalisations by sniffer dogs, $F(1.60, 11.19) = 3.79$, $p = 0.063$. The effect size suggested that the entity sniffed accounted for about a third of the variance in vocalisations, $\eta^2_G = 0.33$.

Interpretation



Contrasts

```
sniff_emm <- emmeans::emmeans(sniff_afx, ~entity, model = "multivariate")
sniff_cons <- list(
  aliens_vs_non = c(1/2, -1/2, -1/2, 1/2),
  alien_vs_shape = c(1/2, 0, 0, -1/2),
  human_vs_manquin = c(0, 1/2, -1/2, 0)
  )
emmeans::contrast(sniff_emm, sniff_cons)
```

contrast	estimate	SE	df	t.ratio	p.value
aliens_vs_non	2.750	0.641	7	4.291	0.004
alien_vs_shape	1.187	0.896	7	1.325	0.227
human_vs_manquin	0.062	0.601	7	0.104	0.920

Post hoc tests

```
pairs(sniff_emm, adjust = "holm")
```

contrast	estimate	SE	df	t.ratio	p.value
Alien - Human	3.875	0.811	7	4.775	0.010
Alien - Mannequin	4.000	0.732	7	5.465	0.006
Alien - Shapeshifter	2.375	1.792	7	1.325	0.907
Human - Mannequin	0.125	1.202	7	0.104	0.920
Human - Shapeshifter	-1.500	1.336	7	-1.122	0.907
Mannequin - Shapeshifter	-1.625	1.822	7	-0.892	0.907

Robust models

```r
WRS2::rmanova(
y = sniff_tib$vocalisations,
groups = sniff_tib$entity,
blocks = sniff_tib$dog_name
)
```

```
## Call:
## WRS2::rmanova(y = sniff_tib$vocalisations, groups = sniff_tib$entity,
##     blocks = sniff_tib$dog_name)
##
## Test statistic: F = 2.7528
## Degrees of freedom 1: 2.31
## Degrees of freedom 2: 11.55
## p-value: 0.1002
```

Robust *post hoc* tests

```
WRS2::rmmcp(
  y = sniff_tib$vocalisations,
  groups = sniff_tib$entity,
  blocks = sniff_tib$dog_name
  )
```

```
## Call:
## WRS2::rmmcp(y = sniff_tib$vocalisations, groups = sniff_tib$entity,
##     blocks = sniff_tib$dog_name)
##
##                               psihat  ci.lower ci.upper p.value  p.crit   sig
## Alien vs. Human              3.66667  -0.48300  7.81633 0.01360 0.01020 FALSE
## Alien vs. Mannequin          4.00000  -0.35728  8.35728 0.01172 0.00851 FALSE
## Alien vs. Shapeshifter       2.00000  -8.09920 12.09920 0.44148 0.01690 FALSE
## Human vs. Mannequin          0.00000  -5.38802  5.38802 1.00000 0.05000 FALSE
## Human vs. Shapeshifter      -1.83333  -9.23266  5.56599 0.34371 0.01270 FALSE
## Mannequin vs. Shapeshifter -2.00000 -12.54827  8.54827 0.46001 0.02500 FALSE
```

Scenting a victory … factorial repeated measures designs

ANDY FIELD

Can scents distract the sniffer dogs?

- 50 sniffer dogs
  - Participated in all conditions
  - Sniffed 9 different 'things'
- Predictor: entity
  - Human: the dog sniffs a human
  - Shapeshifter the dog sniffs an alien in humanoid form
  - Alien the dog sniffs an alien in lizard form
- Predictor: scent_mask
  - The entity had no masking scent (none)
  - The entity was smeared with human pheromones
  - The entity was smeared with fox pheromones
- Outcome:
  - Number of vocalisations during each 1 minute sniff

The model

- Let's simplify things by ignoring the fact that entity and scent_mask will be represented by two dummy variables each (and the interaction by 4!)
- We can model individual differences in all parameters

$$
\begin{aligned}
\widehat{\text{vocalisations}}_{ij} &= \hat{b}_{0j} + \hat{b}_{1j}\text{entity}_{ij} + \hat{b}_{2j}\text{scent}_{ij} + \hat{b}_{3j}(\text{entity}_{ij} \times \text{scent}_{ij}) + e_{ij} \\
\hat{b}_{0j} &= \hat{b}_0 + \hat{u}_{0j} \\
\hat{b}_{1j} &= \hat{b}_1 + \hat{u}_{1j} \\
\hat{b}_{2j} &= \hat{b}_2 + \hat{u}_{2j} \\
\hat{b}_{3j} &= \hat{b}_3 + \hat{u}_{3j}
\end{aligned}
$$

- This model may be too complex to fit
- The simplest version of the repeated measures model instead treats the effects of predictor variables as fixed, but acknowledges that dogs, overall, will vary in their vocalisations:

$$
\begin{aligned}
\widehat{\text{vocalisations}}_{ij} &= \hat{b}_{0j} + \hat{b}_{1}\text{entity}_{ij} + \hat{b}_{2}\text{scent}_{ij} + \hat{b}_{3}(\text{entity}_{ij} \times \text{scent}_{ij}) + e_{ij} \\
\hat{b}_{0j} &= \hat{b}_0 + \hat{u}_{0j}
\end{aligned}
$$

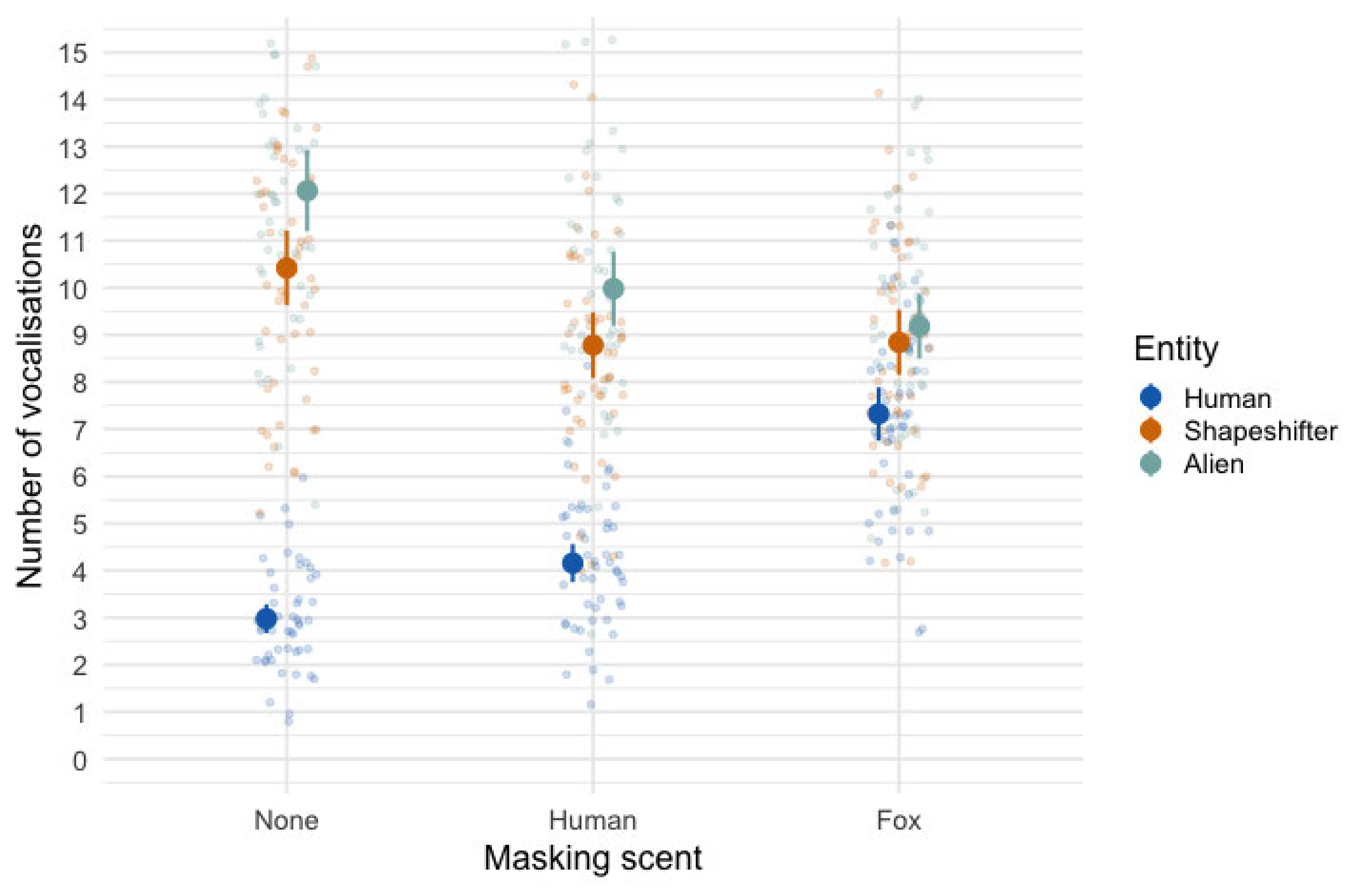
The data

Table 7: Data for the scent masking example

	dog_id	entity	scent_mask	vocalisations
1	56f9p	Alien	Fox	7
2	56f9p	Alien	Human	9
3	56f9p	Alien	None	8
4	56f9p	Shapeshifter	Fox	11
5	56f9p	Shapeshifter	Human	6
6	56f9p	Shapeshifter	None	8
7	56f9p	Human	Fox	3
8	56f9p	Human	Human	3
9	56f9p	Human	None	3
10	2m89y	Alien	Fox	8

Previous 1 2 3 4 5 ... 45 Next

The data



Number of vocalisations
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
None
Human
Fox
Masking scent
Entity
Human
Shapeshifter
Alien

Overall model summary

```
scent_afx <- afex::aov_4(vocalisations ~ entity*scent_mask + (entity*scent_mask|dog_id), data =
scent_tib)
scent_afx
```

```
## Anova Table (Type 3 tests)
##
## Response: vocalisations
##              Effect            df  MSE          F  ges p.value
## 1            entity  1.98, 96.88 4.23 315.95 *** .522   <.001
## 2        scent_mask  1.95, 95.75 2.71  12.91 *** .027   <.001
## 3 entity:scent_mask 3.66, 179.27 3.37  60.26 *** .235   <.001
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
##
## Sphericity correction method: GG
```

Overall model summary


```
scent_afx <- afex::aov_4(vocalisations ~ entity*scent_mask + (entity*scent_mask|dog_id), data =
scent_tib)
scent_afx
```

```
## Anova Table (Type 3 tests)
##
## Response: vocalisations
##                Effect          df  MSE          F  ges p.value
## 1              entity  1.98, 96.88 4.23 315.95 *** .522   <.001
## 2          scent_mask  1.95, 95.75 2.71  12.91 *** .027   <.001
## 3 entity:scent_mask 3.66, 179.27 3.37  60.26 *** .235   <.001
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
##
## Sphericity correction method: GG
```

Repeat the following mantra:

"It is never sensible to interpret main effects in the presence of a significant interaction effect."

Overall model summary

```
## Anova Table (Type 3 tests)
##
## Response: vocalisations
##               Effect           df  MSE          F  ges p.value
## 1             entity  1.98, 96.88 4.23 315.95 *** .522   <.001
## 2         scent_mask  1.95, 95.75 2.71  12.91 *** .027   <.001
## 3 entity:scent_mask 3.66, 179.27 3.37  60.26 *** .235   <.001
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
##
## Sphericity correction method: GG
```

✎ The interaction effect suggests that the effect of entity on vocalisations was significantly moderated by what scent the entity was wearing, *F*(3.66, 179.27) = 60.26, *p* < .001. The effect size suggested that the interaction accounted for about 24% of the variance in vocalisations not accounted for by other variables, $\eta^2_G = 0.24$.

Simple effects analysis

The effect of entity within type of scent

```
emmeans::joint_tests(scent_afx, "scent_mask")
```

```
## scent_mask = None:
##  model term df1   df2 F.ratio p.value
##  entity       2 287.5 339.626 <.0001
##
## scent_mask = Human:
##  model term df1   df2 F.ratio p.value
##  entity       2 287.5 136.965 <.0001
##
## scent_mask = Fox:
##  model term df1   df2 F.ratio p.value
##  entity       2 287.5  14.228 <.0001
```

- The effect of entity is significant for all three scents.

Simple effects analysis

The effect of entity within type of scent

```
emmeans::joint_tests(scent_afx, "scent_mask")
```

```
## scent_mask = None:
##  model term df1   df2 F.ratio p.value
##  entity        2 287.5 339.626 <.0001
##
## scent_mask = Human:
##  model term df1   df2 F.ratio p.value
##  entity        2 287.5 136.965 <.0001
##
## scent_mask = Fox:
##  model term df1   df2 F.ratio p.value
##  entity        2 287.5  14.228 <.0001
```

- The effect of entity is significant for all three scents.
- That's not a helpful finding.

Simple effects analysis

The effect of scent used within entity

```
emmeans::joint_tests(scent_afx, "entity")
```

```
## entity = Human:
##  model term df1    df2 F.ratio p.value
##  scent_mask   2 292.61  85.715 <.0001
##
## entity = Shapeshifter:
##  model term df1    df2 F.ratio p.value
##  scent_mask   2 292.61  14.723 <.0001
##
## entity = Alien:
##  model term df1    df2 F.ratio p.value
##  scent_mask   2 292.61  37.620 <.0001
```

- The effect of the type of scent worn is significant in all three entities.

Simple effects analysis

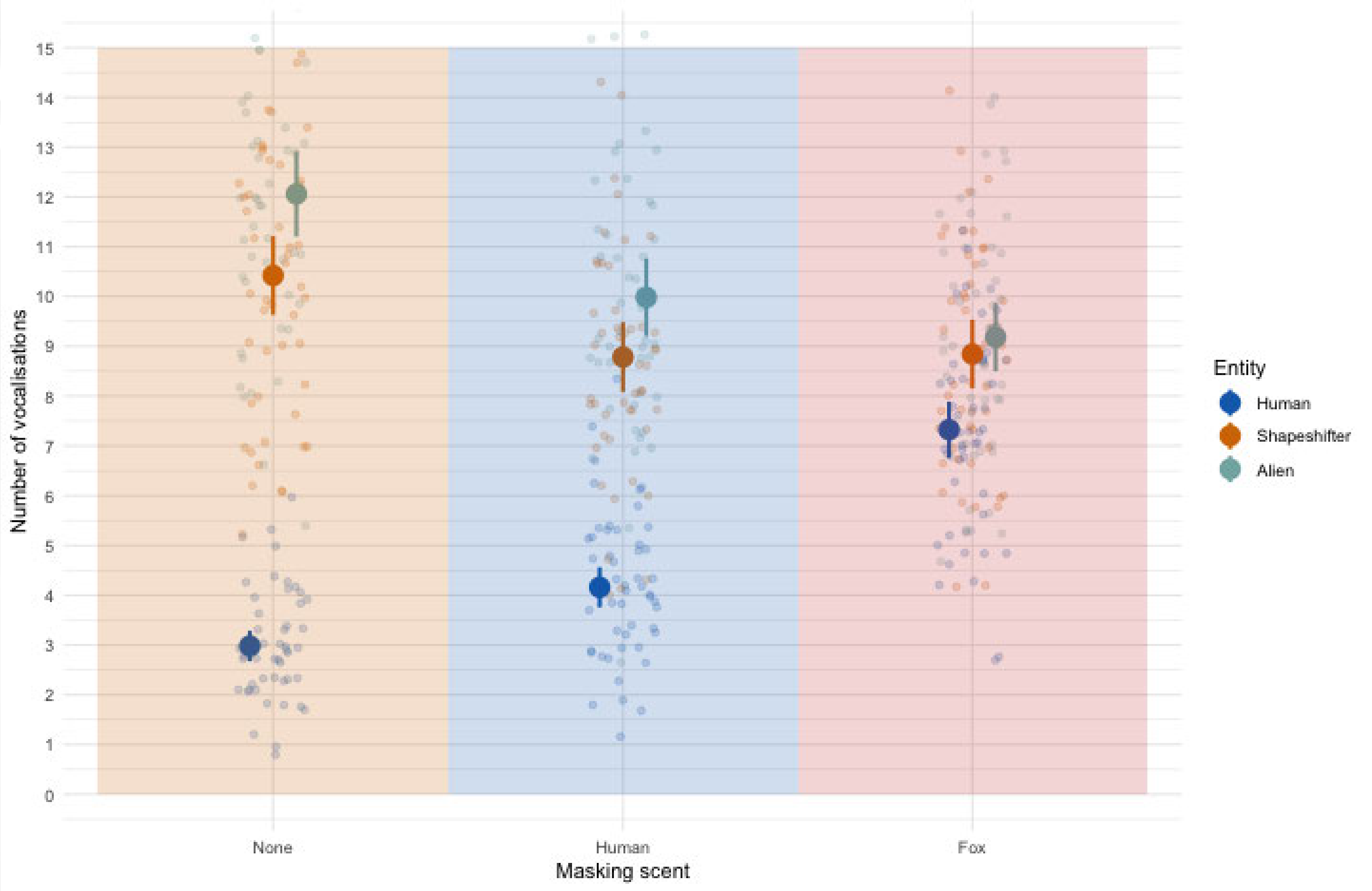
The effect of scent used within entity

```
emmeans::joint_tests(scent_afx, "entity")
```

```
## entity = Human:
##  model term df1    df2 F.ratio p.value
##  scent_mask   2 292.61  85.715 <.0001
##
## entity = Shapeshifter:
##  model term df1    df2 F.ratio p.value
##  scent_mask   2 292.61  14.723 <.0001
##
## entity = Alien:
##  model term df1    df2 F.ratio p.value
##  scent_mask   2 292.61  37.620 <.0001
```

- The effect of the type of scent worn is significant in all three entities.
- This is also not a helpful finding!

Post hoc tests across an interaction



Post hoc tests across an interaction

```
int_emm <- emmeans::emmeans(scent_afx, ~entity|scent_mask, method = "multivariate")
pairs(int_emm, adjust = "holm")
```

```
## scent_mask = None:
##  contrast             estimate    SE  df t.ratio p.value
##  Human - Shapeshifter    -7.44 0.371 288 -20.036 <.0001
##  Human - Alien           -9.08 0.371 288 -24.453 <.0001
##  Shapeshifter - Alien    -1.64 0.371 288  -4.417 <.0001
##
## scent_mask = Human:
##  contrast             estimate    SE  df t.ratio p.value
##  Human - Shapeshifter    -4.62 0.371 288 -12.442 <.0001
##  Human - Alien           -5.82 0.371 288 -15.673 <.0001
##  Shapeshifter - Alien    -1.20 0.371 288  -3.232 0.0014
##
## scent_mask = Fox:
##  contrast             estimate    SE  df t.ratio p.value
##  Human - Shapeshifter    -1.52 0.371 288  -4.093 0.0001
##  Human - Alien           -1.86 0.371 288  -5.009 <.0001
##  Shapeshifter - Alien    -0.34 0.371 288  -0.916 0.3606
##
## P value adjustment: holm method for 3 tests
```

Post hoc tests across an interaction

✏ When no scent is worn, mean vocalisations differ between all entities: aliens elicit significantly more vocalisations than both shapeshifters and humans, and shapeshifters elicit significantly more vocalisations than humans.

✏ This pattern of findings is the same when a human scent is worn.

✏ When fox scent is worn, there are still significantly more vocalisations when sniffing aliens and shapeshifters compared to humans, but the difference between shapeshifters and aliens is *not significant*.

✏ To sum up, the scents don't distract the sniffer dogs from detecting aliens compared to humans, but confuses them when distinguishing aliens in their lizard form compared to when in humanoid form.

Diagnostic plots and robust models

ANDY FIELD

Summary

- Repeated measures designs violate the assumption of independent errors
- Sphericity
  - The degree to which the variances of differences between pairs of conditions are similar
  - Apply the Greenhouse-Geisser correctiona nd forget about it
- Interpret *F*-statistics as you would for other models
- Follow-up tests
  - Simple effect analysis
  - Selective *post hoc* tests
  - Contrasts (but for factorial designs goes beyond what we teach)