Third party imitation: R codes and analysis

1. ***Intercoder reliability***

> kappa2(ratings)

Cohen's Kappa for 2 Raters (Weights: unweighted)

Subjects = 251

Raters = 2

Kappa = 0.968

z = 15.3

p-value = 0

1. ***Number of actions learnt***

| **Descriptive Statistics** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | | **Behaviours\_learnt** | | | |
|  | | **Control** | | **Test** | |
| Valid |  | 5 |  | 6 |  |
| Missing |  | 0 |  | 0 |  |
| Mean |  | 2.200 |  | 4.167 |  |
| Std. Error of Mean |  | 0.200 |  | 0.401 |  |
| Std. Deviation |  | 0.447 |  | 0.983 |  |
| Minimum |  | 2.000 |  | 3.000 |  |
| Maximum |  | 3.000 |  | 5.000 |  |
|  | | | | | |
| *Note.*  Excluded 76 rows from the analysis that correspond to the missing values of the split-by variable Groups\_BL | | | | | |

*Normality test for number of actions learnt in test and control group*

> control <- ds[which(ds$Groups\_BL==0),]

> test <- ds[which(ds$Groups\_BL==1),]

> shapiro.test(control$Behaviours\_learnt)

Shapiro-Wilk normality test

data: control$Behaviours\_learnt

W = 0.55218, p-value = 0.000131

> shapiro.test(test$Behaviours\_learnt)

Shapiro-Wilk normality test

data: test$Behaviours\_learnt

W = 0.77516, p-value = 0.03473

*Non parametric test to compare means of number of actions learnt by two groups*

> wilcox.test(test$Behaviours\_learnt, control$Behaviours\_learnt, "greater")

Wilcoxon rank sum test with continuity correction

data: test$Behaviours\_learnt and control$Behaviours\_learnt

W = 29, p-value = 0.005052

alternative hypothesis: true location shift is greater than 0

1. ***Response accuracy***

*Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']*

Family: gaussian ( identity )

Formula: Response\_Rate ~ Group + Behaviours + (1 | Subjects)

Data: birds

Control: glmerControl(optimizer = "bobyqa")

AIC BIC logLik deviance df.resid

21.2 37.2 -2.6 5.2 47

Scaled residuals:

Min 1Q Median 3Q Max

-1.85902 -0.75988 -0.00395 0.28175 2.89900

Random effects:

Groups Name Variance Std.Dev.

Subjects (Intercept) 0.00000 0.0000

Residual 0.06204 0.2491

Number of obs: 55, groups: Subjects, 11

Fixed effects:

Estimate Std. Error t value Pr(>|z|)

(Intercept) 0.20665 0.08363 2.471 0.013471 \*

GroupTest 0.25639 0.06745 3.801 0.000144 \*\*\*

BehavioursLiftleg 0.12440 0.10621 1.171 0.241476

BehavioursSpin 0.04197 0.10621 0.395 0.692729

BehavioursVocal -0.02872 0.10621 -0.270 0.786854

BehavioursWings -0.27377 0.10621 -2.578 0.009946 \*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:

(Intr) GrpTst BhvrsL BhvrsS BhvrsV

GroupTest -0.440

BehvrsLftlg -0.635 0.000

BehavirsSpn -0.635 0.000 0.500

BehavirsVcl -0.635 0.000 0.500 0.500

BehavrsWngs -0.635 0.000 0.500 0.500 0.500

optimizer (bobyqa) convergence code: 0 (OK)

boundary (singular) fit: see help('isSingular')

1. ***Learning speed***

|  |
| --- |
|  |

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']

Family: poisson ( log )

Formula: sessions ~ Group + Behaviours + (1 | Subjects)

Data: Birds\_sessions

Control: glmerControl(optimizer = "bobyqa")

AIC BIC logLik deviance df.resid

198.8 207.7 -93.4 186.8 27

Scaled residuals:

Min 1Q Median 3Q Max

-2.3834 -0.9291 -0.2011 0.7265 3.7232

Random effects:

Groups Name Variance Std.Dev.

Subjects (Intercept) 0.02927 0.1711

Number of obs: 33, groups: Subjects, 11

Fixed effects:

Estimate Std. Error z value Pr(>|z|)

(Intercept) 2.3324 0.1668 13.979 < 2e-16 \*\*\*

GroupTest -0.5512 0.1718 -3.208 0.00133 \*\*

BehavioursLiftleg -0.4142 0.1980 -2.091 0.03650 \*

BehavioursSpin 0.1006 0.1702 0.591 0.55455

BehavioursVocal -0.2177 0.2346 -0.928 0.35333

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:

(Intr) GrpTst BhvrsL BhvrsS

GroupTest -0.529

BehvrsLftlg -0.487 -0.020

BehavirsSpn -0.608 0.015 0.501

BehavirsVcl -0.409 -0.123 0.350 0.438

**Mean (sessions)**: ≈ 6.85

 **Variance**: ≈ 23.07

 **Overdispersion ratio (Variance / Mean)**: **≈ 3.37**

*Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']*

Family: Negative Binomial(5.6465) ( log )

Formula: sessions ~ Group + Behaviours + (1 | Subjects)

Data: Birds\_sessions

AIC BIC logLik deviance df.resid

189.0 199.5 -87.5 175.0 26

Scaled residuals:

Min 1Q Median 3Q Max

-1.4364 -0.7070 -0.2493 0.5653 2.8549

Random effects:

Groups Name Variance Std.Dev.

Subjects (Intercept) 1.748e-13 4.18e-07

Number of obs: 33, groups: Subjects, 11

Fixed effects:

Estimate Std. Error z value Pr(>|z|)

(Intercept) 2.3045 0.2233 10.319 <2e-16 \*\*\*

GroupTest -0.5363 0.2139 -2.507 0.0122 \*

BehavioursLiftleg -0.3256 0.2867 -1.136 0.2562

BehavioursSpin 0.1352 0.2622 0.516 0.6061

BehavioursVocal -0.1888 0.3292 -0.574 0.5662

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:

(Intr) GrpTst BhvrsL BhvrsS

GroupTest -0.468

BehvrsLftlg -0.552 -0.119

BehavirsSpn -0.649 -0.033 0.522

BehavirsVcl -0.409 -0.258 0.443 0.460

optimizer (Nelder\_Mead) convergence code: 0 (OK)

boundary (singular) fit: see help('isSingular')

**res <- residuals(binom, type = "pearson")**

**> sum(res^2) / df.residual(binom)**

**[1] 1.268428**

*Post hoc analysis- Non parametric test to compare mean session number taken to learn fluff between two groups*

>wilcox.test(Sessions\_fluff~Groups\_fluff,data=fluff,exact=FALSE)

Wilcoxon rank sum test with continuity correction

data: Sessions\_fluff by Groups\_fluff

W = 15, p-value = 0.03466

alternative hypothesis: true location shift is not equal to 0

1. ***Serial order of actions learnt***

*CLM to test whether order of the actions (rank) learnt is predicted by the actions and the formula:*

Order\_effect ~ Actions + Groups

data: ordinal

link threshold nobs logLik AIC niter max.grad cond.H

logit flexible 36 -43.08 104.17 5(0) 1.07e-08 5.8e+01

Coefficients:

Estimate Std. Error z value Pr(>|z|)

ActionsLift leg -1.6713 0.9828 -1.701 0.0890 .

ActionsSpin -0.4975 0.9187 -0.541 0.5882

ActionsVocal 0.6425 0.9695 0.663 0.5075

ActionsWings 4.0764 1.6695 2.442 0.0146 \*

GroupsTest 1.3216 0.7371 1.793 0.0730 .

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Threshold coefficients:

Estimate Std. Error z value

1|2 -0.5041 0.7994 -0.631

2|3 1.2528 0.8145 1.538

3|4 2.7906 0.9459 2.950

4|5 4.5999 1.3134 3.502

> ord\_effect <- clm(Order\_effect ~ Actions + Groups\*Actions, data = ordinal)

> summary(ord\_effect)

*formula: Order\_effect ~ Actions + Groups \* Actions*

*data: ordinal*

link threshold nobs logLik AIC niter max.grad cond.H

logit flexible 36 -40.91 105.82 5(0) 1.02e-08 1.8e+02

Coefficients: (1 not defined because of singularities)

Estimate Std. Error z value Pr(>|z|)

ActionsLift leg -2.3722 1.5682 -1.513 0.13035

ActionsSpin -2.7435 1.5318 -1.791 0.07329 .

ActionsVocal -0.6022 1.8562 -0.324 0.74563

ActionsWings 5.0854 1.8045 2.818 0.00483 \*\*

GroupsTest -0.5000 1.3403 -0.373 0.70911

ActionsLift leg:GroupsTest 1.3498 1.9669 0.686 0.49255

ActionsSpin:GroupsTest 3.9517 2.0278 1.949 0.05133 .

ActionsVocal:GroupsTest 2.2687 2.2478 1.009 0.31282

ActionsWings:GroupsTest NA NA NA NA

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Threshold coefficients:

Estimate Std. Error z value

1|2 -1.5880 1.0336 -1.536

2|3 0.3837 0.9601 0.400

3|4 1.9586 1.0747 1.822

4|5 3.7893 1.4141 2.680

***Supplementary Results***

# *One-sided binomial tests (p > 0.5)*

individuals <- data.frame(

+ ID = c("Marvel", "Carrot", "IM"),

+ successes = c(35, 30, 29),

+ trials = c(40, 40, 40)

*+ )*

> individuals$p\_value <- mapply(function(x, n) {

+ binom.test(x, n, p = 0.2, alternative = "greater")$p.value

+ }, individuals$successes, individuals$trials)

> m <- nrow(individuals)

> individuals$sidak\_adjusted <- 1 - (1 - individuals$p\_value)^m

> alpha <- 0.05

> individuals$significant <- individuals$sidak\_adjusted < alpha

> print(individuals)

# ID successes trials p\_value sidak\_adjusted significant

1 Marvel 35 40 7.672843e-20 0.000000e+00 TRUE

2 Carrot 30 40 1.061995e-13 3.187450e-13 TRUE

3 IM 29 40 1.172330e-12 3.516853e-12 TRUE

individuals2 <- data.frame(

+ ID = c("Natasha", "Thor", "Pickle"),

+ successes = c(37, 36, 33),

+ trials = c(40, 40, 40)

+ )

> individuals2$p\_value <- mapply(function(x, n) {

+ binom.test(x, n, p = 0.33, alternative = "greater")$p.value

+ }, individuals2$successes, individuals2$trials)

> m <- nrow(individuals2)

> individuals2$sidak\_adjusted <- 1 - (1 - individuals2$p\_value)^m

> alpha <- 0.05

> individuals2$significant <- individuals2$sidak\_adjusted < alpha

> print(individuals2)

# ID successes trials p\_value sidak\_adjusted significant

1 Natasha 37 40 4.731293e-15 1.432188e-14 TRUE

2 Thor 36 40 9.017886e-14 2.704503e-13 TRUE

3 Pickle 33 40 1.620184e-10 4.860552e-10 TRUE

*Linear Regression to test whether age predicts the number of behaviours learnt*

| **Model Summary - No.of behaviors** | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | | **R** | | **R²** | | **Adjusted R²** | | **RMSE** | |
| H₀ |  | 0.000 |  | 0.000 |  | 0.000 |  | 1.813 |  |
| H₁ |  | 0.788 |  | 0.621 |  | 0.545 |  | 1.223 |  |
|  | | | | | | | | | |

| **ANOVA** | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | |  | | **Sum of Squares** | | **df** | | **Mean Square** | | **F** | | **p** | |
| H₁ |  | Regression |  | 12.240 |  | 1 |  | 12.240 |  | 8.188 |  | 0.035 |  |
|  |  | Residual |  | 7.474 |  | 5 |  | 1.495 |  |  |  |  |  |
|  |  | Total |  | 19.714 |  | 6 |  |  |  |  |  |  |  |
|  | | | | | | | | | | | | | |
| *Note.*  The intercept model is omitted, as no meaningful information can be shown. | | | | | | | | | | | | | |

| **Coefficients** | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | |  | | **Unstandardized** | | **Standard Error** | | **Standardized** | | **T** | | **p** | |
| H₀ |  | (Intercept) |  | 3.571 |  | 0.685 |  |  |  | 5.213 |  | 0.002 |  |
| H₁ |  | (Intercept) |  | 4.891 |  | 0.653 |  |  |  | 7.491 |  | < .001 |  |
|  |  | Age |  | -0.240 |  | 0.084 |  | -0.788 |  | -2.861 |  | 0.035 |  |
|  | | | | | | | | | | | | | |