Math 530/630: CM 5.4
ANOVA:
2- and 3-way

Let’s add a second predictor!

PLOS ONE

RESEARCH ARTICLE

Being Sticker Rich: Numerical Context Influences Children's Sharing Behavior

Tasha Posid[1]*, Allyse Fazio[2], Sara Cordes[2]

1 Department of Psychology, The Ohio State University, Columbus, Ohio, United States of America,
2 Department of Psychology, Boston College, Chestnut Hill, Massachusetts, United States of America

* posid.1@osu.edu

Being sticker rich

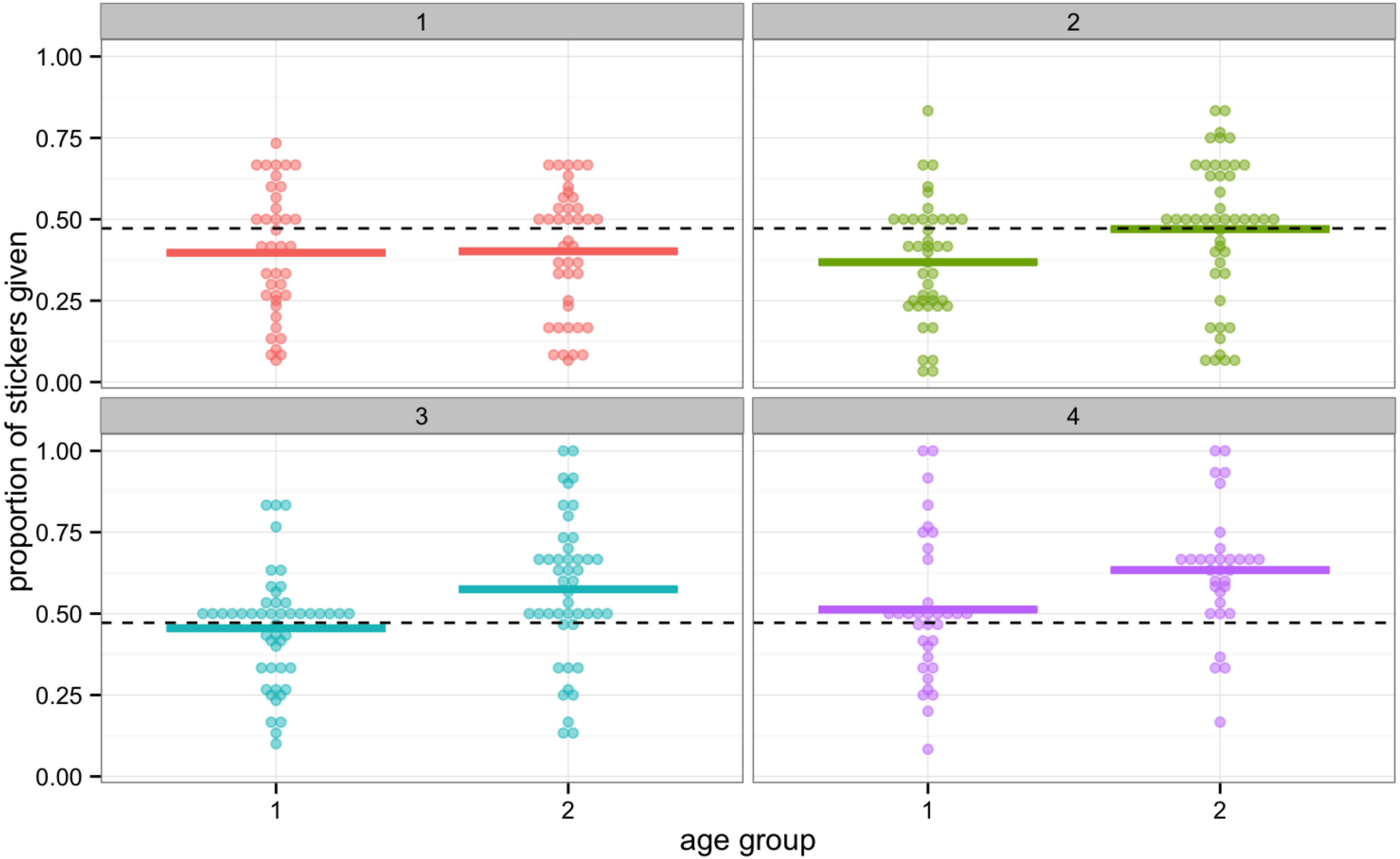
- Children (ages 3–11) received a small (12, “sticker poor”) or large (30, “sticker rich”) number of stickers, and were then given the opportunity to share their windfall with either one or multiple anonymous recipients (Dictator Game).
- Do the number of available resources and/or the number of potential recipients alter the likelihood of a child donating and/or the amount they donate?

Approach

- “Givers” only analyzed
- *“A univariate [read: one response/outcome variable] ANOVA was conducted investigating the impact of the between-subjects factors [read: all levels of factors are measured from independent samples] of age (4: 3–4 years, 5–6 years, 7–8 years, 9–11 years), number of resources (2: 12 or 30 stickers), number of recipients (2: 1 or 2 anonymous recipients), and gender (2: female, male) on the proportion of resources shared.”*

Adding a covariate

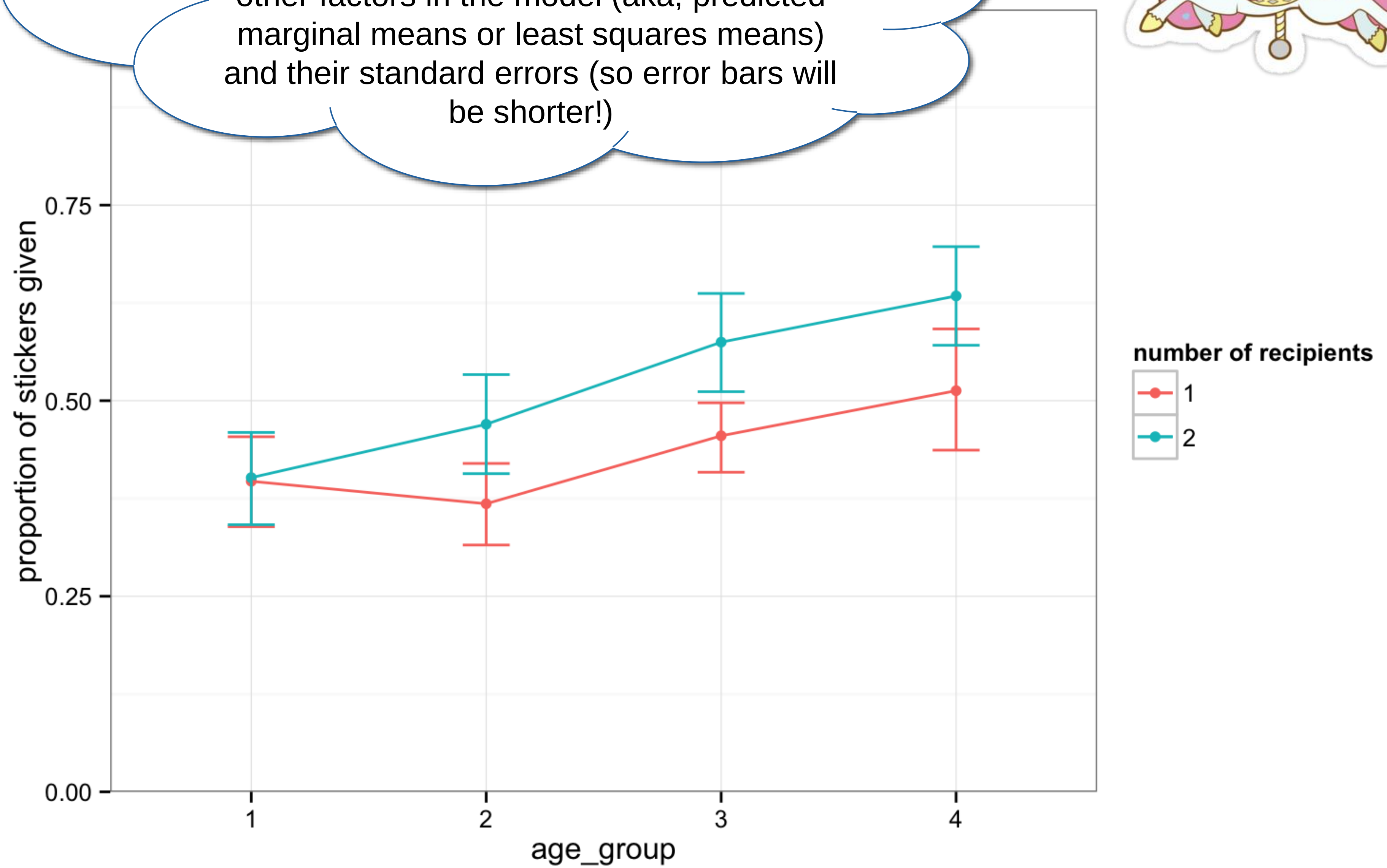
- Past: one-way ANOVA
- Now: analysis of covariance (ANCOVA)
- Variables as covariates are typically added (+), and their effects are assumed to be additive
- No interaction term (meaning we are not allowing for estimated non-parallel lines)
- If we include another variable in our model, say 1 vs. 2 recipients, then our estimate of the effect of age group is interpreted holding the value of the covariate fixed.
  - Just like in multiple regression.


1
2
3
4
1.00
0.75
0.50
0.25
0.00
proportion of stickers given
1
2
age group

ggplot code for previous plot

```
ggplot(givers, aes(x = factor(num_env), y = prop_given, fill = age_group, colour =
age_group)) +
  geom_dotplot(stackdir = "center", binaxis = "y",
    binwidth = .01, binpositions = "all", stackratio = 1, dotsize = 3, alpha = .5) +
  stat_summary(fun.y = mean, fun.ymin = mean,
    fun.ymax = mean, geom = "crossbar", width = 0.75, lwd = .75) +
  scale_x_discrete(name = "age group") +
  scale_y_continuous(name = "proportion of stickers given") +
  geom_hline(yintercept = mean(givers$prop_given), lty = "dashed") +
  theme_bw() +
  theme(legend.position = "none") +
  facet_wrap(~ age_group)
```

These are the “unadjusted” means plus 95% confidence intervals: in ANCOVA, you’ll also want to plot means adjusted for other factors in the model (aka, predicted marginal means or least squares means) and their standard errors (so error bars will be shorter!)



ggplot code for previous plot

```
ggplot(givers, aes(x = age_group, y = prop_given, colour = num_env)) +
  stat_summary(fun.y = mean, geom = "point") +
  stat_summary(fun.y = mean, geom = "line", aes(group = num_env)) +
  stat_summary(fun.data = mean_cl_boot, geom = "errorbar", width = 0.2) +
  labs(x = "age_group",
       y = "proportion of stickers given",
       colour = "number of recipients") +
  theme_bw() +
  coord_cartesian(ylim = c(0, 1))
```

ANCOVA in R

```
anova(lm(prop_given ~ age_group + num_env, data = givers))
Analysis of Variance Table

Response: prop_given
           Df  Sum Sq Mean Sq F value        Pr(>F)
age_group   3  1.5111 0.50370  12.575 0.00000008525 ***
num_env     1  0.5940 0.59399  14.829      0.000142 ***
Residuals 323 12.9383 0.04006
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Don’t try this one at home…

ANCOVA in R

```
anova(lm(prop_given ~ age_group + num_env, data = givers))
Analysis of Variance Table

Response: prop_given
           Df  Sum Sq Mean Sq F value        Pr(>F)
age_group   3  1.5111 0.50370  12.575 0.00000008525 ***
num_env     1  0.5940 0.59399  14.829      0.000142 ***
Residuals 323 12.9383 0.04006
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
# eep order changes coefficient estimates!
anova(lm(prop_given ~ num_env + age_group, data = givers))
Analysis of Variance Table

Response: prop_given
           Df  Sum Sq Mean Sq F value        Pr(>F)
num_env     1  0.5451 0.54510  13.608     0.0002642 ***
age_group   3  1.5600 0.52000  12.982 0.00000005007 ***
Residuals 323 12.9383 0.04006
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

order

matters

!

Don’t try this one at home…

What is happening here?

- Remember, the **`anova()`** command as we used it before was used to compare two nested models. The null hypothesis was that the more complicated model was not better than the less complicated model…

```
lm_age <- lm(prop_given ~ age_group, data = givers)
lm_age_env <- lm(prop_given ~ age_group + num_env, data = givers)
anova(lm_age, lm_age_env)
Analysis of Variance Table

Model 1: prop_given ~ age_group
Model 2: prop_given ~ age_group + num_env
  Res.Df    RSS Df Sum of Sq      F   Pr(>F)
1    324 13.532
2    323 12.938  1   0.59399 14.829 0.000142 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
```

This is exactly what we got with num_env second!

Wait, did lm() do this awful thing to us?

```
lm(prop_given ~ num_env + age_group, data = givers)

Call:
lm(formula = prop_given ~ num_env + age_group, data = givers)

Coefficients:
(Intercept)      num_env2    age_group1    age_group2    age_group3
    0.43472       0.08515      -0.07899      -0.05790       0.03841

lm(prop_given ~ age_group + num_env, data = givers)

Call:
lm(formula = prop_given ~ age_group + num_env, data = givers)

Coefficients:
(Intercept)    age_group1    age_group2    age_group3      num_env2
    0.43472      -0.07899      -0.05790       0.03841       0.08515
```

Types of sums of squares

- Don’t bring this up on stack overflow ☺
- Type 1: sequential (order matters) [this is the default in R!]
  - This is rarely what you will be interested in if you are not doing a nested models comparison intentionally
- Type II:
  - This type tests for each main effect *after* the other main effect.
  - Note that *no significant interaction* is assumed (in other words, you should test for interaction first) and only if AB is not significant, continue with the analysis for main effects).
- Type III:
  - This type tests for the presence of a main effect *after* the other main effect and interaction.
  - However, it is often not interesting to interpret a main effect if interactions are present (generally speaking, if a significant interaction is present, the main effects should not be further analysed).
  - If the interactions are not significant, type II gives a more powerful test.

ANCOVA in R the better way

```
# library(car)
sticker_mod <- lm(prop_given ~ age_group + num_env, data = givers,
+                  contrasts = list(age_group = contr.sum,
+                                   num_env = contr.sum))
Anova(sticker_mod, type = 2)
Anova Table (Type II tests)

Response: prop_given
          Sum Sq  Df F value        Pr(>F)
age_group  1.560   3  12.982 0.00000005007 ***
num_env    0.594   1  14.829      0.000142 ***
Residuals 12.938 323
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

order

does not

matter !

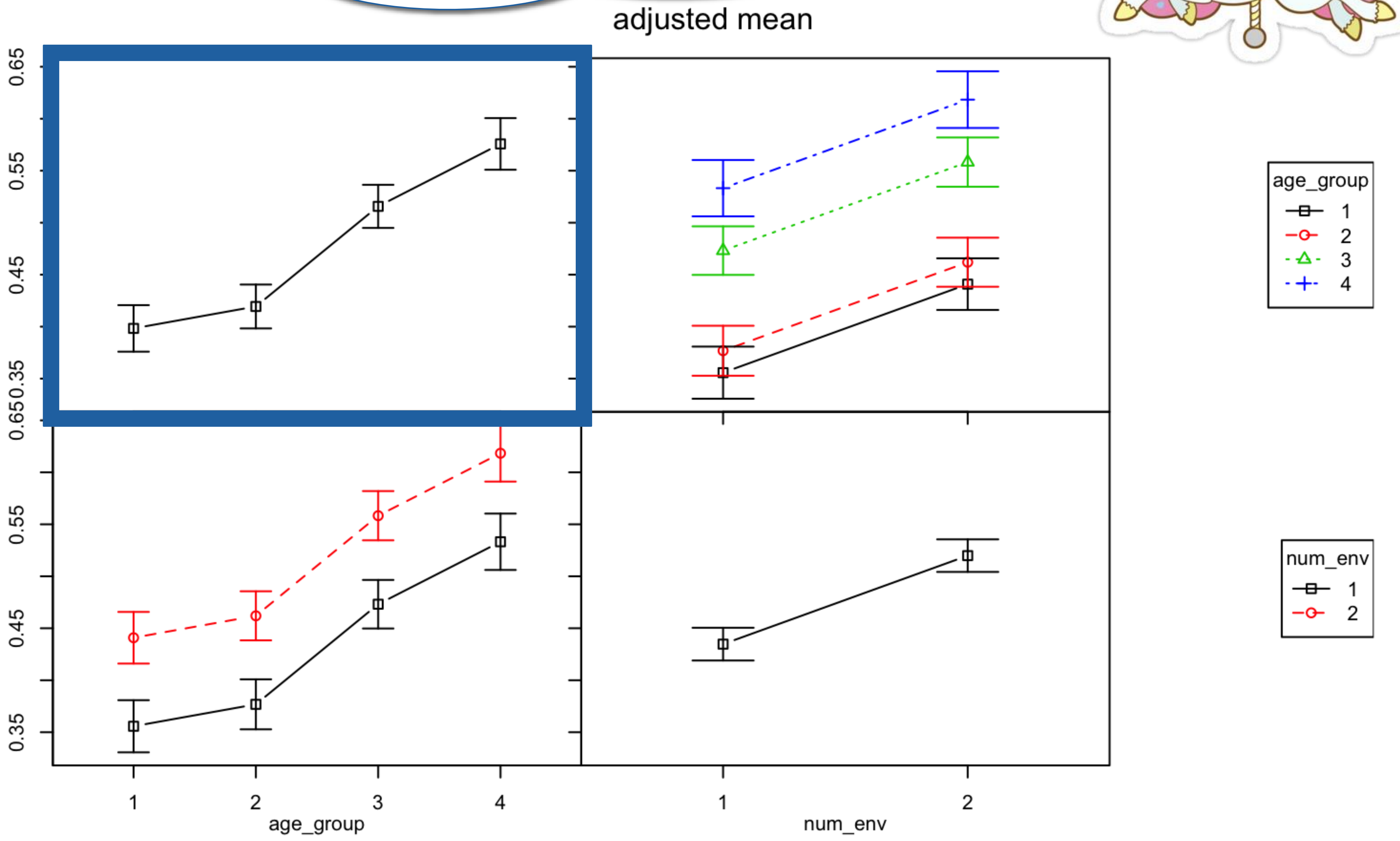
This is the one!! You can also try type = 3 (but…)

ANCOVA in R the better way

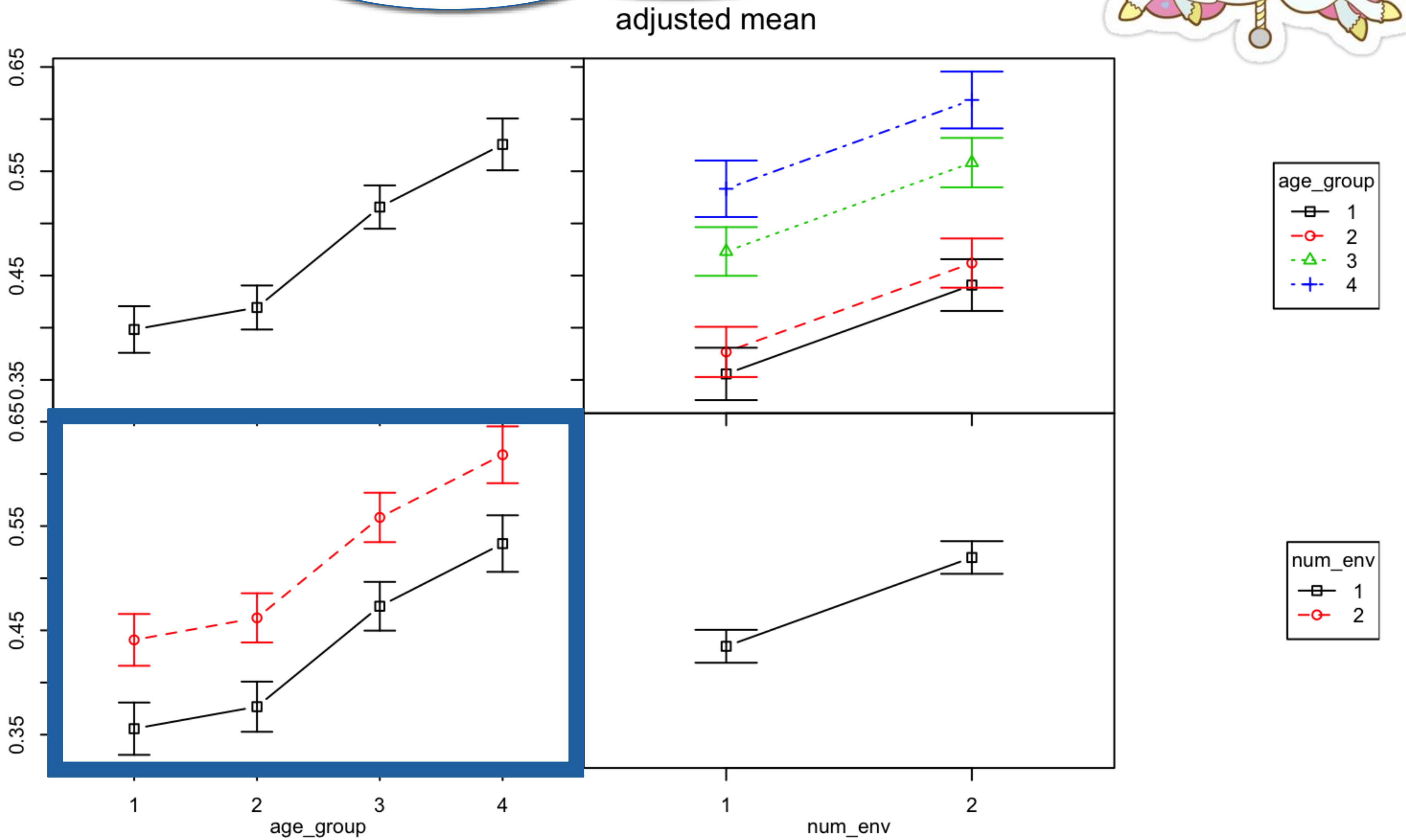
```
sticker_ancova <- lm(prop_given ~ age_group + num_env, data = givers, contrasts =
list(age_group = contr.sum, num_env = contr.sum))
summary(sticker_ancova)

Call:
lm(formula = prop_given ~ age_group + num_env, data = givers,
    contrasts = list(age_group = contr.sum, num_env = contr.sum))

Residuals:
     Min       1Q   Median       3Q      Max
-0.45169 -0.12351  0.02687  0.12317  0.46679

Coefficients:
            Estimate Std. Error t value             Pr(>|t|)
(Intercept)  0.47730    0.01116  42.768 < 0.0000000000000002 ***
age_group1  -0.07899    0.01936  -4.079            0.0000569 ***
age_group2  -0.05790    0.01864  -3.107             0.002058 **
age_group3   0.03841    0.01844   2.083             0.038061 *
num_env1    -0.04258    0.01106  -3.851             0.000142 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2001 on 323 degrees of freedom
Multiple R-squared:  0.1399,  Adjusted R-squared:  0.1293
F-statistic: 13.14 on 4 and 323 DF,  p-value: 0.0000000006306
```

$$\beta_0 = (\mu_1 + \mu_2 + \mu_3 + \mu_4)/4$$
$$\beta_1 = \mu_1 - (\mu_1 + \mu_2 + \mu_3 + \mu_4)/4$$
$$\beta_2 = \mu_2 - (\mu_1 + \mu_2 + \mu_3 + \mu_4)/4$$
$$\beta_3 = \mu_3 - (\mu_1 + \mu_2 + \mu_3 + \mu_4)/4$$

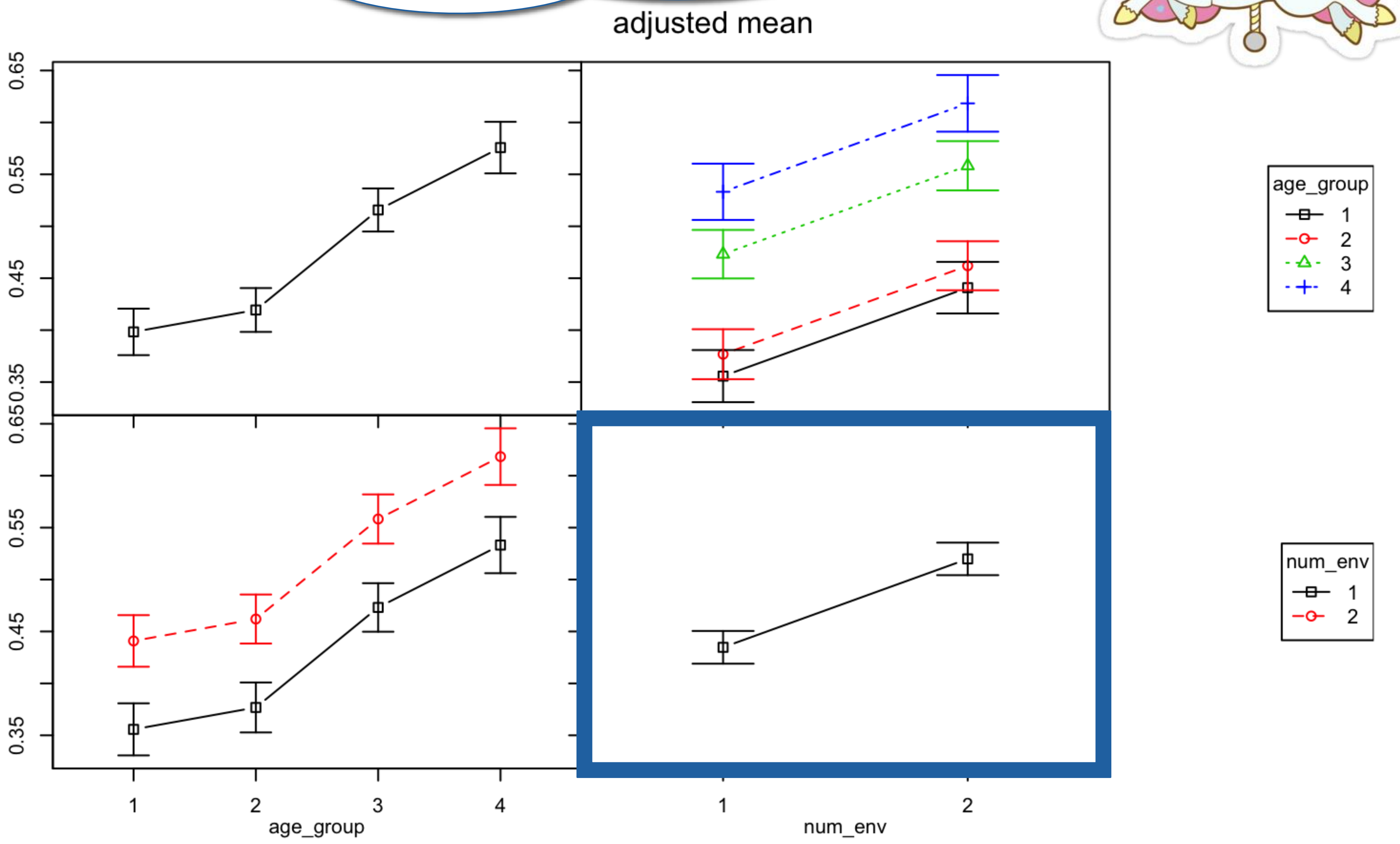
Means by age group
after adjusting for number of envelopes:
seems to increase as age goes up? (bars
show S.E.M.)



Means by age group
after adjusting for number of envelopes:
same increasing pattern (bars show
S.E.M.)

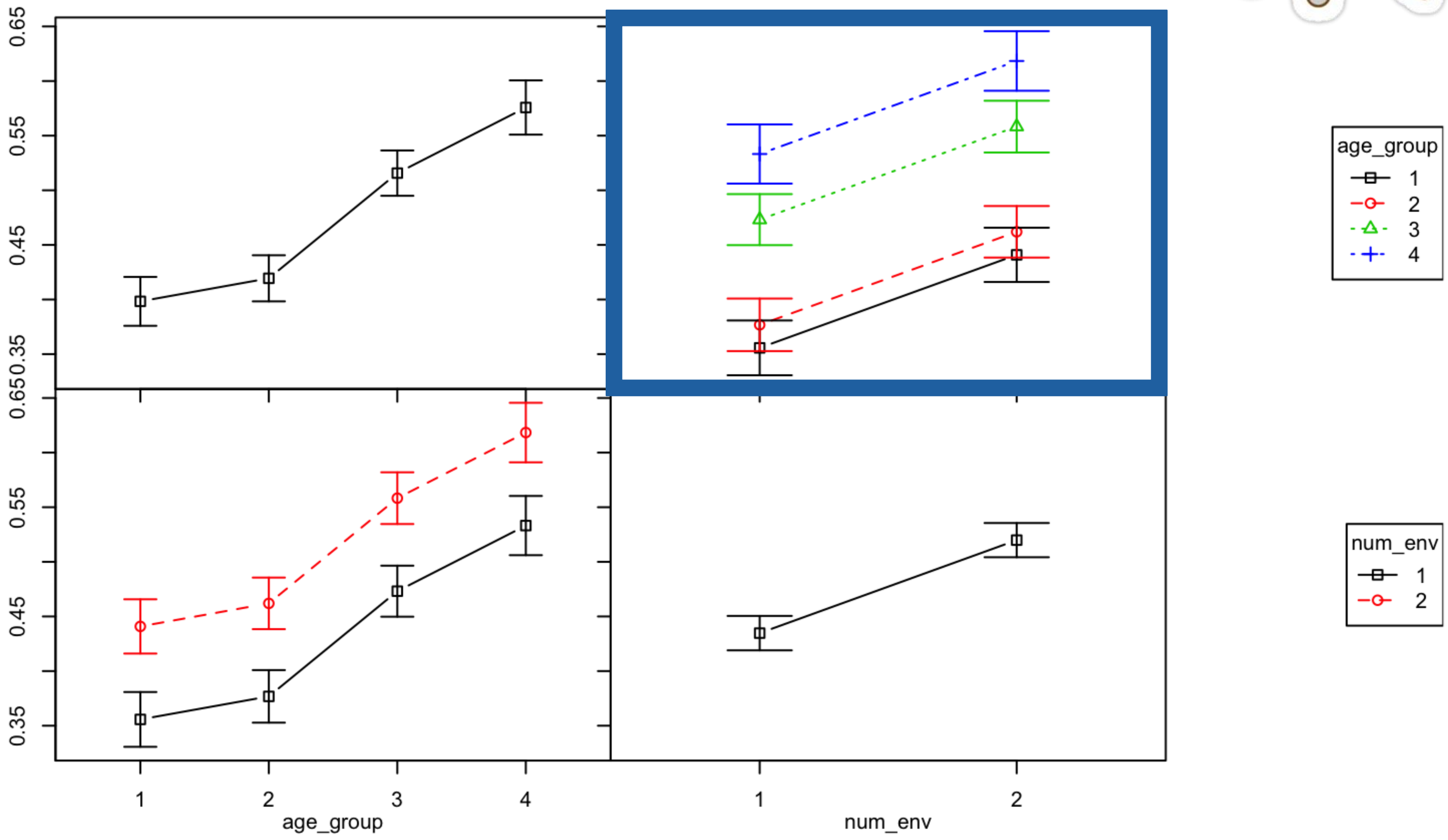


Means by number of envelopes after adjusting for age group: increases from 1 to 2 envelopes (bars show S.E.M.)

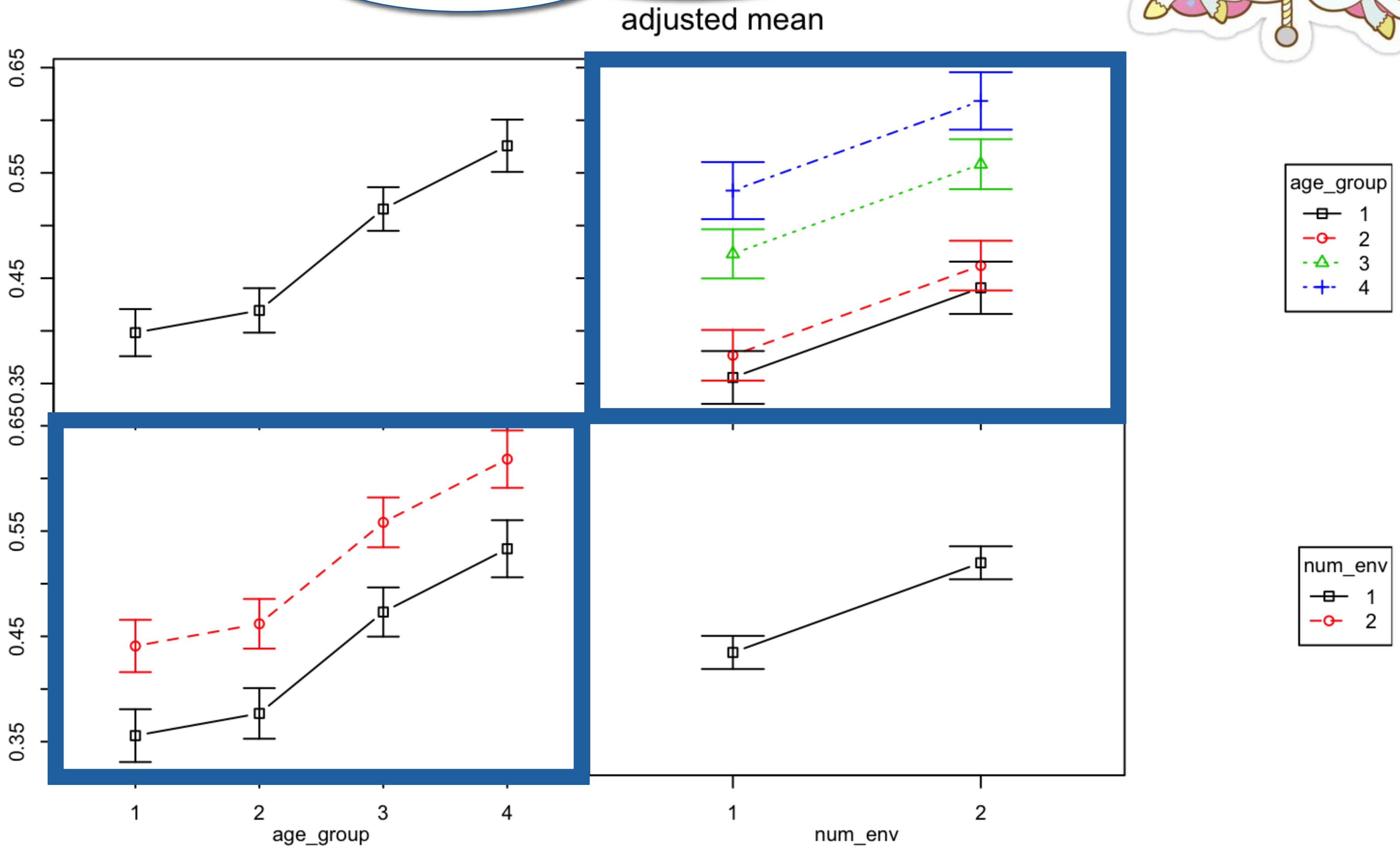


Means by number of envelopes
after adjusting for age group:
same increasing pattern for each age
group (bars show S.E.M.)

adjusted mean



Note: lines in the lower left and upper right panels **must be parallel**- we made the model that way!



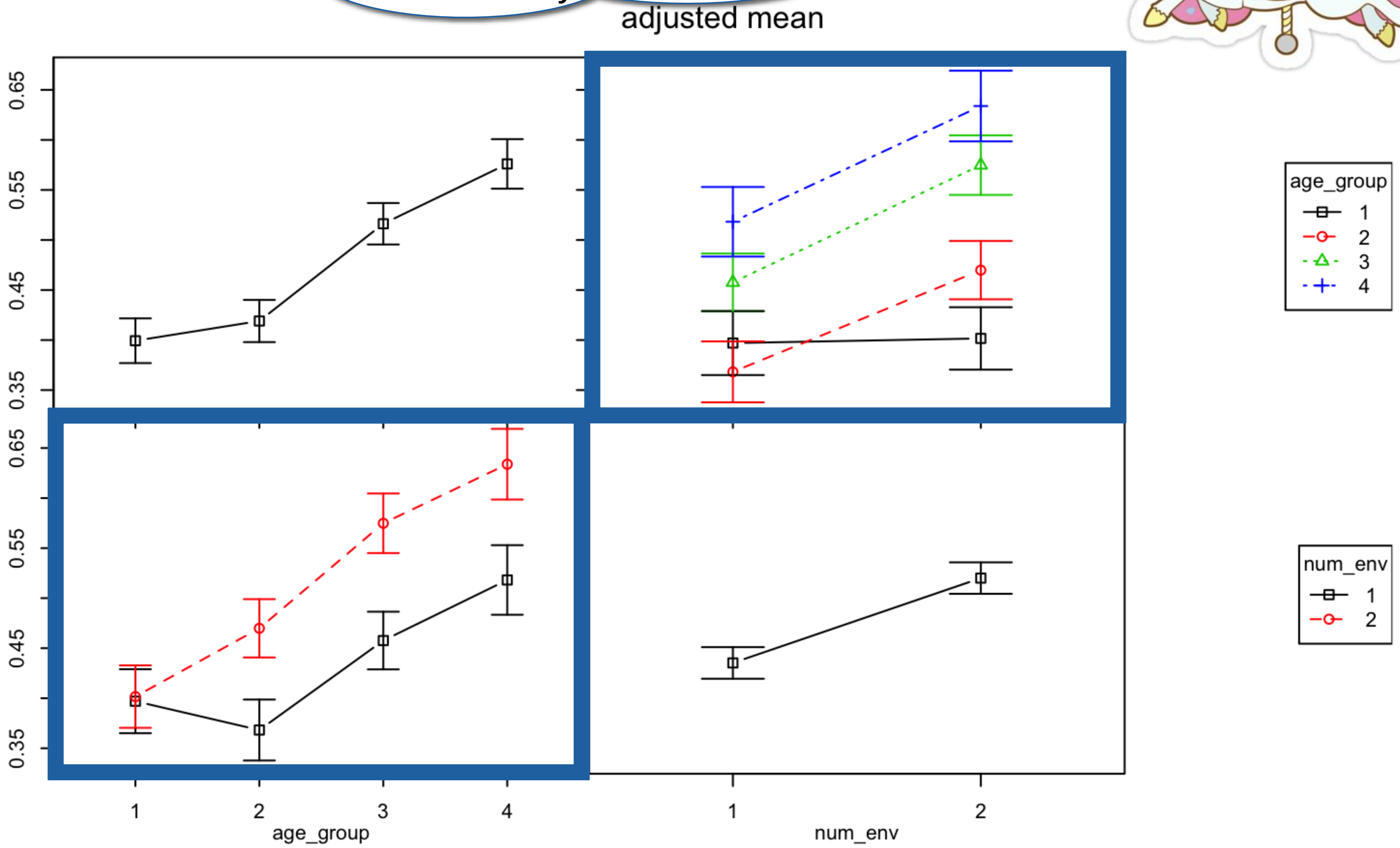
ANCOVA assumption: Homogeneity of regression slopes

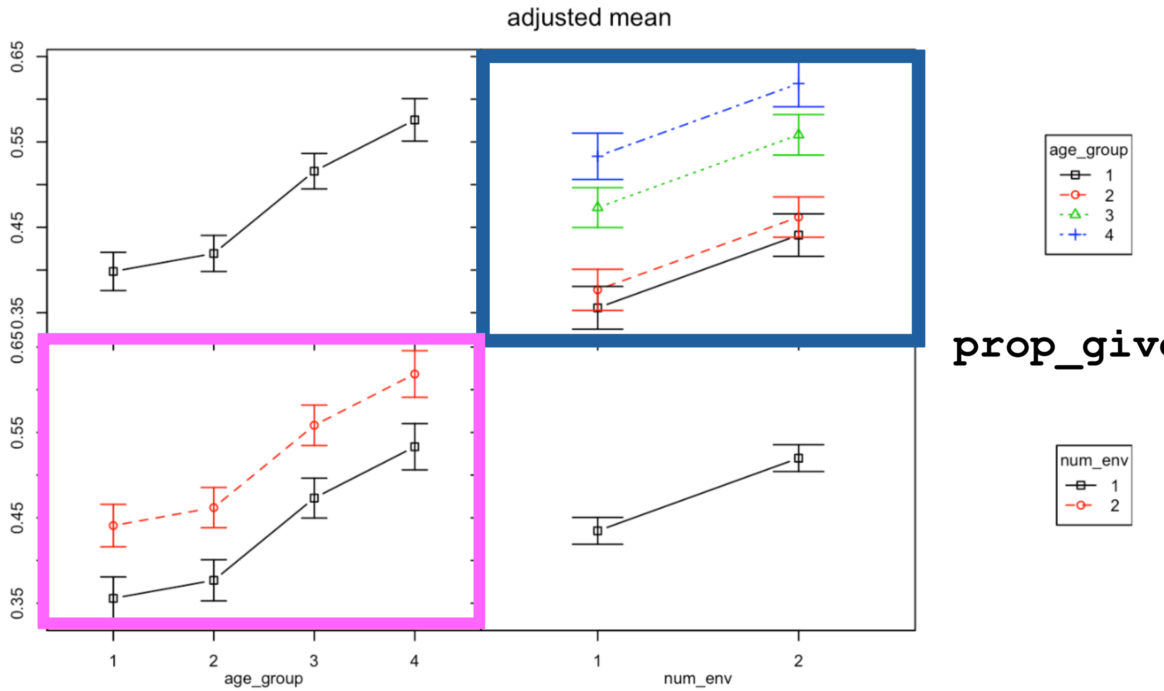
Safe assumption? Let's include interaction term…

```
# library(car)
sticker_int <- lm(prop_given ~ age_group*num_env, data = givers,
+                    contrasts = list(age_group = contr.sum,
+                                     num_env = contr.sum))
Anova(sticker_int, type = 2)
Anova Table (Type II tests)

Response: prop_given
                   Sum Sq  Df F value      Pr(>F)
age_group          1.5600   3 13.0371 0.000000047 ***
num_env            0.5940   1 14.8920   0.0001377 ***
age_group:num_env  0.1747   3  1.4599   0.2254215
Residuals         12.7636 320
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

age_group*num_env is equivalent to:
`age_group + num_env + age_group:num_env`

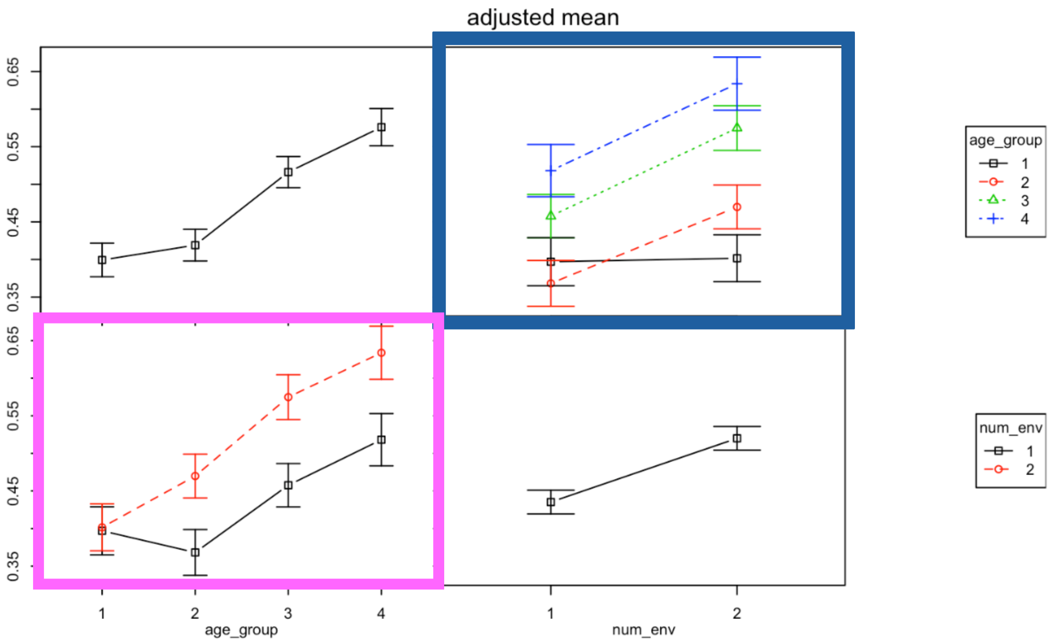
Lines in the lower left and upper right panels now can have non-parallel slopes- again, we **made** the model that way!





prop_given ~ age_group + num_env

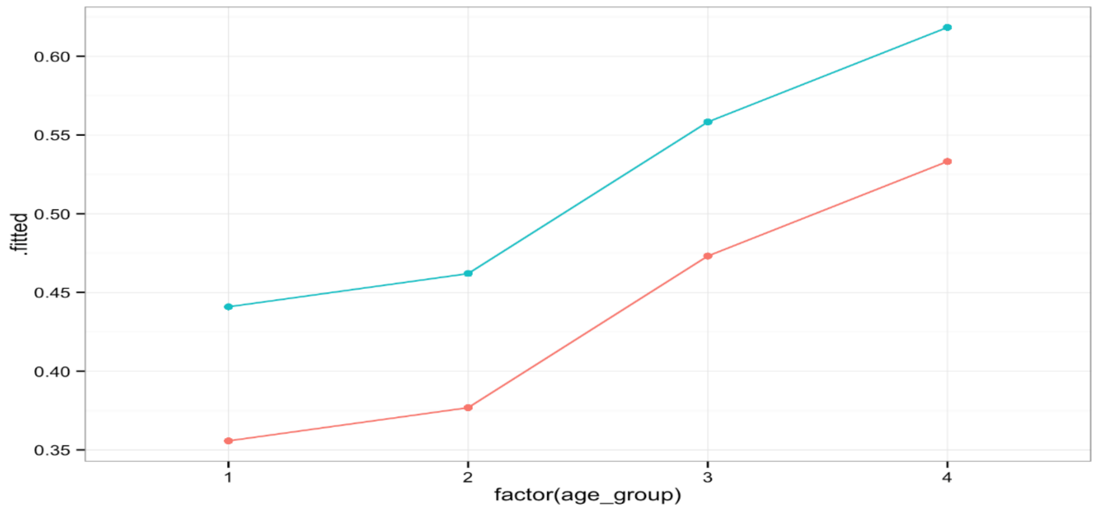
prop_given ~ age_group * num_env

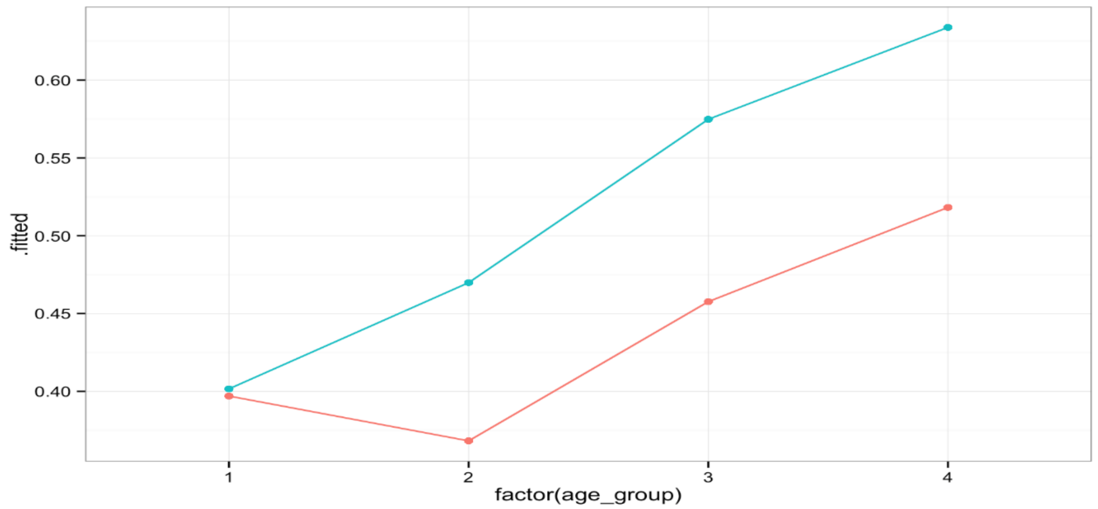


Covariate versus interacting terms

Covariates allow only for different intercepts, not slopes

Interactions allow for both different intercepts & slopes





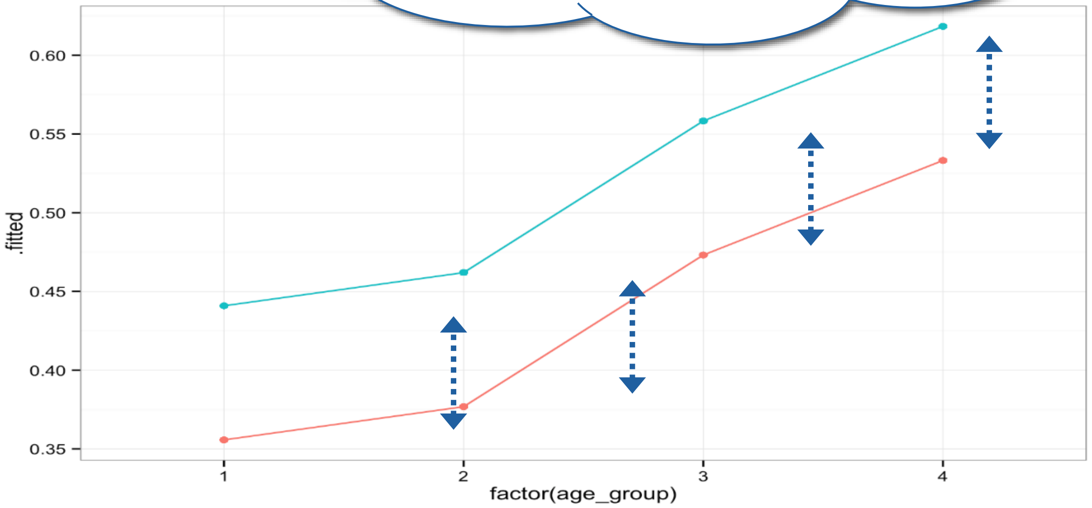
`prop_given ~ age_group + num_env`

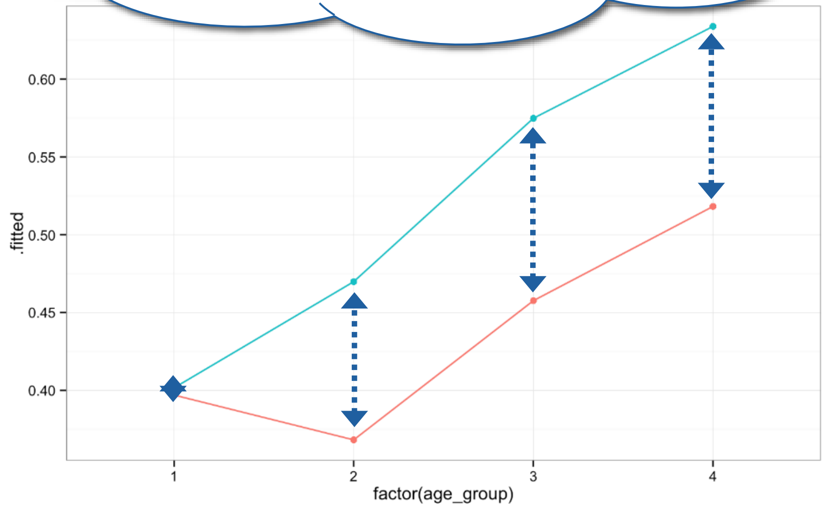
`prop_given ~ age_group * num_env`

Covari

Models says: the effect of num_env is the same at every age

Model says: the effect of num_env could differ depending on age





	age_group (fctr)	num_env (fctr)	fit_means (dbl)	fit_diff (dbl)
1	1	1	0.3557333	NA
2	1	2	0.4408878	0.08515454
3	2	1	0.3768267	NA
4	2	2	0.4619812	0.08515454
5	3	1	0.4731331	NA
6	3	2	0.5582876	0.08515454
7	4	1	0.5332060	NA
8	4	2	0.6183605	0.08515454

	age_group (fctr)	num_env (fctr)	fit_means (dbl)	diff_fit (dbl)
1	1	1	0.3970085	NA
2	1	2	0.4016260	0.004617469
3	2	1	0.3682171	NA
4	2	2	0.4698582	0.101641102
5	3	1	0.4576389	NA
6	3	2	0.5748148	0.117175926
7	4	1	0.5181818	NA
8	4	2	0.6338542	0.115672348

Including interaction term changes interpretation

- including an interaction changes the interpretation of coefficients for main effects
- the coefficient on the constitutive term X cannot be interpreted as an unconditional marginal effect since it indicates only the effect of a one-unit change in X on Y when the conditioning variable is zero.
- If the modifying variable is dichotomous, this simply requires the analyst to present four numbers—the marginal effect of X when Z is 0 and when Z is 1, along with the two corresponding standard errors.

Bottom line

- Using a variable as a covariate (+) rather than letting it interact with other variables (*) is an assumption called "**homogeneity of regression slopes**", which is what we just observed- they are assumed to be parallel
- If the variable you want to be covariate interacts with your other predictor, you cannot do an ANCOVA
- Since the interaction effect here was not significant, we can proceed with interpreting the main effects of each of our predictors separately (sticking with **`sticker_int`** model):
  - age group
  - number of recipients

Back to our results…

```
# library(car)
sticker_int <- lm(prop_given ~ age_group*num_env, data = givers,
+                    contrasts = list(age_group = contr.sum,
+                                     num_env = contr.sum))
Anova(sticker_int, type = 2)
Anova Table (Type II tests)

Response: prop_given
                   Sum Sq  Df F value       Pr(>F)
age_group          1.5600   3 13.0371 0.000000047 ***
num_env            0.5940   1 14.8920   0.0001377 ***
age_group:num_env  0.1747   3  1.4599   0.2254215
Residuals         12.7636 320
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

The significant main effect of age group tells us that the mean proportion given differed by age group

Back to our results…

```
# library(car)
sticker_int <- lm(prop_given ~ age_group*num_env, data = givers,
+                    contrasts = list(age_group = contr.sum,
+                                     num_env = contr.sum))
Anova(sticker_int, type = 2)
Anova Table (Type II tests)

Response: prop_given
                  Sum Sq  Df F value      Pr(>F)
age_group         1.5600   3 13.0371 0.000000047 ***
num_env           0.5940   1 14.8920   0.0001377 ***
age_group:num_env 0.1747   3  1.4599   0.2254215
Residuals        12.7636 320
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The significant main effect of number of recipients tells us that the mean proportion given differed based on number

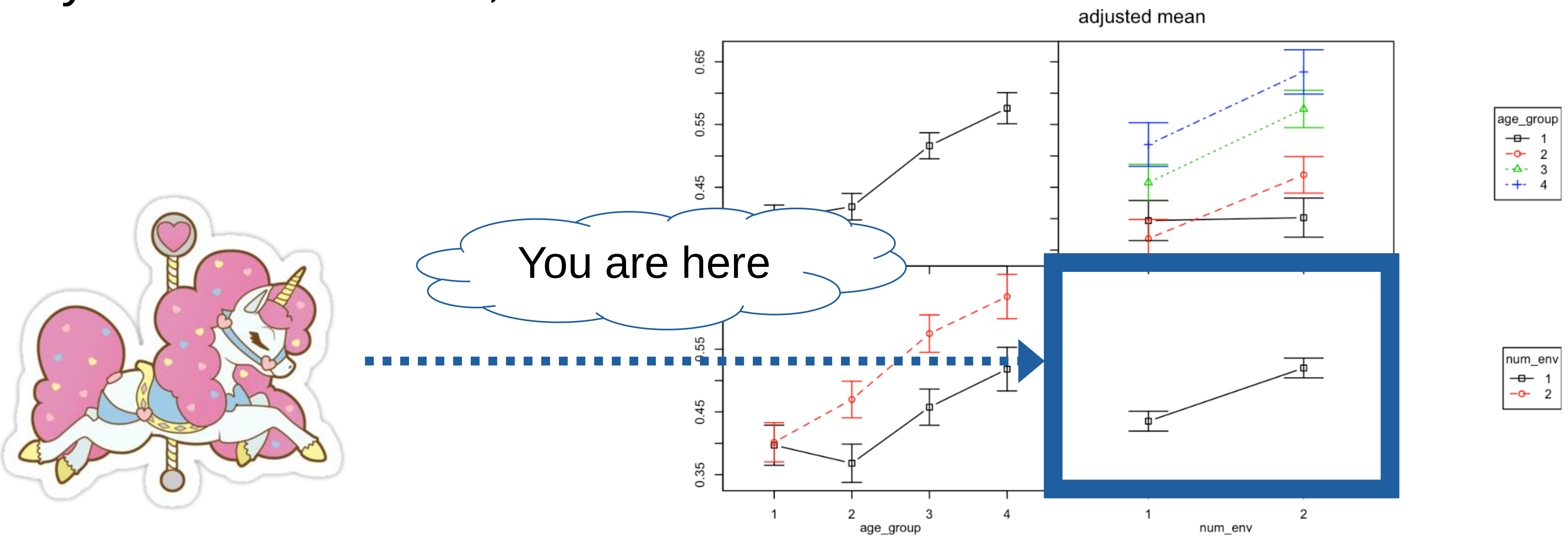
Back to our results…

```
# library(car)
sticker_int <- lm(prop_given ~ age_group*num_env, data = givers,
+                    contrasts = list(age_group = contr.sum,
+                                     num_env = contr.sum))
Anova(sticker_int, type = 2)
Anova Table (Type II tests)

Response: prop_given
                   Sum Sq  Df F value      Pr(>F)
age_group          1.5600   3 13.0371 0.000000047 ***
num_env            0.5940   1 14.8920   0.0001377 ***
age_group:num_env  0.1747   3  1.4599   0.2254215
Residuals         12.7636 320
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

No interaction-
the effect of each
variable did not
depend on the
level of the other

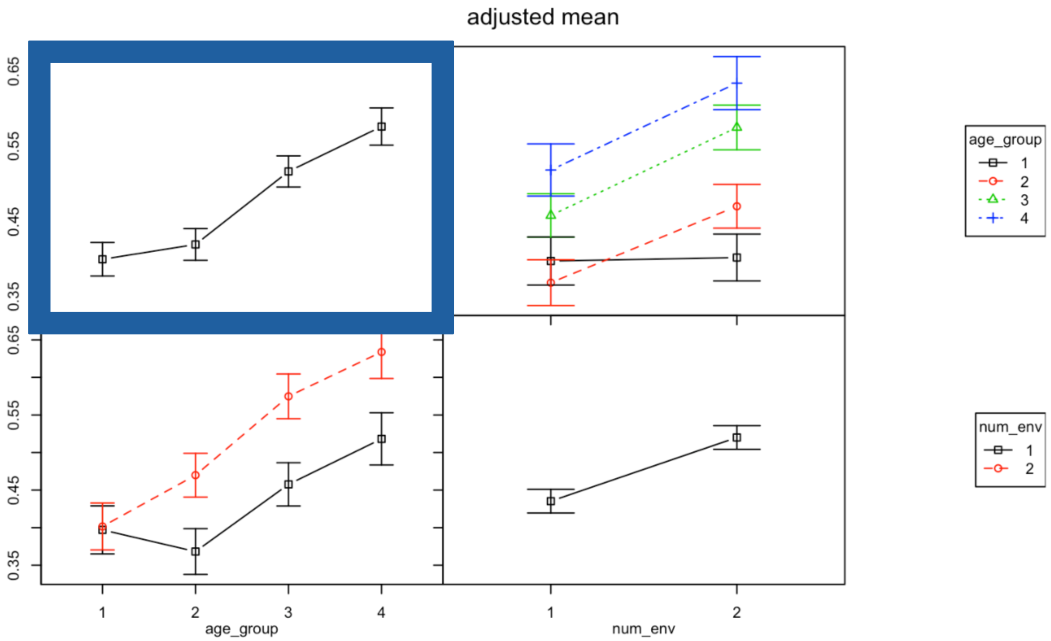
Interpreting effect of number of recipients

- No need!
- Only two levels, so F = 14.89 tells us that that proportion of stickers given was different when there was 1 vs. 2 recipients
- Plots/adjusted means tell us: 2 recipients > 1 (p = 0.0001)
- Note: if you did do a t-test, t = $\sqrt{F}$ = 3.86



Interpreting effect of age group

- F = 13 tells us that proportion of stickers given differed depending on age group, but which age groups were different from each other?
- Need post hoc comparisons to examine main effect of group



Follow-up contrasts: multcomp?

```
# library(multcomp)
```

- The mcp function must be used with care when defining parameters of interest in two-way ANOVA or ANCOVA models. Here, the definition of treatment differences (such as Tukey's all-pair comparisons or Dunnett's comparison with a control) might be problem specific. Because it is impossible to determine the parameters of interest automatically in this case, mcp in multcomp version 1.0-0 and higher generates comparisons for the main effects only, ignoring covariates and interactions (older versions automatically averaged over interaction terms).

- A warning is given.

Follow-up contrasts: phia?

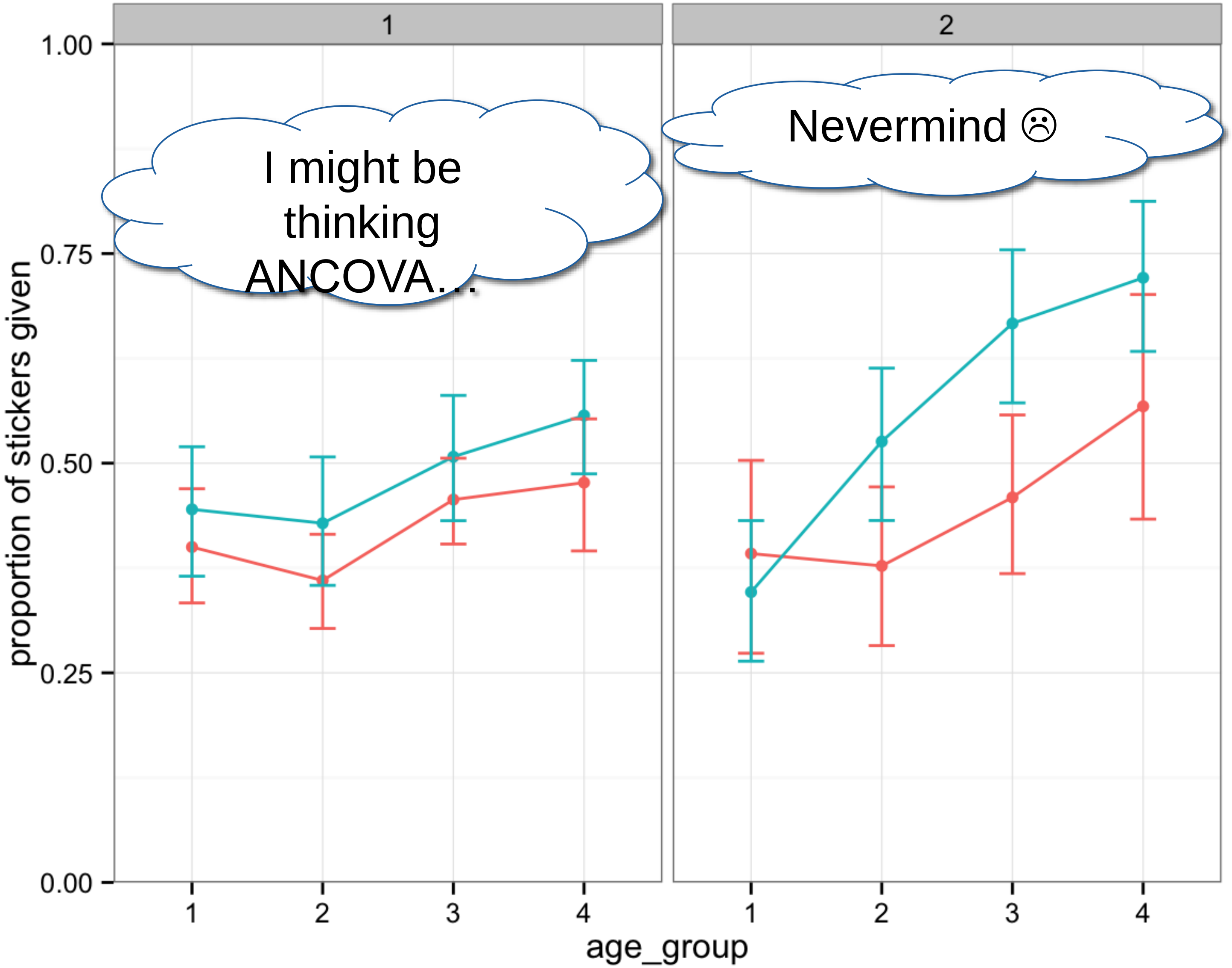
```
# library(phia)
interactionMeans(sticker_2int, factors =
"age_group")
  age_group adjusted mean std. error
1         1     0.3993173 0.02233588
2         2     0.4190376 0.02107271
3         3     0.5162269 0.02072034
4         4     0.5760180 0.02477462
```

Follow-up contrasts: phia?

```
# library(phia)
testInteractions(sticker_2int, pairwise = "age_group")
F Test:
P-value adjustment method: holm
                Value  Df Sum of Sq       F      Pr(>F)
1-2         -0.019720   1    0.0164  0.4124   0.5212039
1-3         -0.116910   1    0.5873 14.7247   0.0005996 ***
1-4         -0.176701   1    1.1193 28.0614 0.000001315 ***
2-3         -0.097189   1    0.4314 10.8150   0.0033566 **
2-4         -0.156980   1    0.9292 23.2954 0.000010775 ***
3-4         -0.059791   1    0.1367  3.4272   0.1301025
Residuals             320   12.7636
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Let’s include 3 predictors!

Girls = 1, Boys = 2
(again, unadjusted means)



```
# library(car)
sticker_3 <- lm(prop_given ~ age_group*num_env*gender, data = givers,
+                  contrasts = list(age_group = contr.sum, num_env =
contr.sum, gender = contr.sum))
Anova(sticker_3)
Anova Table (Type II tests)

Response: prop_given
                          Sum Sq  Df F value         Pr(>F)
age_group                 1.5400   3 13.3548 0.00000003188 ***
num_env                   0.6064   1 15.7770 0.00008854555 ***
gender                    0.2059   1  5.3556       0.02131 *
age_group:num_env         0.1591   3  1.3795       0.24902
age_group:gender          0.3371   3  2.9234       0.03413 *
num_env:gender            0.0701   1  1.8242       0.17779
age_group:num_env:gender  0.1652   3  1.4325       0.23325
Residuals                11.9924 312
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

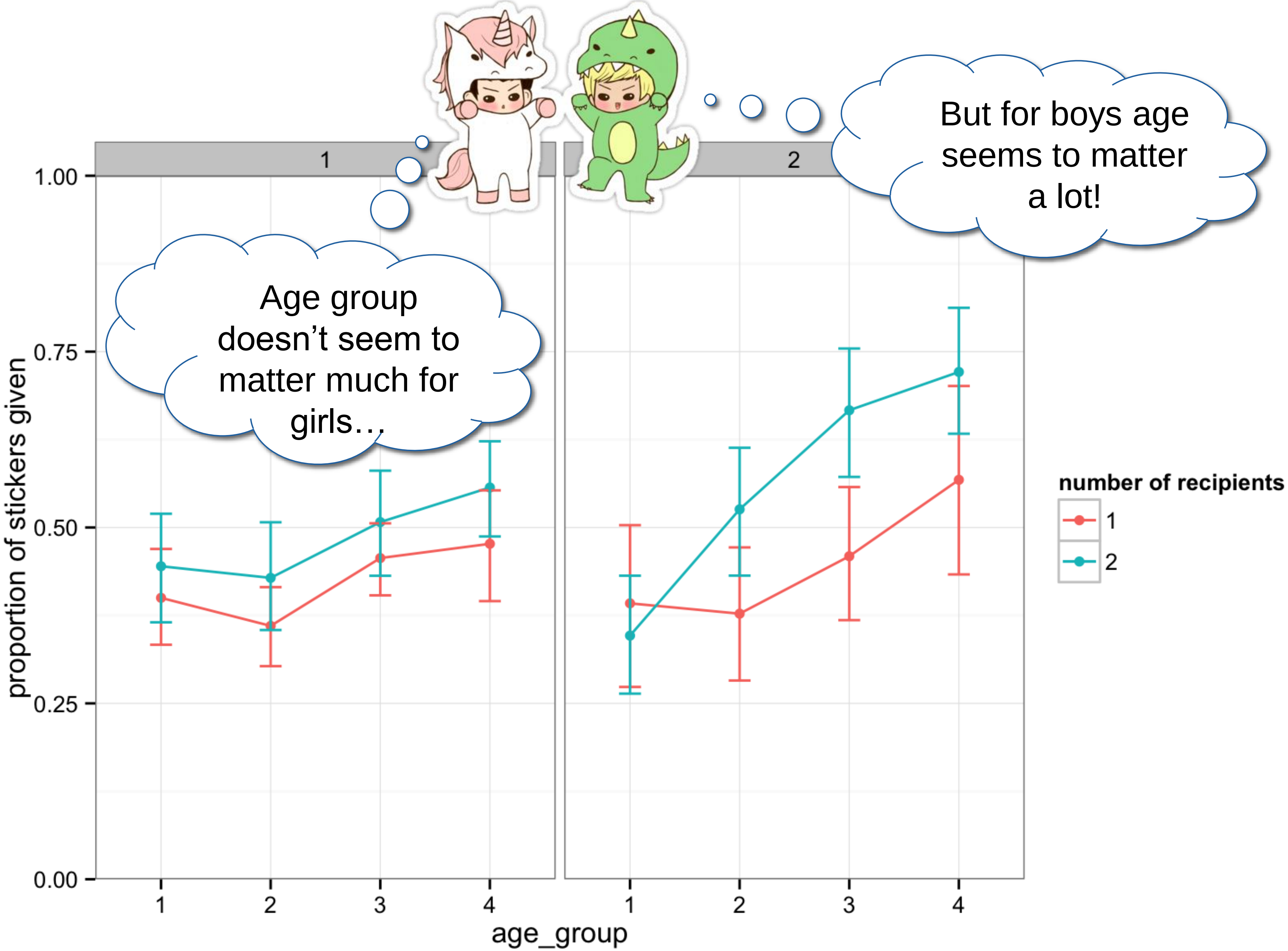
We have a significant interaction effect!

```
Sum Sq  Df F value          Pr(>F)
age_group                1.5400   3 13.3548 0.00000003188 ***
num_env                  0.6064   1 15.7770 0.00008854555 ***
gender                   0.2059   1  5.3556       0.02131 *
age_group:num_env        0.1591   3  1.3795       0.24902
age_group:gender         0.3371   3  2.9234       0.03413 *
num_env:gender           0.0701   1  1.8242       0.17779
age_group:num_env:gender 0.1652   3  1.4325       0.23325
Residuals               11.9924 312
---
Signif. codes:  0 '**      1 '      01 '*' 0
```

The interaction between age group and gender means we cannot fully interpret those main effects without using words like “but” or “depends on”

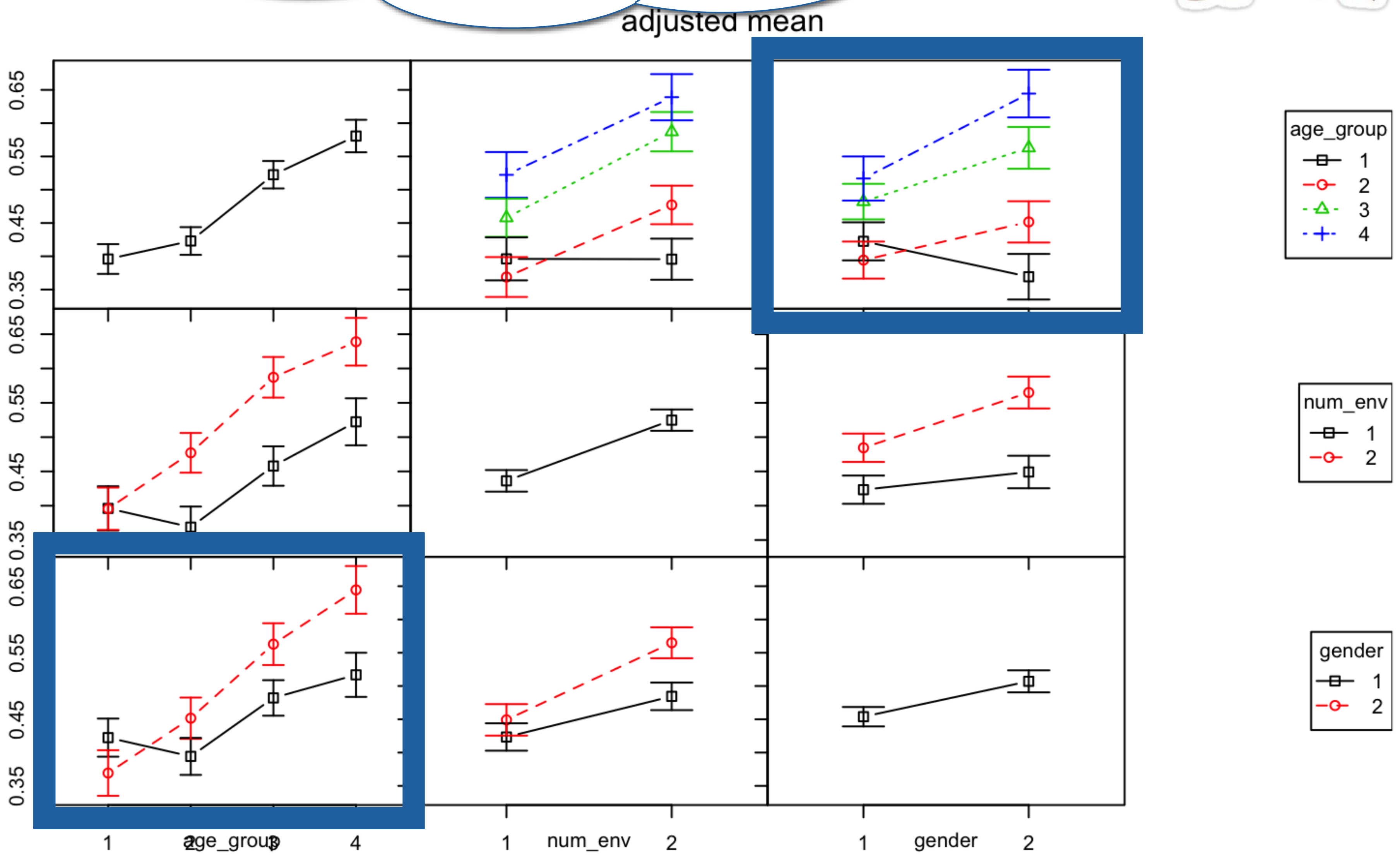
Principle of marginality

- The separate partial effects, or *main effects*, of age group and gender are *marginal* to the age group-by-gender interaction.
- In general, we neither test nor interpret main effects of explanatory variables that interact.
  - If we can rule out interaction either on theoretical or empirical grounds, then we can proceed to test, estimate, and interpret main effects.
- It does not generally make sense to specify and fit models that include interaction regressors but that delete main effects that are marginal to them.
  - Such models — which violate the *principle of marginality* — are interpretable, but they are not broadly applicable.
- **`?Anova`**: “Type-II tests are calculated according to the principle of marginality, testing each term after all others, except ignoring the term's higher-order relatives; so-called type-III tests violate marginality, testing each term in the model after all of the others.”



But for boys age seems to matter a lot!
Age group doesn't seem to matter much for girls…
1
2
1.00
0.75
0.50
0.25
0.00
proportion of stickers given
1
2
3
4
1
2
3
4
age_group
number of recipients
1
2

The effect of gender depends on age
&
the effect of age depends on gender



Adjusted interaction means

```
sticker_3int <- lm(prop_given ~ age_group*num_env*gender, data = givers,
                contrasts = list(age_group = contr.sum, num_env = contr.sum,
gender = contr.sum))
# library(phia)
interactionMeans(sticker_3int)
```

```
   age_group num_env gender adjusted mean std. error
1          1       1      1     0.4000000 0.04001938
2          2       1      1     0.3601449 0.04088011
3          3       1      1     0.4565476 0.03705074
4          4       1      1     0.4768519 0.04621039
5          1       2      1     0.4449275 0.04088011
6          2       2      1     0.4283951 0.03773063
7          3       2      1     0.5076923 0.03844937
8          4       2      1     0.5568627 0.04755010
9          1       1      2     0.3922222 0.05062095
10         2       1      2     0.3775000 0.04383903
11         3       1      2     0.4591667 0.04383903
12         4       1      2     0.5677778 0.05062095
13         1       2      2     0.3462963 0.04621039
14         2       2      2     0.5258333 0.04383903
15         3       2      2     0.6666667 0.04497790
16         4       2      2     0.7211111 0.05062095
```

Simple effects analysis

- A simple effects analysis looks at the main effect of one factor at a given level of a second factor (“pick a point” analysis)
- This is our way of breaking down a significant interaction
- It *can* be frowned upon to do this analysis post-hoc when you do not have a significant interaction effect in your omnibus ANOVA, or did not have a darn good reason to hypothesize one a priori

The effect of gender depends on age
&
the effect of age depends on gender

```
# library(phia)
testInteractions(sticker_3, fixed = "age_group", across = "gender", adjustment =
"bonferroni")
F Test:
P-value adjustment method: bonferroni
               Value  Df  Sum of Sq      F   Pr(>F)
1           0.053205   1     0.0546 1.4206 0.93686
2          -0.057397   1     0.0730 1.8992 0.67663
3          -0.080797   1     0.1477 3.8423 0.20345
4          -0.127587   1     0.2629 6.8387 0.03741 *
Residuals            312    11.9924
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Boys shared 12.7% more stickers than girls on average at ages 9 – 11…
where does this number come from?

```
interactionMeans(sticker_3int)
   age_group num_env gender adjusted mean std. error
1          1       1      1     0.4000000 0.04001938
2          2       1      1     0.3601449 0.04088011
3          3       1      1     0.4565476 0.03705074
4          4       1      1     0.4768519 0.04621039
5          1       2      1     0.4449275 0.04088011
6          2       2      1     0.4283951 0.03773063
7          3       2      1     0.5076923 0.03844937
8          4       2      1     0.5568627 0.04755010
9          1       1      2     0.3922222 0.05062095
10         2       1      2     0.3775000 0.04383903
11         3       1      2     0.4591667 0.04383903
12         4       1      2     0.5677778 0.05062095
13         1       2      2     0.3462963 0.04621039
14         2       2      2     0.5258333 0.04383903
15         3       2      2     0.6666667 0.04497790
16         4       2      2     0.7211111 0.05062095
```

Girls

Boys

```
(0.4768519 + 0.5568627)/2 - (0.5677778 + 0.7211111)/2
[1] -0.1275871
```

Boys shared 12.7% more stickers than girls on average at ages 9 – 11…
where does this number come from? See above!

The effect of gender depends on age
&
the effect of age depends on gender

```
# library(phia)
testInteractions(sticker_3, fixed = "gender", across = "age_group", adjustment = "bonferroni")
F Test:
P-value adjustment method: bonferroni
          age_group1 age_group2 age_group3  Df  Sum of Sq       F        Pr(>F)
1          -0.094394   -0.12259  -0.034737   3     0.4019   3.4851       0.03236 *
2          -0.275185   -0.19278  -0.081528   3     1.4307 12.4076 0.0000002189 ***
Residuals                                   312   11.9924
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

This is not so helpful- but does show significant effect of age is there for both boys and girls;
more breakdown is needed…

The effect of gender depends on age
&
the effect of age depends on gender

```
# library(phia)
testInteractions(sticker_3, pairwise = "age_group", fixed = "gender", adjustment = "bonferroni")
F Test:
P-value adjustment method: bonferroni
             Value  Df Sum of Sq       F      Pr(>F)
1-2 : 1   0.028194   1    0.0192  0.4993   1.0000000
1-3 : 1  -0.059656   1    0.0894  2.3246   1.0000000
1-4 : 1  -0.094394   1    0.1786  4.6473   0.3824087
2-3 : 1  -0.087850   1    0.1996  5.1919   0.2804164
2-4 : 1  -0.122587   1    0.3084  8.0242   0.0589929 .
3-4 : 1  -0.034737   1    0.0256  0.6660   1.0000000
1-2 : 2  -0.082407   1    0.1222  3.1802   0.9060965
1-3 : 2  -0.193657   1    0.6672 17.3571   0.0004817 ***
1-4 : 2  -0.275185   1    1.1853 30.8371 0.0000007213 ***
2-3 : 2  -0.111250   1    0.2443  6.3563   0.1463387
2-4 : 2  -0.192778   1    0.6371 16.5747   0.0007121 ***
3-4 : 2  -0.081528   1    0.1127  2.9314   1.0000000
Residuals          312   11.9924
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Significant differences in proportion of stickers given differs between certain age groups, but only observed among boys not girls

Effect size

- Omnibus: ω^2

```
omega.squared(sticker_3int) # seriously ugly
[1] 0.1119718
```

- But also needed for all contrasts! Need either mean difference, t, or F statistic to calculate **Cohen's d** for standard ANOVA/ANCOVA

```
library(compute.es)
```

What do we conclude?

- Overall, model accounted for 11.2% of variance in proportion of stickers donated
- Main effect: proportion of stickers donated depended on number of envelopes
  - Kids tended to share more proportionally when they thought there were 2 recipients rather than just 1
- Interaction effect: suggested that…
  - Boys shared proportionally more than girls, but only in the oldest age group (age_group = 4)
  - For boys only, increased sharing at older ages, but only between certain age groups:
    - 1 vs 3
    - 1 vs 4
    - 2 vs 4

Interactions with continuous variables…

- Let's take age as a continuous variable (in months) rather than as a factor with 4 levels

```
sticker_spot <- lm(prop_given ~ age*gender*num_env, data = givers, contrasts = list(gender =
contr.sum, num_env = contr.sum))
Anova(sticker_spot, type = 2)
```

```
Sum Sq  Df F value          Pr(>F)
age                  1.7416   1 46.5137 0.0000000000455 ***
gender               0.1804   1  4.8175       0.0288896 *
num_env              0.5689   1 15.1934       0.0001182 ***
age:gender           0.2815   1  7.5184       0.0064506 **
age:num_env          0.1256   1  3.3538       0.0679813 .
gender:num_env       0.0734   1  1.9615       0.1623209
age:gender:num_env   0.0713   1  1.9044       0.1685502
Residuals           11.9814 320
```

For spotlight analysis, you'll want to **dummy code** (not deviation/effect code) your contrasts, and use `summary()` or `tidy()`

Spotlight analysis for ANOVA

- We'll use a spotlight analysis to better understand the interaction between gender (girls/boys) and age (in months)

```
sticker_spot <- lm(prop_given ~ age*gender*num_env, data = givers, contrasts = list(gender =
contr.sum, num_env = contr.sum))
Anova(sticker_spot, type = 2)
Anova Table (Type II tests)
Response: prop_given
                    Sum Sq  Df F value           Pr(>F)
age                 1.7416   1 46.5137 0.0000000000455 ***
gender              0.1804   1  4.8175       0.0288896 *
num_env             0.5689   1 15.1934       0.0001182 ***
age:gender          0.2815   1  7.5184       0.0064506 **
age:num_env         0.1256   1  3.3538       0.0679813 .
gender:num_env      0.0734   1  1.9615       0.1623209
age:gender:num_env  0.0713   1  1.9044       0.1685502
Residuals          11.9814 320
```

Spotlight analysis to understand interactions

- *"scaling changes do not affect significance tests, slopes, etc. of that variable or of any other variable in the model."*
- Yes, this heuristic is accurate for simple models without interactions, but **not for models with interactions**.
- Spotlight analysis exploits this fact by re-scaling your continuous covariate (here, age in months)
- We want to evaluate main effect of gender when age takes on different values- the interaction tells us that the main effect of gender depends on the age we are looking at
- We can pick any value to re-scale by
  - First, I'll center age around the mean age
  - Can also use ±1 SD around the mean (usually of interest)

```
# center age at mean
givers <- givers %>%
  mutate(age_mean = age - mean(age))
sticker_ctr <- lm(prop_given ~ age_mean*gender*num_env, data = givers, contrasts =
list(gender = contr.treatment, num_env = contr.treatment))
tidy(sticker_ctr)
                    term     estimate   std.error  statistic      p.value
1            (Intercept) 0.4233064576 0.0200749628 21.0862885 1.727537e-62
2               age_mean 0.0014530306 0.0007301193  1.9901278 4.742745e-02
3                gender2 0.0170252154 0.0306598615  0.5552933 5.790820e-01
4                num_env2 0.0563245651 0.0283894013  1.9839998 4.810997e-02
5        age_mean:gender2 0.0010712428 0.0011078819  0.9669287 3.343099e-01
6       age_mean:num_env2 0.0004548377 0.0010525807  0.4321167 6.659476e-01
7         gender2:num_env2 0.0599339324 0.0431639747  1.3885175 1.659453e-01
8 age_mean:gender2:num_env2 0.0021650305 0.0015688634  1.3799994 1.685502e-01
```

Shine your spotlight on the gender effect:
It is not significant when age = mean
age. So sharing in boys is
0.017 > girls(ns) @ 81 months (6.7
years)
when num_env = 1

```
# center at 1 sd below mean
givers <- givers %>%
+     mutate(age_lowsd = age - (mean(age) - sd(age)))
sticker_lowsd <- lm(prop_given ~ age_lowsd*gender*num_env, data = givers, contrasts =
list(gender = contr.treatment, num_env = contr.treatment))
tidy(sticker_lowsd)
```

	term	estimate	std.error	statistic	p.value
1	(Intercept)	0.3834153744	0.0279165665	13.734331329	4.587631e-34
2	age_lowsd	0.0014530306	0.0007301193	1.990127755	4.742745e-02
3	gender2	-0.0123843751	0.0436644743	-0.283625883	7.768805e-01
4	num_env2	0.0438375824	0.0398851848	1.099094378	2.725531e-01
5	age_lowsd:gender2	0.0010712428	0.0011078819	0.966928707	3.343099e-01
6	age_lowsd:num_env2	0.0004548377	0.0010525807	0.432116699	6.659476e-01
7	gender2:num_env2	0.0004958102	0.0612413124	0.008096009	9.935454e-01
8	age_lowsd:gender2:num_env2	0.0021650305	0.0015688634	1.379999352	1.685502e-01

Shine your spotlight on the gender effect: It is not significant when age is 1 SD below mean age. So sharing is 0.012 < (ns) in boys than girls @ 53 months (4.4 years) when num_env = 1

```
# center at 1 sd above mean
givers <- givers %>%
+     mutate(age_hisd = age - (mean(age) + sd(age)))
sticker_hisd <- lm(prop_given ~ age_hisd*gender*num_env, data = givers, contrasts =
list(gender = contr.treatment, num_env = contr.treatment))
tidy(sticker_hisd)
                     term     estimate    std.error  statistic      p.value
1             (Intercept) 0.4631975409 0.0288138200 16.0755339 4.864763e-43
2                age_hisd 0.0014530306 0.0007301193  1.9901278 4.742745e-02
3                 gender2 0.0464348058 0.0427044430  1.0873530 2.776991e-01
4                 num_env2 0.0688115479 0.0411240801  1.6732666 9.525206e-02
5         age_hisd:gender2 0.0010712428 0.0011078819  0.9669287 3.343099e-01
6        age_hisd:num_env2 0.0004548377 0.0010525807  0.4321167 6.659476e-01
7         gender2:num_env2 0.1193720545 0.0607124566  1.9661872 5.014113e-02
8 age_hisd:gender2:num_env2 0.0021650305 0.0015688634  1.3799994 1.685502e-01
```

Shine your spotlight on the gender effect:
It is significant when age is 1 SD above
mean age. So sharing in boys is
0.046 > girls @ 108 months (9 years)
when num_env = 1

Cautionary notes on spotlight analysis

- Unless you are using them as a classification variable (i.e., they are not being used as actual numbers), dummy codes are **only** useful if you are planning to spotlight 1 group in an interaction
- Otherwise, all of your purported "main effects" are actually the effects of a variable at one level of the other (i.e., when the other is 0).
- Caution: if there are other variables in your model, recall that the main effect when dummy coded is just a marginal effect where the other variables are set to lowest level (here, num_env = 1)

They are called dummy codes for a reason ☺ That is, you may not be testing what you think you are testing.

What about when num_env = 2?
The spotlight method works best for a two-way analysis with two factors or one factor and one continuous variable.

What we really want to know is how the interaction between age and gender can be interpreted, averaging across number of envelopes (since that does not contribute to any interactions with other model predictors). There has to be a better way…

```
sticker_spot <- lm(prop_given ~ age*gender*num_env, data = givers, contrasts =
list(gender = contr.sum, num_env = contr.sum))
interactionMeans(sticker_spot, factors = "gender", covariates = c(age = 80.9878))
  gender adjusted mean std. error
1      1     0.4514687 0.01419470
2      2     0.4984609 0.01625708
```

@ mean age (81 mos):
diff = -.047

```
interactionMeans(sticker_spot, factors = "gender", covariates = c(age = 53))
  gender adjusted mean std. error
1      1     0.4044366 0.02014112
2      2     0.3911499 0.02345927
```

@ 1 sd below mean (53 mos):
diff = .013

```
interactionMeans(sticker_spot, factors = "gender", covariates = c(age = 108))
  gender adjusted mean std. error
1      1     0.4968614 0.02039448
2      2     0.6020313 0.02215622
```

@ 1 sd above mean (108 mos):
diff = -.105

Here, the default is to calculate adjusted mean values across all unspecified other factors in your model, so this is averaging across num_env 1 & 2

```
testInteractions(sticker_spot, pairwise = "gender", covariates = c(age = 80.9878),
adjustment = "none") # at mean age
F Test:
P-value adjustment method: none
              Value  Df  Sum of Sq      F   Pr(>F)
1-2       -0.046992   1     0.1775  4.741  0.03018 *
Residuals           320    11.9814
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

@ mean age (81 mos):
diff = -.047

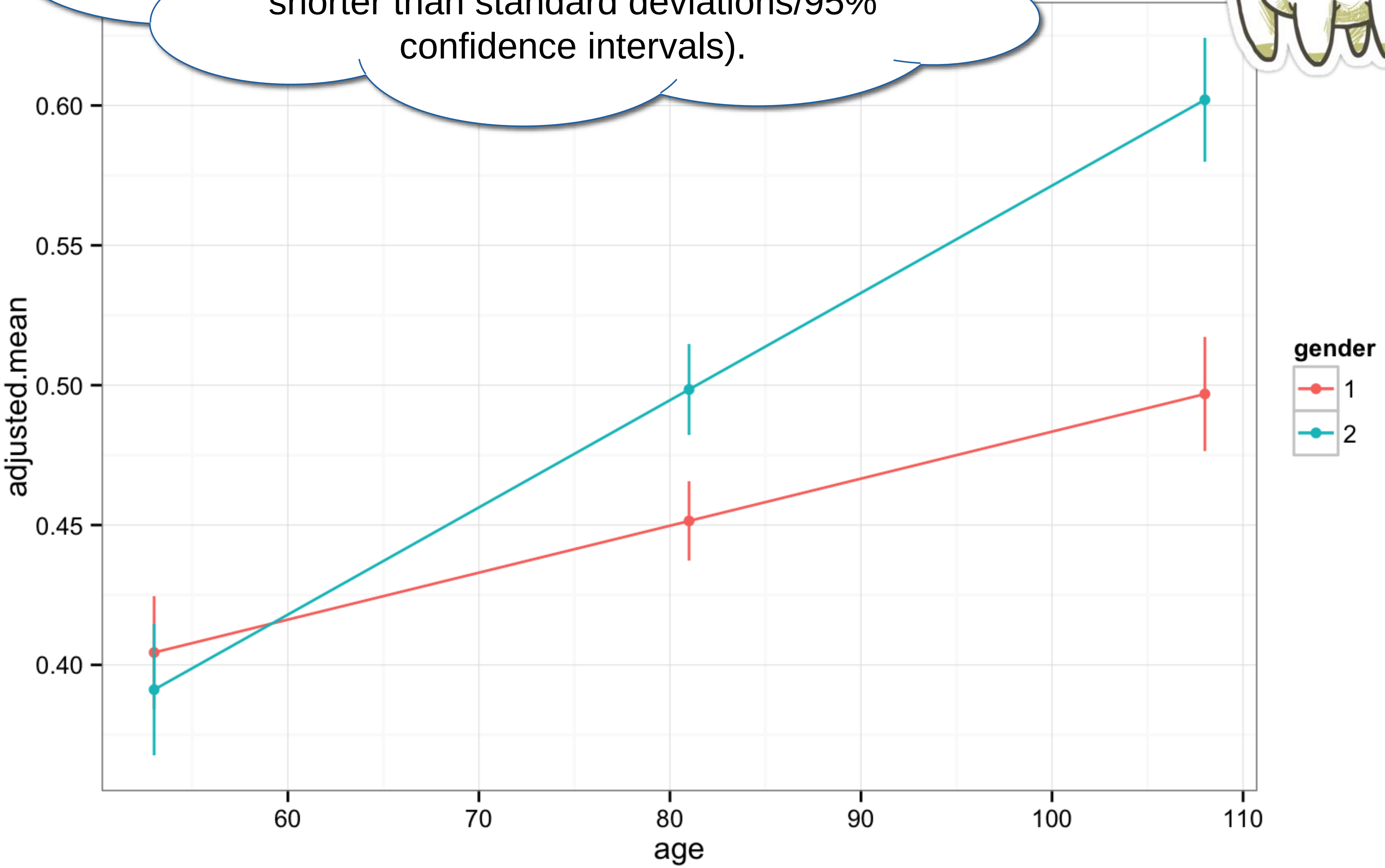
```
testInteractions(sticker_spot, pairwise = "gender",  covariates = c(age = 53),
adjustment = "none")
F Test:
P-value adjustment method: none
             Value  Df  Sum of Sq       F  Pr(>F)
1-2       0.013287   1     0.0069  0.1847  0.6677
Residuals          320    11.9814
```

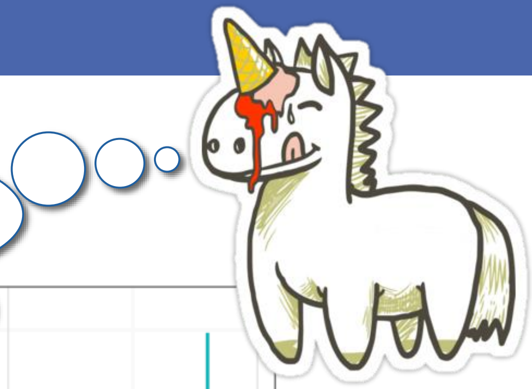
@ 1 sd below mean (53 mos):
diff = .013

```
testInteractions(sticker_spot, pairwise = "gender",  covariates = c(age = 108),
adjustment = "none")
F Test:
P-value adjustment method: none
             Value  Df  Sum of Sq       F    Pr(>F)
1-2       -0.10517   1     0.4567  12.197  0.000546 ***
Residuals          320    11.9814
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

@ 1 sd above mean (108 mos):
diff = -.105

You can plot these adjusted conditional means! Use the estimated standard errors (realize these bars will always be shorter than standard deviations/95% confidence intervals).





Spotlight analysis summary

- Great way to understand interactions with continuous covariates; also works for factors
- Powerful way to see what is happening **without subsetting your data**
    - i.e., splitting by above/below mean/median/etc can be weak, unstable, potentially very misleading
- This is not multiple testing- you do not need to worry about multiple comparisons here- it is used to interpret an already significant interaction
- So don't report them as if you ran different analyses with different results