Contrast coding

Professor Andy Field

@profandyfield

www.youtube.com/user/ProfAndyField/

www.discoveringstatistics.com
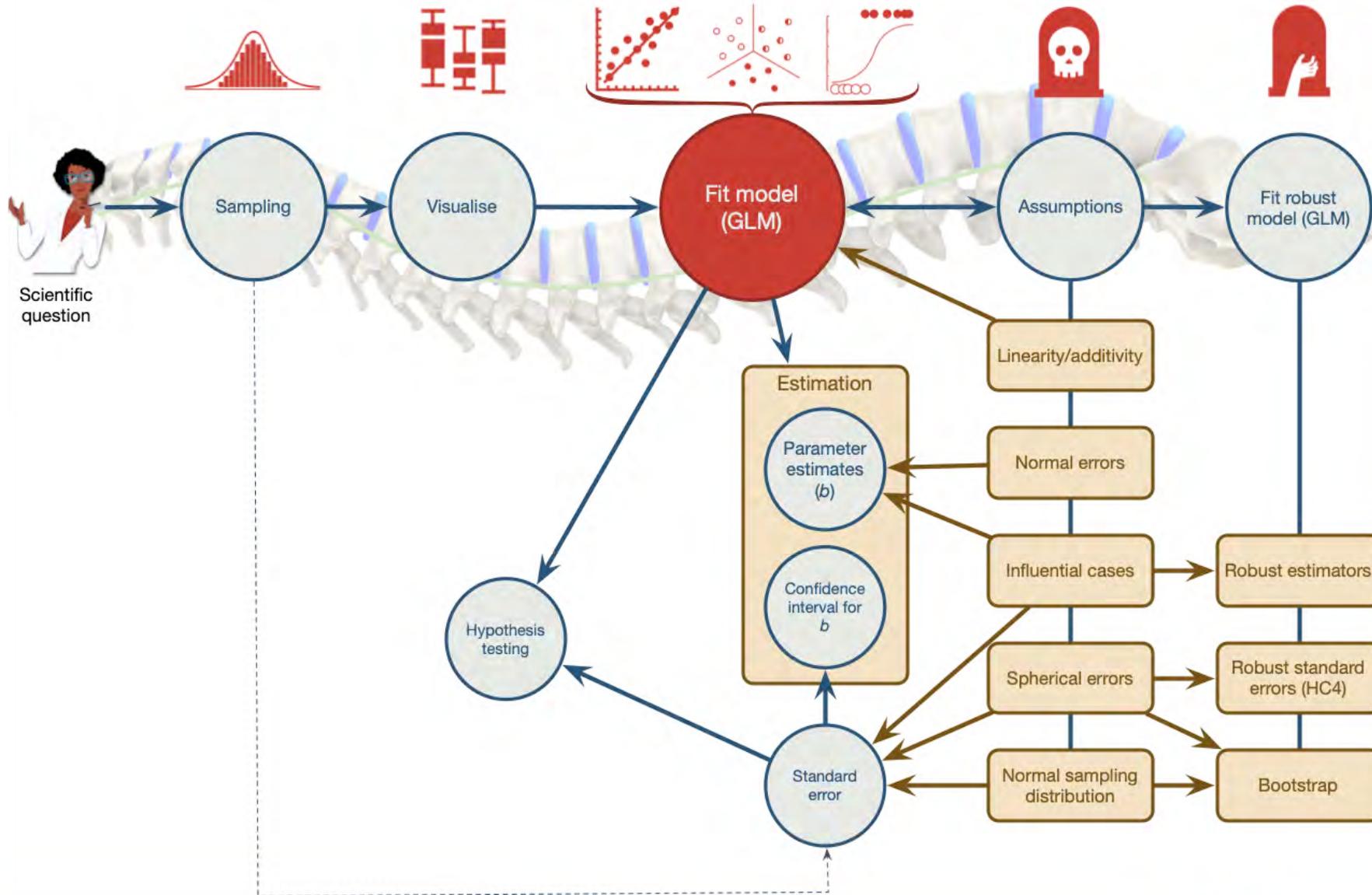
www.milton-the-cat.rocks

www.discovr.rocks

Learning outcomes

- Explain the different ways to break down categorical predictors in a linear model

  - Planned contrasts/comparisons)
  - Choosing contrasts
  - Contrast coding

- *Post hoc* tests

- Polynomial contrasts (trend analysis)

Contrast coding

- The *F*-statistic tests the overall fit of the model

    - i.e. It is a general test of model fits/whether group means significantly differ

- Model parameters tells us about specific differences between means

    - Dummy coding compares each category to a baseline

- What do we do when dummy coding does not reflect our *a priori* hypotheses?

Options for breaking down categorical predictors

- Orthogonal contrasts (contrast coding)

  - Hypothesis driven
  - Planned *a priori*
  - Control Type I error rate

- *Post hoc* tests

  - Not planned (not hypothesis driven)
  - Compare all pairs of means
  - Multiple *t*-tests adjusted for the number of tests

- Trend analysis

  - Useful only for ordered means

A puppy-tastic example

- A puppy therapy RCT in which we randomized people into three groups:

  - A control group
  - 15 minutes of puppy therapy
  - 30 minutes of puppy contact.

- The outcome is happiness (0 = unhappy) to 10 (happy).

- Predictions:
  - Any form of puppy therapy should be better than the control (i.e. higher happiness scores)
  - A dose-response hypothesis that as exposure time increases (from 15 to 30 minutes) happiness will increase too.

The data

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

	No puppies	15 mins	30 mins
	3	5	7
	2	2	4
	1	4	5
	1	2	3
	4	3	6
Mean	2.20	3.20	5.00
Variance (s^2)	1.70	1.70	2.50
Standard deviation (s)	1.30	1.30	1.58

Overall mean $(\bar{X}_{\text{grand}}) = 3.467$

The general linear model

Dummy coding

Therapy group	Long (30 mins vs. no puppies)	Short 1 (15 mins vs. no puppies)
No Puppies	0	0
15 mins	0	1
30 mins	1	0

$$\hat{\text{Happiness}}_i = \hat{b}_0 + \hat{b}_1 \text{Long}_i + \hat{b}_2 \text{Short}_i + e_i$$
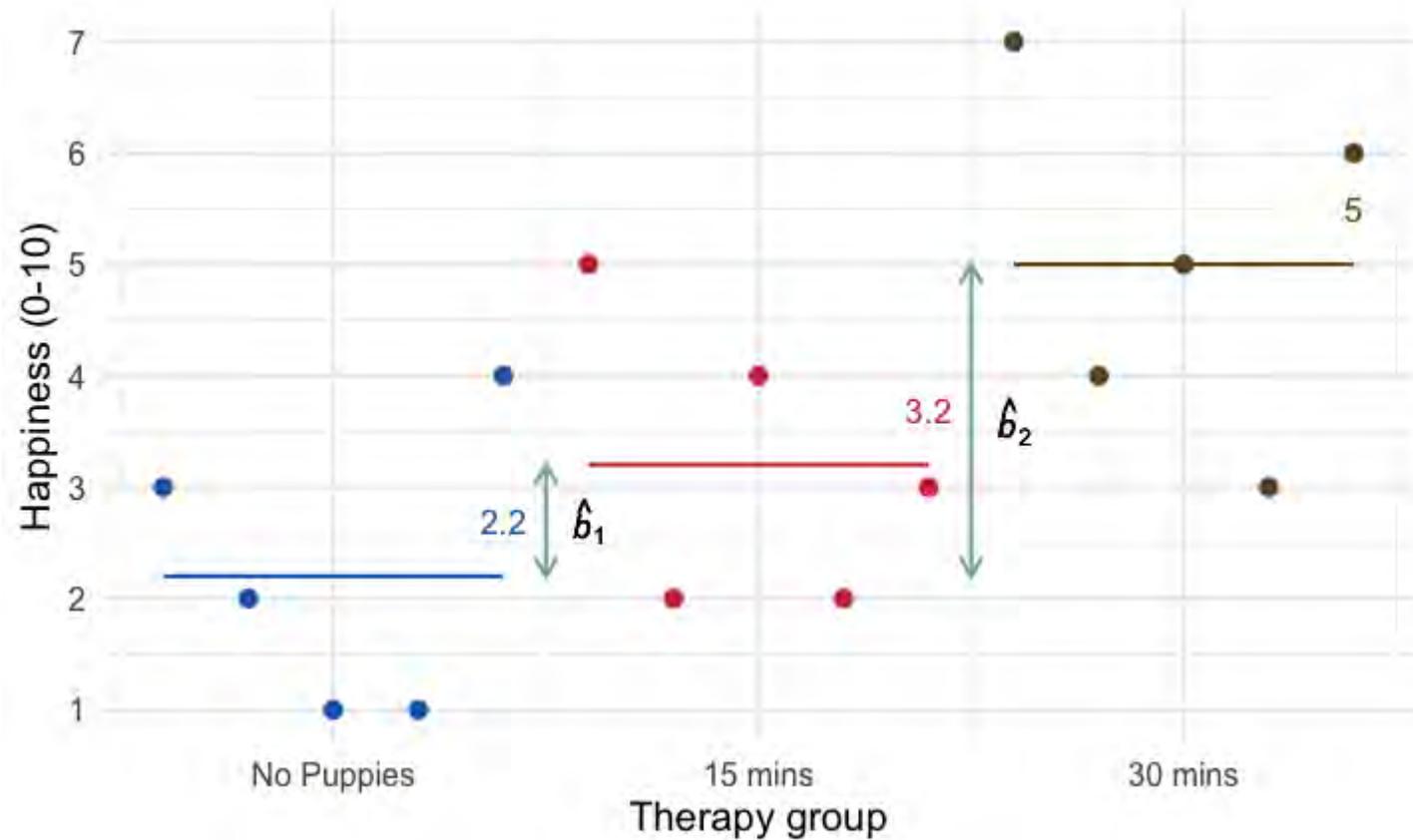
The 'dummy' model
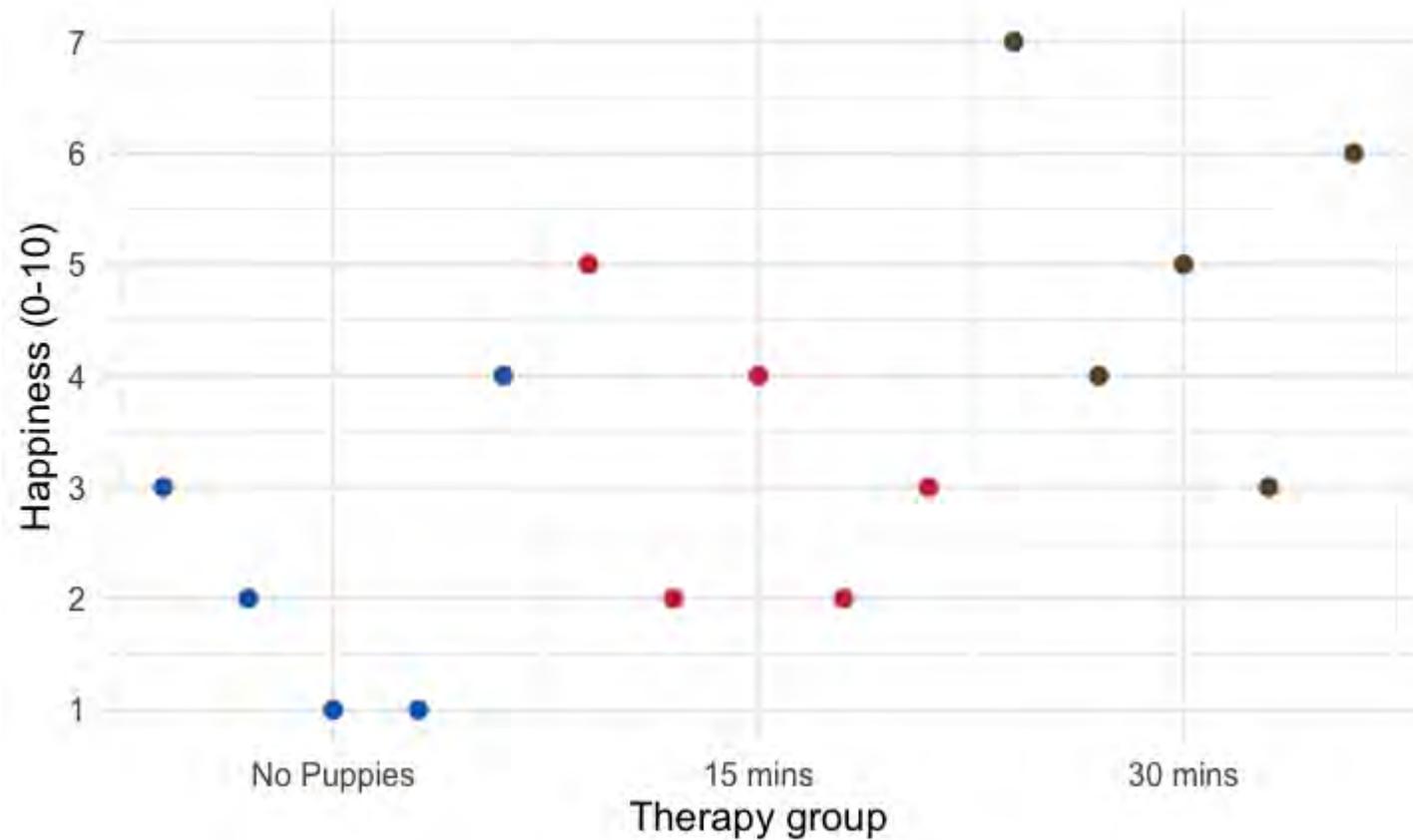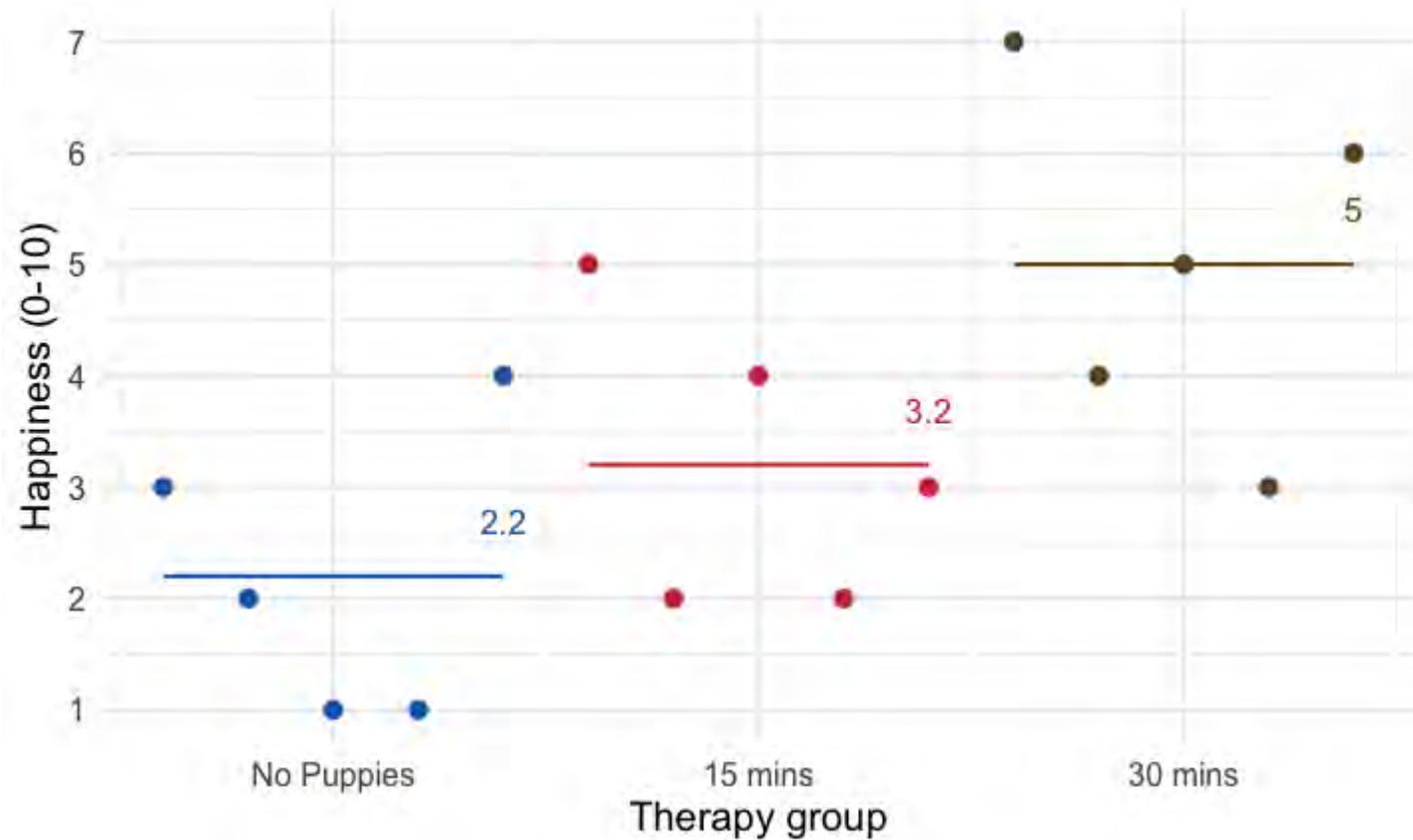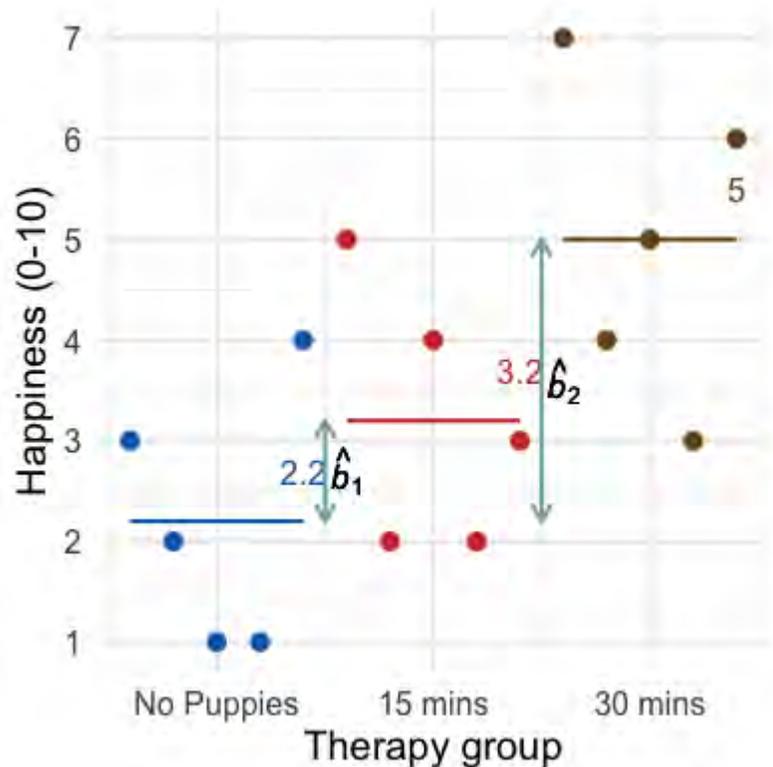
The 'dummy' model
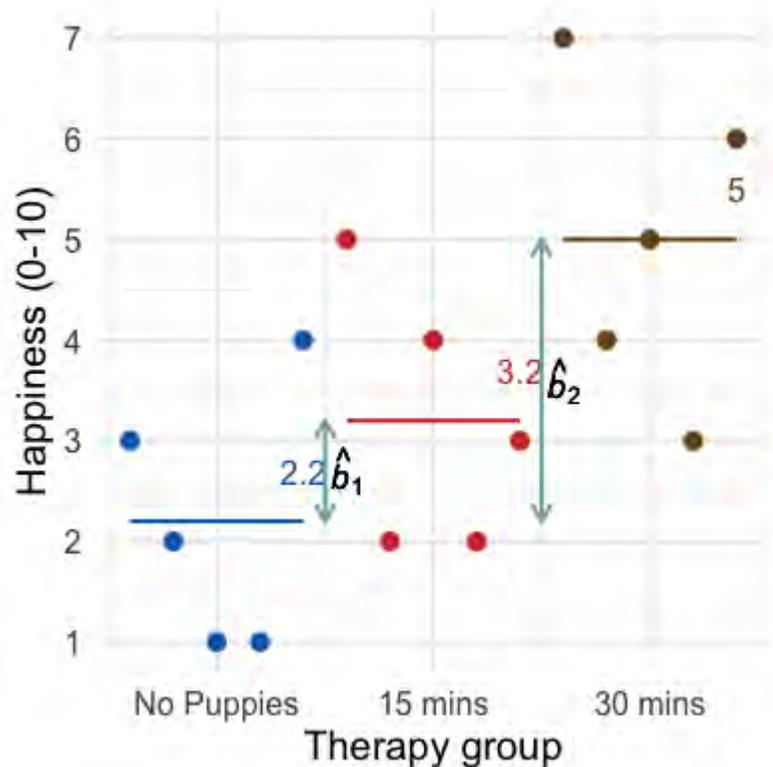
The 'dummy' model

The model

$$\text{Happ\hat{i}ness}_i = \hat{b}_0 + \hat{b}_1 \text{Long}_i + \hat{b}_2 \text{Short}_i + e_i$$

The model

$$\text{Happ\hat{i}ness}_i = \hat{b}_0 + \hat{b}_1\text{Long}_i + \hat{b}_2\text{Short}_i + e_i$$

$$\hat{b}_0 = \bar{X}_{\text{No puppies}}$$
$$\hat{b}_1 = 5.0 - 2.2 = 2.8$$
$$\hat{b}_2 = 3.2 - 2.2 = 1.0$$

Model fit: *F*-statistic

```
puppy_lm <- lm(happiness ~ dose, data = puppy_tib)
broom::glance(puppy_lm)
```

r.squared	adj.r.squared	sigma	statistic	p.value	df	logLik	AIC	BIC	deviance	df.residual	nobs
0.46	0.37	1.4	5.12	0.02	2	-24.68	57.37	60.2	23.6	12	15

Parameter estimates

```
broom::tidy(puppy_lm, conf.int = TRUE)
```

| term	estimate	std.error	statistic	p.value	conf.low	conf.high
(Intercept)	2.2	0.63	3.51	0.00	0.83	3.57
dose15 mins	1.0	0.89	1.13	0.28	-0.93	2.93
dose30 mins	2.8	0.89	3.16	0.01	0.87	4.73

Planned contrasts

- The variability explained by the model, SS_M is due to participants being assigned to different groups

  - This variability sometimes represents an experimental manipulation

- This variability (SS_M) can be broken down further to test specific hypotheses about which groups might differ

- We break down the variance according to hypotheses made *a priori* (before the experiment)

- It's like cutting up a cake (yum yum!)

The cake analogy again

SS_M
Improvement due to Model

SS_R
Error in Model

SS_T
Total variance
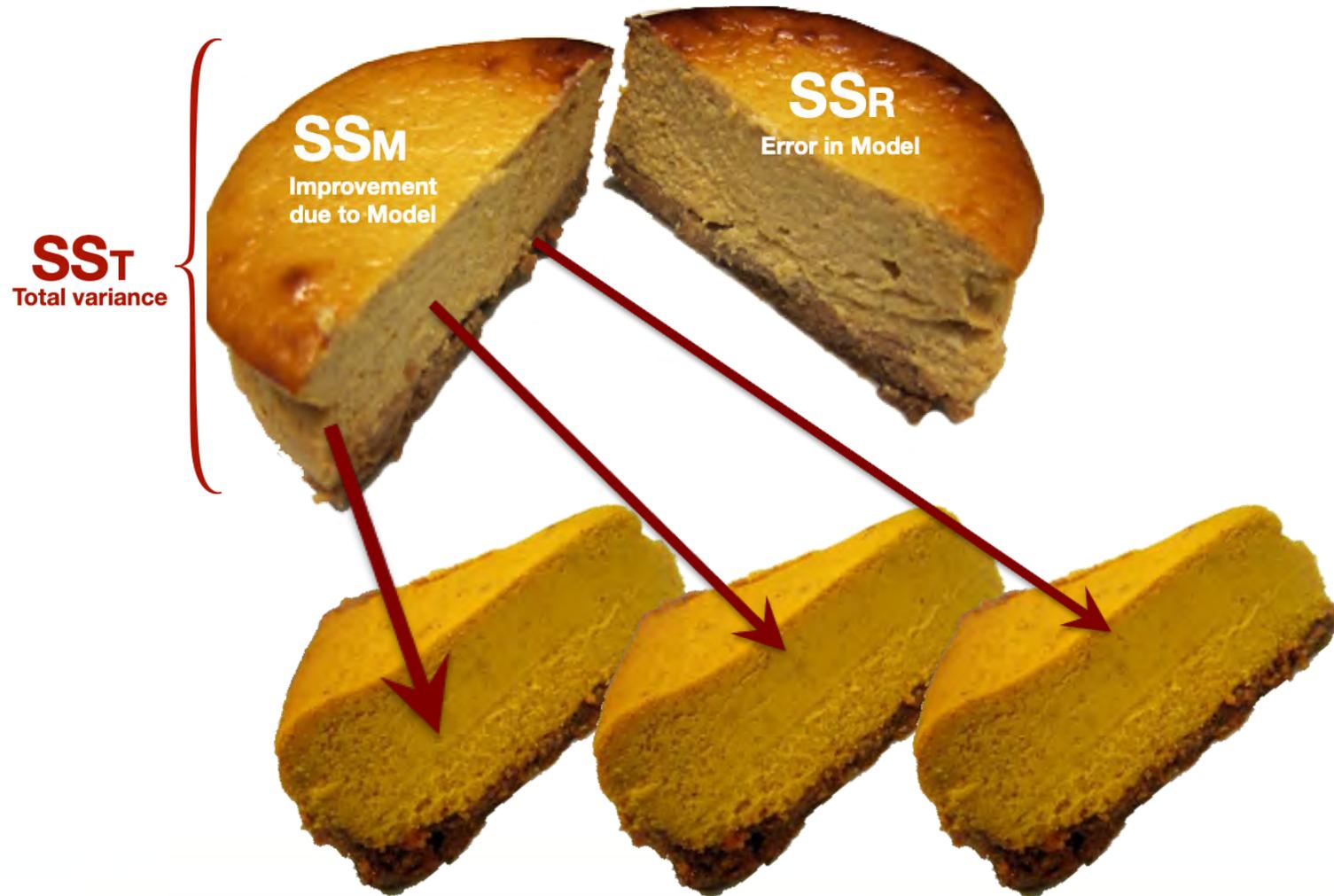
Choosing contrasts

- Independent

  - To control Type I error rates contrasts must be independent (they must test unique hypotheses)
  - If a group is singled out in a contrast, then that group should not be used in any subsequent contrasts

- Only 2 Chunks

  - Each contrast should compare only 2 chunks of variation (why?)

- *K*-1
  - You should always end up with one less contrast than the number of groups

How do I choose contrasts?

Big Hint:

- Most experimental designs typically have one or more control groups

- The logic of control groups means that we expect scores within them to differ from those in the groups we've manipulated

- The first contrast will usually compare any control conditions (chunk 1) with any experimental ones (chunk 2)

Hypotheses

Hypothesis 1:

- People who have puppy therapy will be happier (have have higher happiness scores) than those who don't

- Control $\neq$ (15 mins, 30 mins)

Hypothesis 2:

- People receiving a high dose of puppy therapy (30 mins) will be happier than those receiving a low dose (15 mins)
- 15 mins $\neq$ 30 mins

SS$_M$
Variance explained by the experiment
Four groups: E1, E2, C1 and C2

Experimental Groups
E1 and E2

Control Groups
C1 and C2

Contrast 1

E1

E2

Contrast 2

C1
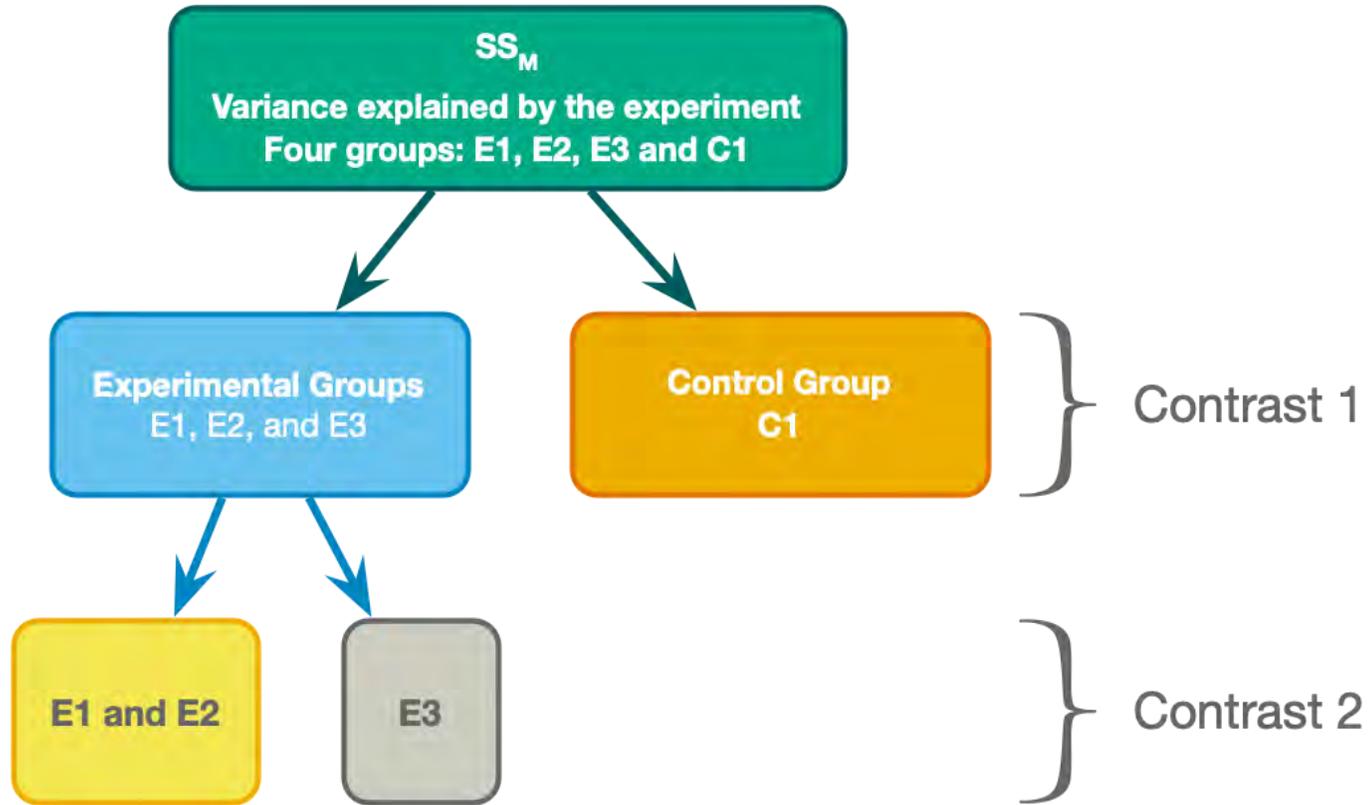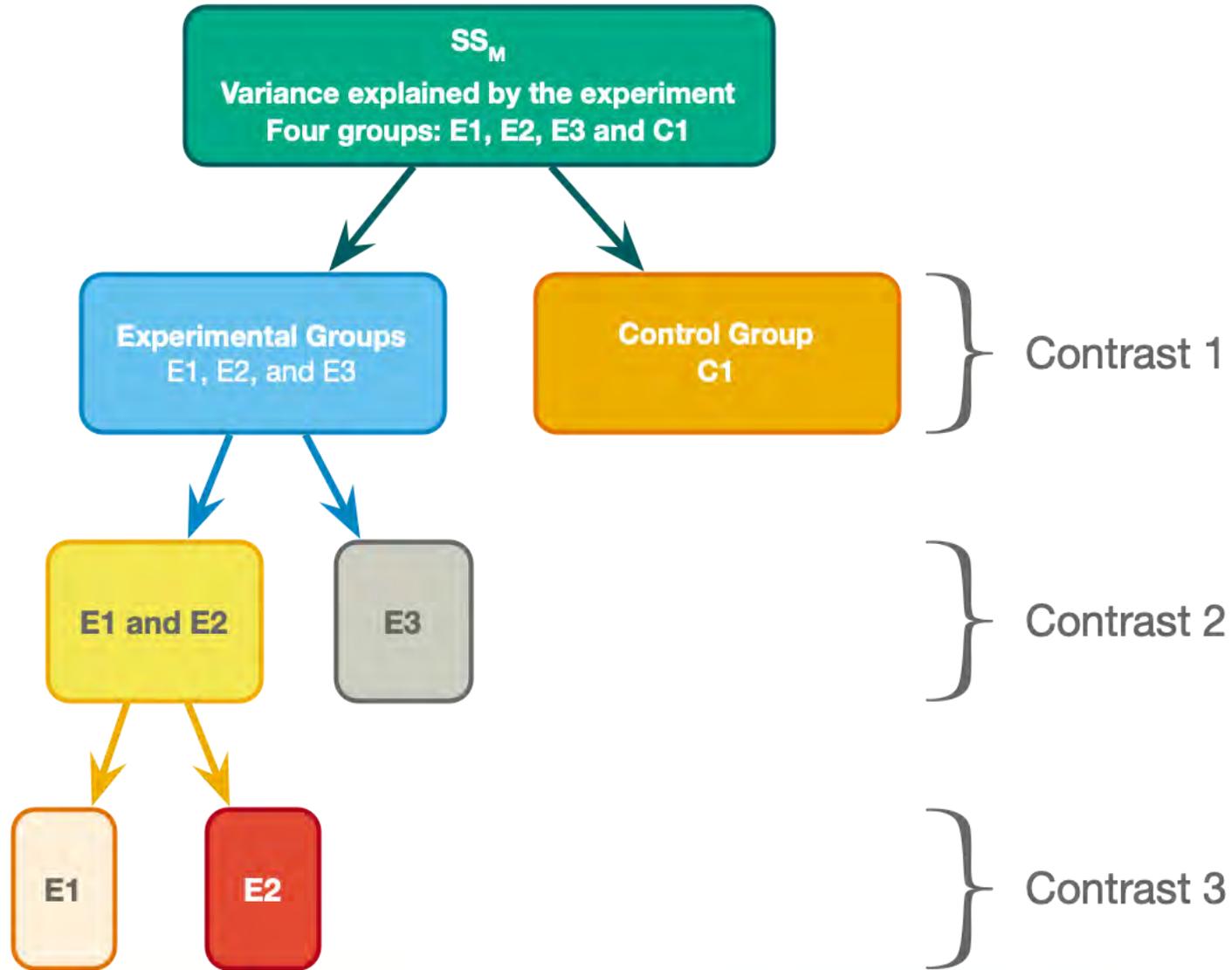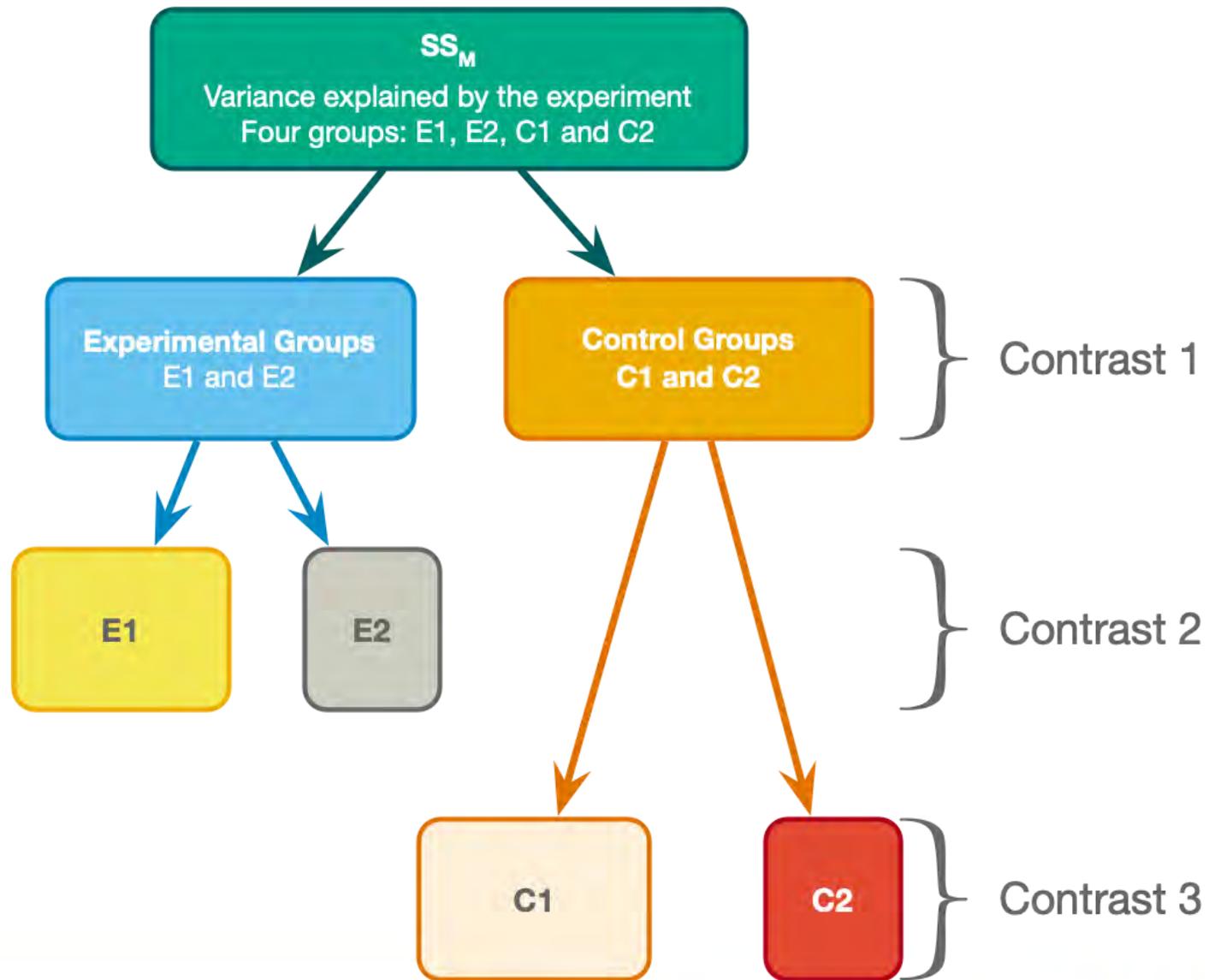
C2

Contrast 3

Coding planned contrasts

- Rule 1

  - Groups coded with positive weights compared to groups coded with negative weights

- Rule 2

  - The sum of weights for a comparison should be zero

- Rule 3

  - If a group is not involved in a comparison, assign it a weight of zero

- Rule 4

  - For a given contrast, the initial weight assigned to the group(s) in one chunk of variation should be equal to the number of groups in the opposite chunk of variation

- Rule 5

  - To get the final weight, divide the initial weights by the number of groups with non-zero weights

Chunk 1 15 mins + 30 mins	vs.	Chunk 2 No puppies (control)	}	Contrast 1
Positive		Negative		Sign of weight
1		2		Magnitude
+1 +1		-2		Initial weight
+1/3 +1/3		-2/3		Final weight

Chunk 1 **30 mins**	vs.	**Chunk 2** **15 mins**	Contrast 2	No puppies (not in contrast)
Positive		Negative	Sign of weight	
1		1	Magnitude	0
+1		-1	Initial weight	0
+1/2		-1/2	Final weight	0

What the coding does

Dummy coding

Therapy group	Long (30 mins vs. no puppies)	Short 1 (15 mins vs. no puppies)
No Puppies	0	0
15 mins	0	1
30 mins	1	0

Contrast coding

Therapy group	Contrast 1 (Puppies vs. no puppies)	Contrast 2 (15 mins vs. 30 mins)
No Puppies	-2/3	0
15 mins	1/3	-1/2
30 mins	1/3	1/2

The 'dummy' model

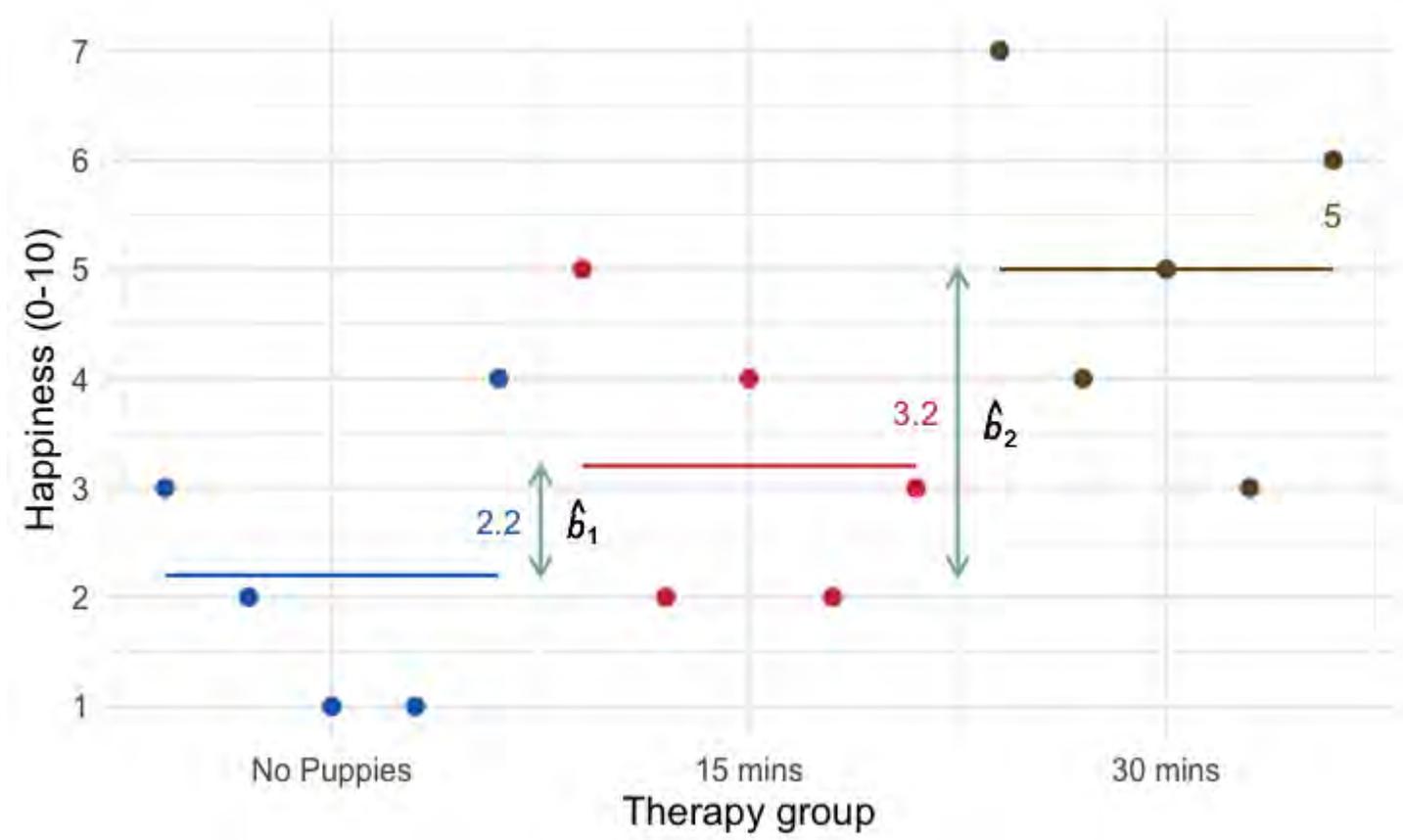$$\hat{\text{Happiness}}_i = \hat{b}_0 + \hat{b}_1 \text{Long}_i + \hat{b}_2 \text{Short}_i + e_i$$

$$\hat{\text{Happiness}}_i = \hat{b}_0 + \hat{b}_1 30 \text{ vs. control}_i + \hat{b}_2 15 \text{ vs.control}_i + e_i$$

The 'contrast' model

$$\hat{\text{Happiness}}_i = \hat{b}_0 + \hat{b}_1 \text{Contrast } 1_i + \hat{b}_2 \text{Contrast } 2_i + e_i$$

$$\hat{\text{Happiness}}_i = \hat{b}_0 + \hat{b}_1 \text{Therapy vs. control}_i + \hat{b}_2 15 \text{ vs. } 30 \text{ mins}_i + e_i$$

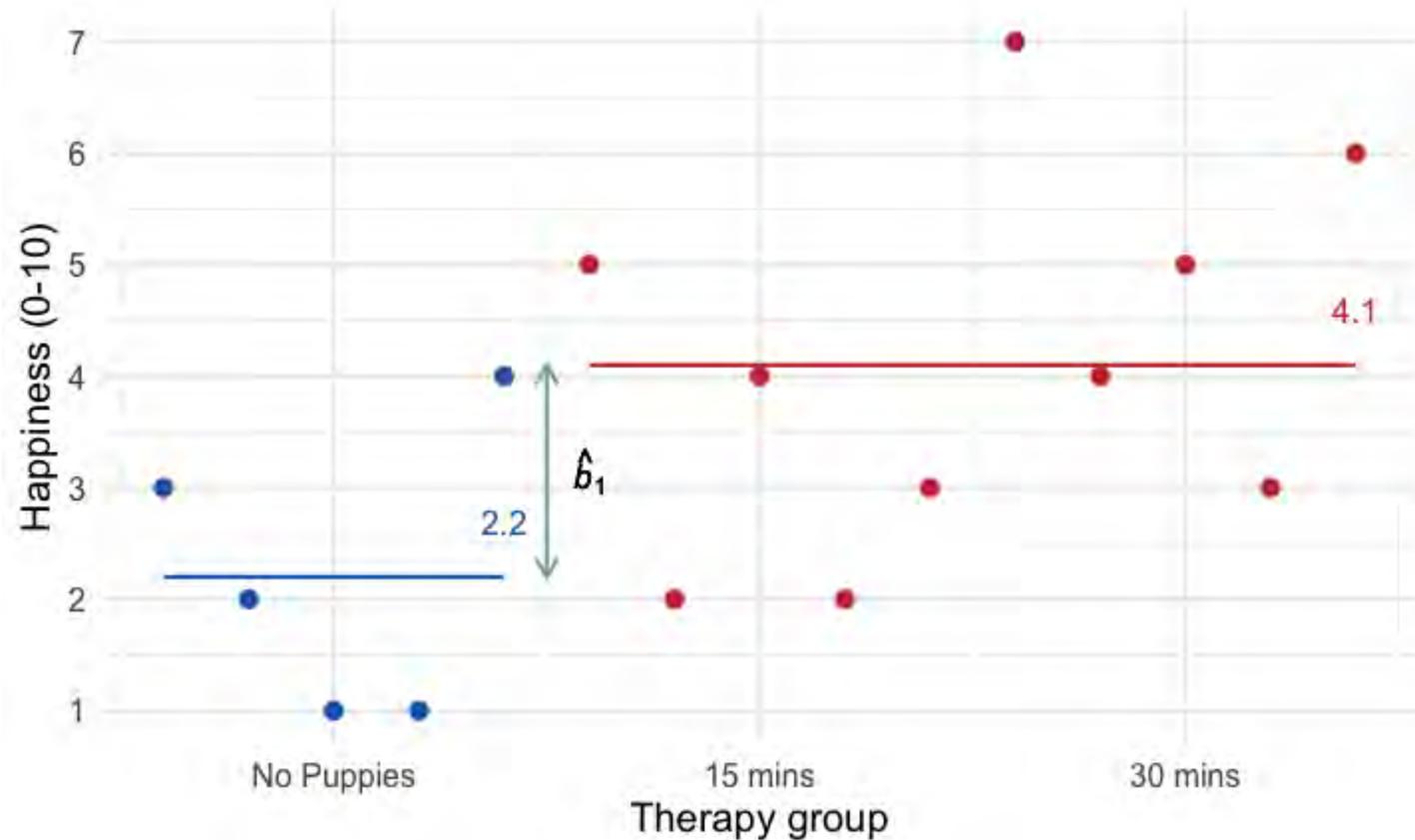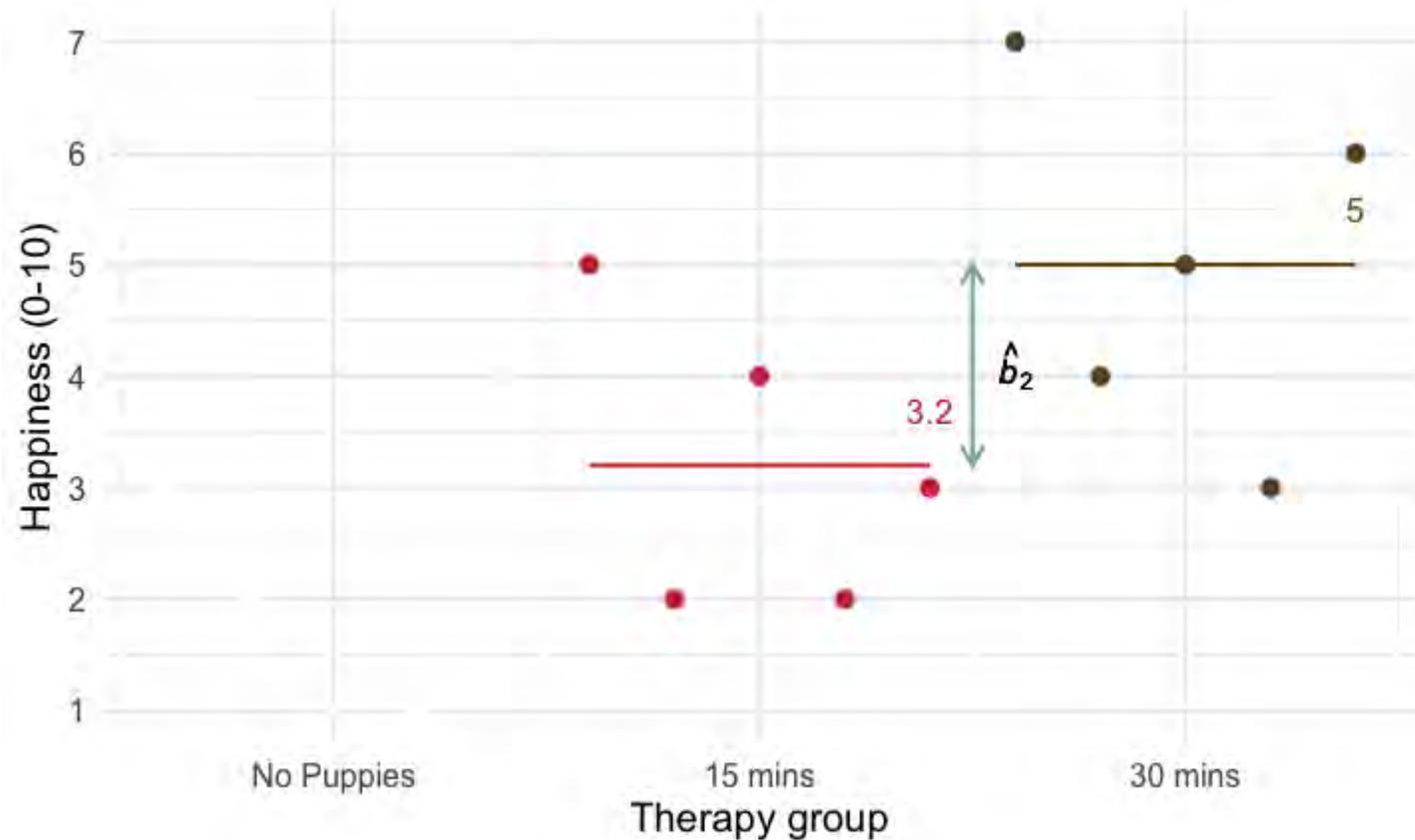# The 'dummy' model

The contrast model: Contrast 1

$$\hat{b}_1 = 4.1 - 2.2 = 1.9$$

The contrast model: Contrast 2

$$\hat{b}_2 = 5 - 3.2 = 1.8$$

Parameter estimates

```
puppy_vs_none <- c(-2/3, 1/3, 1/3)
short_vs_long <- c(0, -1/2, 1/2)

contrasts(puppy_tib$dose) <- cbind(puppy_vs_none, short_vs_long)

puppy_lm <- lm(happiness ~ dose, data = puppy_tib)
broom::tidy(puppy_lm, conf.int = TRUE)
```

term	estimate	std.error	statistic	p.value	conf.low	conf.high
(Intercept)	3.47	0.36	9.57	0.00	2.68	4.26
dosepuppy_vs_none	1.90	0.77	2.47	0.03	0.23	3.57
doseshort_vs_long	1.80	0.89	2.03	0.07	-0.13	3.73

Post hoc tests

- In the absence of specific hypotheses
    - Compare all pairs of means to see where the specific differences lie

Post hoc tests

- In the absence of specific hypotheses

  - Compare all pairs of means to see where the specific differences lie

- Problem

  - Inflates the Type I error rate

$$\text{Familywise error} = 1 - 0.95^{n}$$

- Solution
  - Adjust the alpha (or test statistic) to be more conservative

$$\text{Bonferroni } \alpha = \frac{\alpha}{\text{number of tests}}$$

Post hoc tests

```
modelbased::estimate_contrasts(puppy_lm, adjust = "bonferroni")
```

Level1	Level2	Difference	SE	CI_low	CI_high	t	df	p	Std_Difference
15 mins	30 mins	-1.8	0.89	-4.27	0.67	-2.03	12	0.20	-1.02
No puppies	15 mins	-1.0	0.89	-3.47	1.47	-1.13	12	0.84	-0.57
No puppies	30 mins	-2.8	0.89	-5.27	-0.33	-3.16	12	0.02	-1.58

Trend analysis (Polynomial contrasts)

- Test for trends in the means
- Makes sense only for ordered groups

```
contrasts(puppy_tib$dose) <- contr.poly(3)
puppy_trend <- lm(happiness ~ dose, data = pup

broom::tidy(puppy_trend, conf.int = TRUE)
```

term	estimate	std.error	statistic	p.value	conf.low
(Intercept)	3.467	0.362	9.574	0.000	2.678
dose.L	1.980	0.627	3.157	0.008	0.613
dose.Q	0.327	0.627	0.521	0.612	-1.040
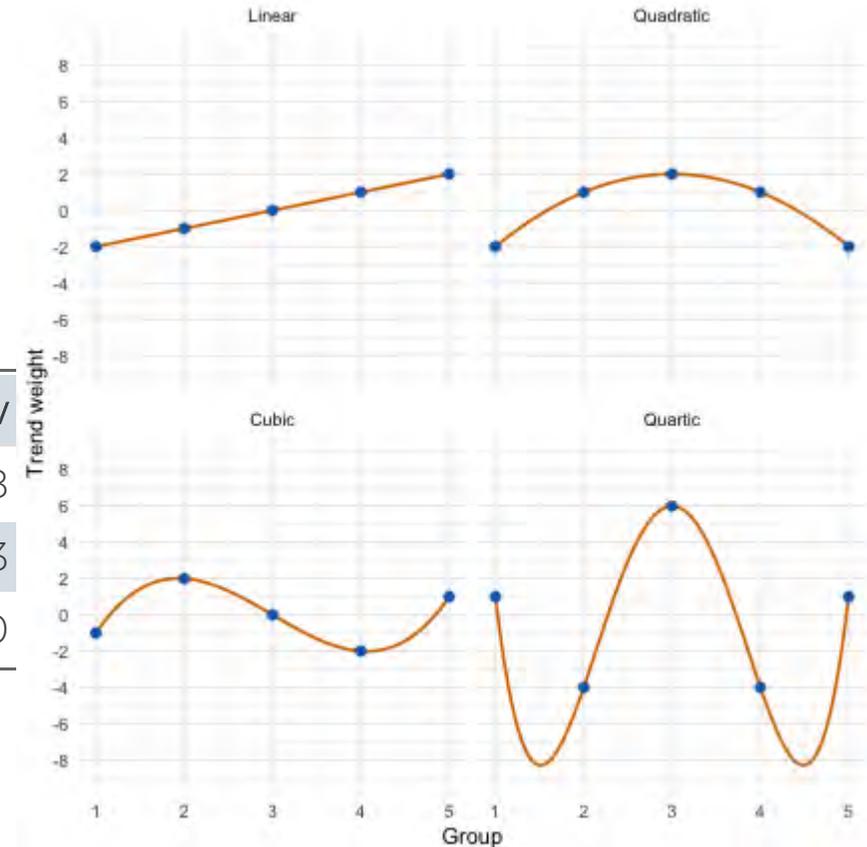
Summary

- Categorical predictors can be coded to test specific a priori hypotheses

- First devise contrasts to test your hypotheses

    - Independent
    - K-1 contrasts
    - Each compares 2 'chunks'

- Assign 'weights' to each group within each contrast

    - Assign 1 chunk positive values and the other negative
    - Assign an initial weight equal to the number of conditions in the opposite chunk
    - Divide the initial weight by the number of groups with non-zero weights

- *Post hoc* tests

    - Compare all pairs of group means but adjusting for multiple tests

- Polynomial contrasts (trend analysis)

    - Test for trends in the means of ordered categories