Stat 412/512

Hotelling's T^2

Mar 4 2015

Charlotte Wickham stat512.cwick.co.nz

Lab today

No new material.

Your TA will be there.

Ask questions about:

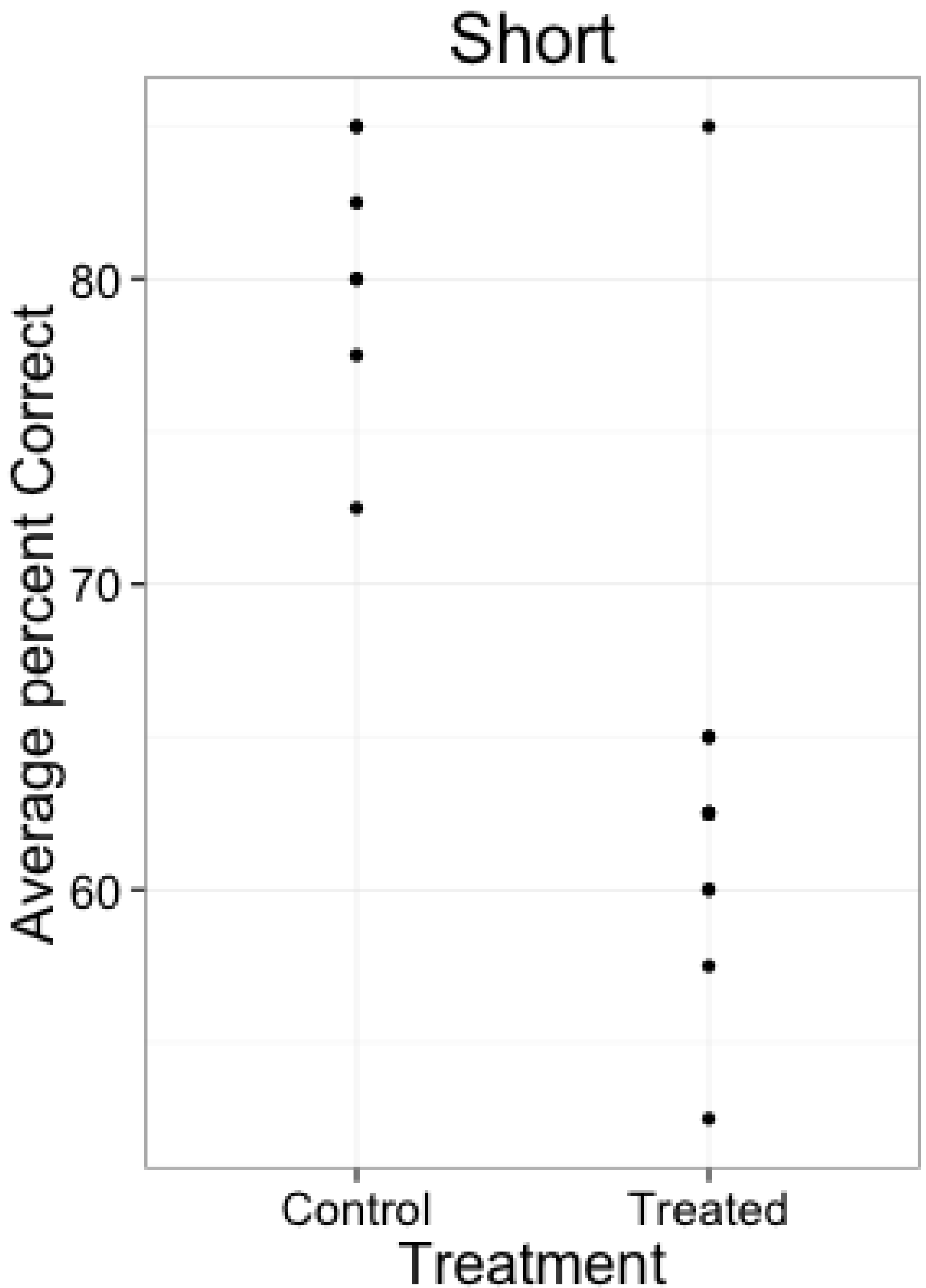
- R
- Course content
- Quiz questions
- Last year’s final

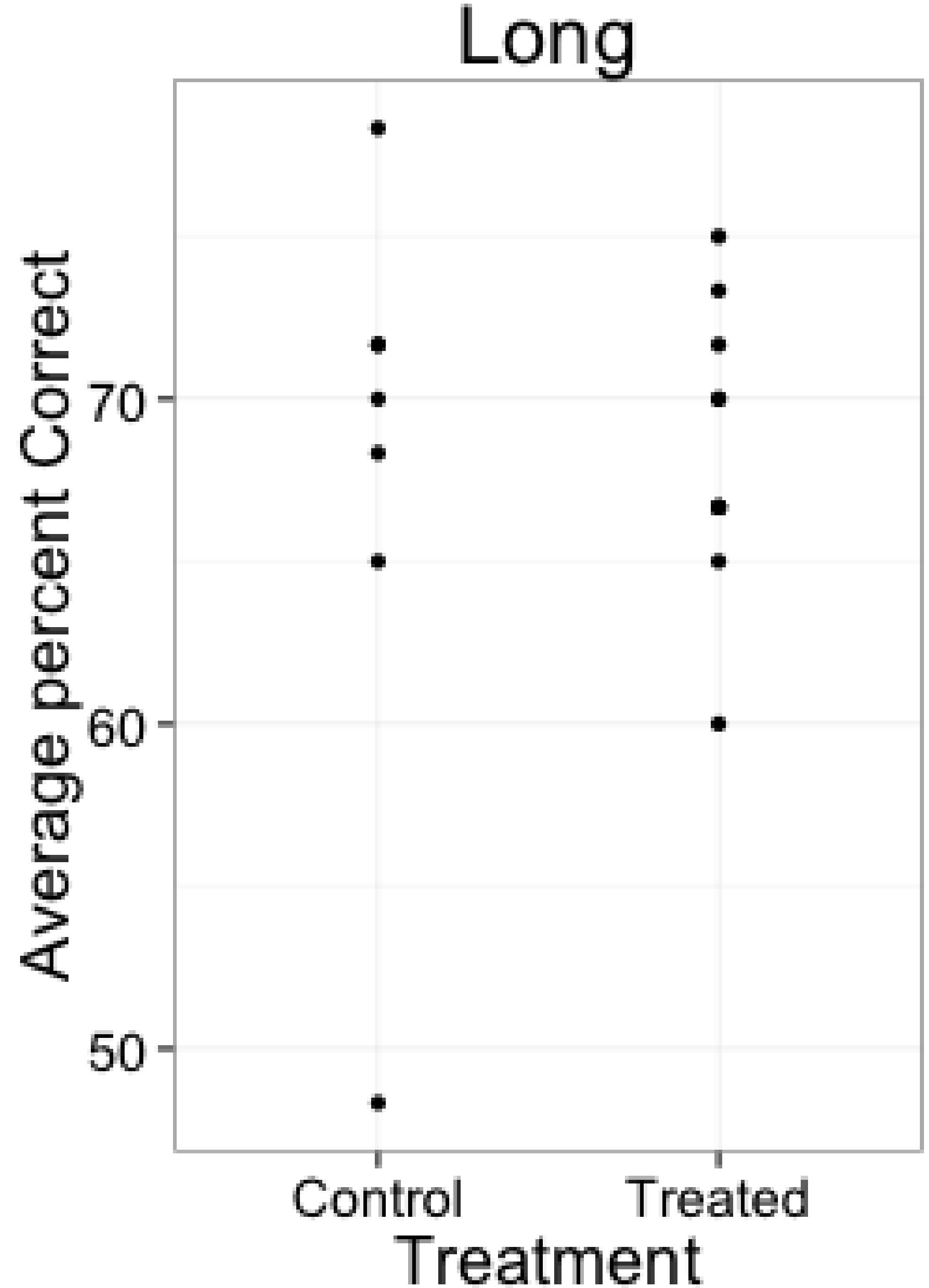
Big Idea

Hotelling's T^2 extends the "two-sample t-test" to multiple "two-sample t-test"s on different response variables.

i.e. the samples/groups/populations are the same in each but,

the response variable is different.





Do treated monkeys have similar short term memory to control monkeys?

Do treated monkey's have similar long term memory to control monkeys?

Hotelling’s T^2

An ellipse is hard to compute and hard to present.

Hotelling’s T^2 adjustment adjusts the univariate confidence intervals to conservatively approximate the ellipse.

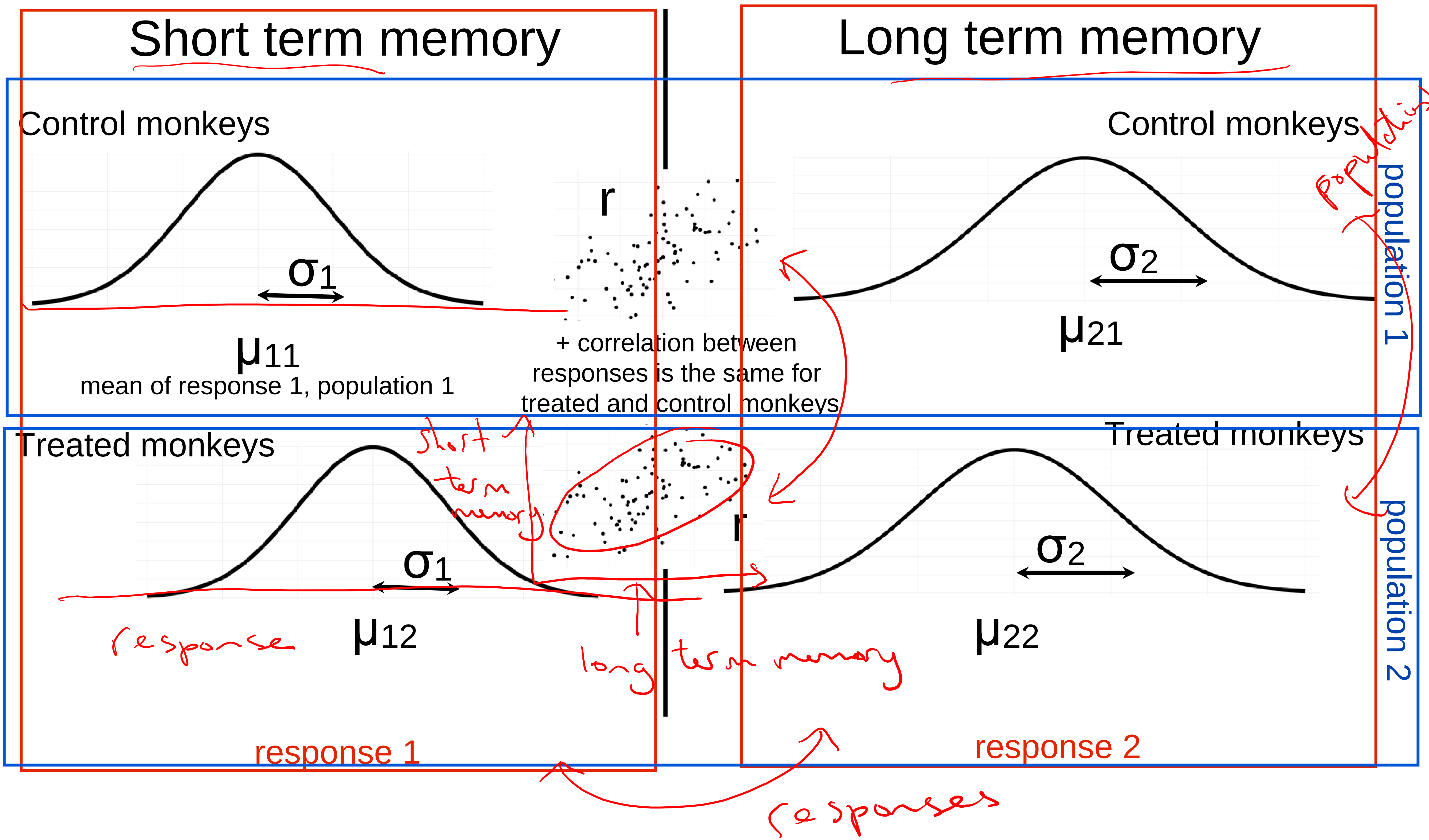
Hotelling’s T^2 statistic provides an joint test for both parameters at once.

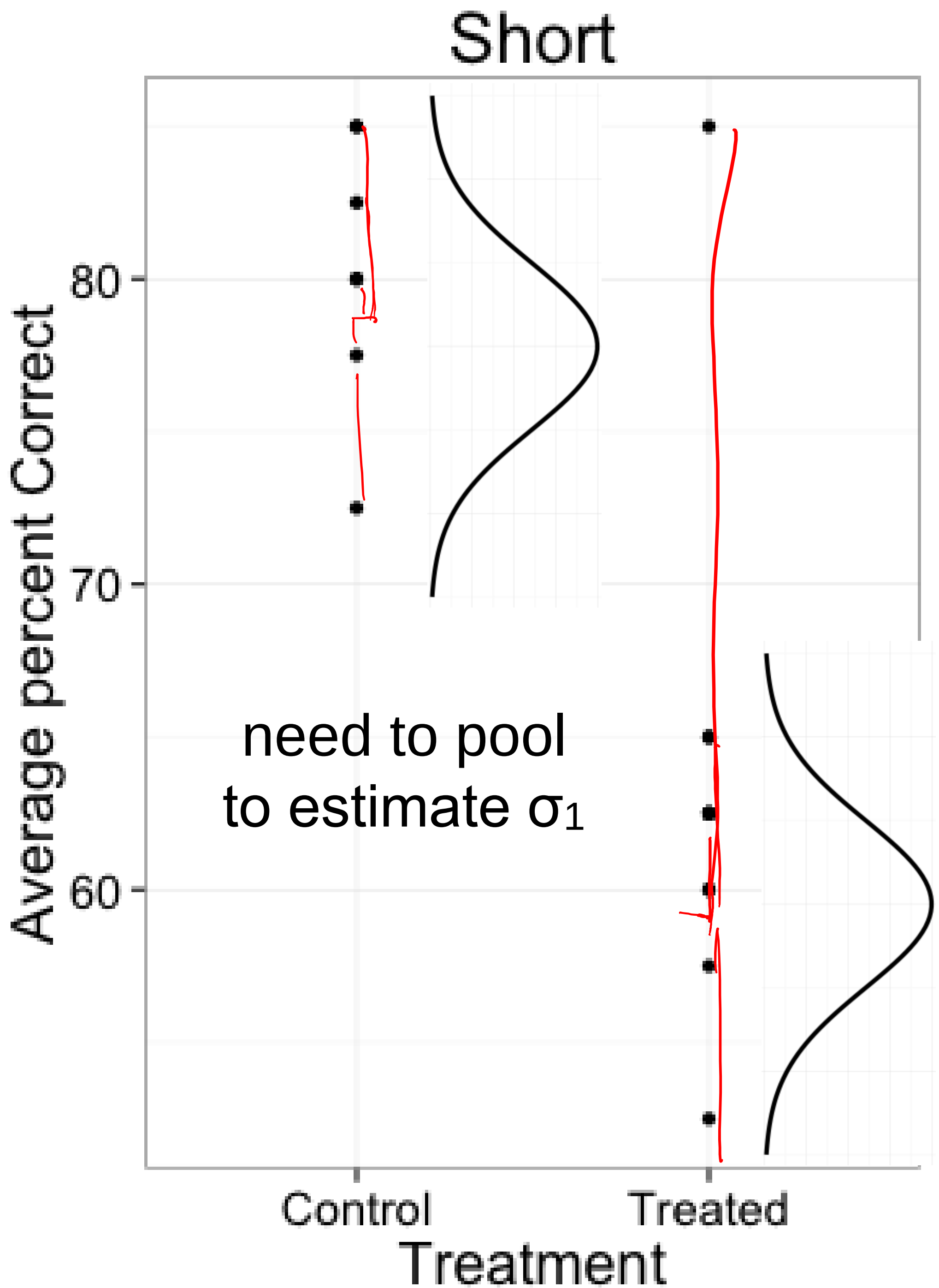
An ideal model for the monkeys

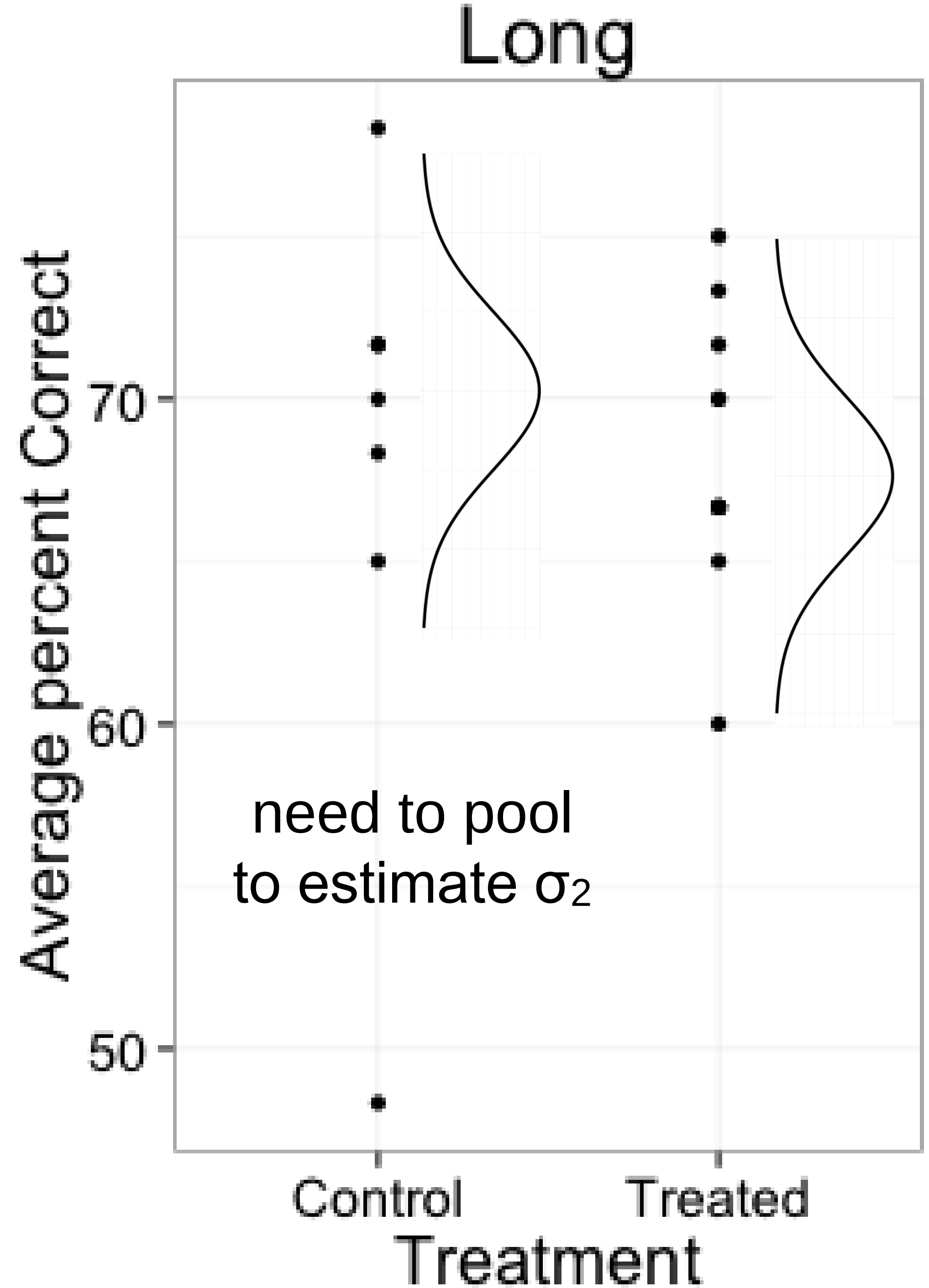
1. The mean of response i, in population j is, μ_{ij}
2. Both populations have the same population standard deviations of response, $\sigma_{i.}$
3. Both populations have the same population correlation between responses (avg long term and avg short term memory).
4. Responses are normally distributed about their means.
5. Subjects are randomly sampled for each population.
6. Two groups are sampled independently.

Two “two sample t-test” assumptions +
same correlation between responses in both populations

An ideal model for the monkeys







Need to pool across both populations to estimate σ_1 and σ_2

Need to pool across both populations to estimate r

Display 16.5

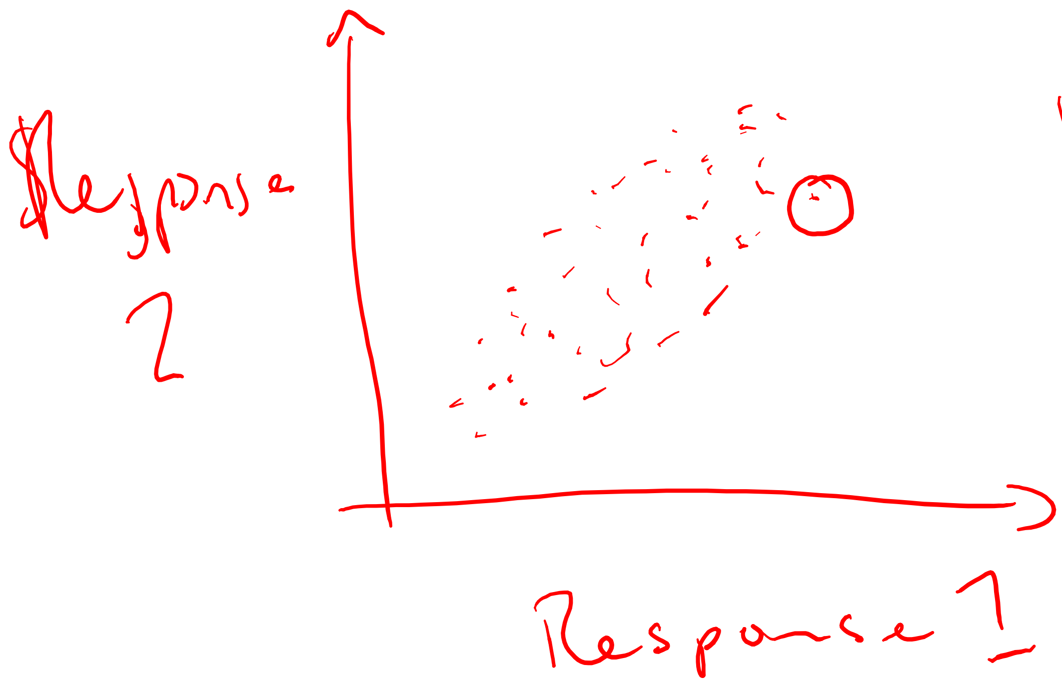
Pooled estimates of variance and covariance for the memory example

	Control	Treated	Pooled	
Sample Size (n)	7	11	18	$= 7 + 11$
Degrees of Freedom	6	10	16	$= 6 + 10$
Sample variances				
short-term:	19.64	65.11	S^2_1 48.06	$= \dfrac{\{6(19.64)+10(65.11)\}}{16}$
long-term:	88.76	17.78	S^2_2 44.39	=
Sample covariance	-15.67	3.75	-3.53 C	= similar pooling

$$\text{Pooled sample correlation} = C/(S_1 S_2)$$
$$= -3.53/(\text{sqrt}(48.06) * \text{sqrt}(44.39))$$
$$= -0.076$$

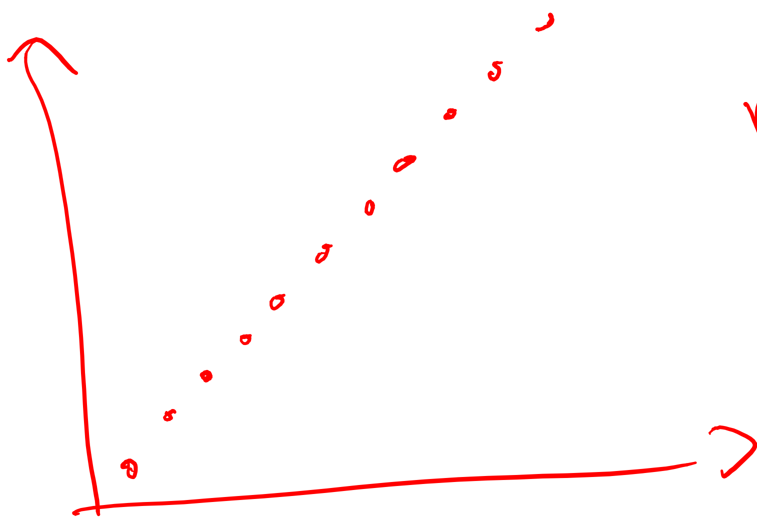
between -1 and 1

0 is no correlation

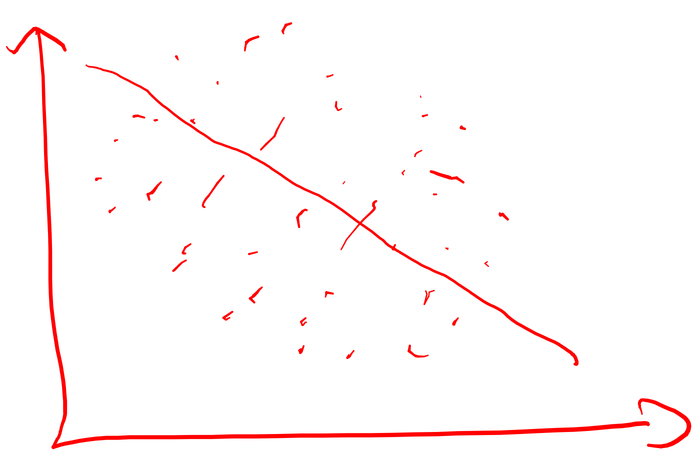


positively correlated

$\sim 0.6 - 0.8$



$r = 1$



our case

Hotelling's T^2 statistic

Null: null is true in both tests

$$T^2 = \boxed{\frac{t_1^2 + t_2^2 - 2Rt_1t_2}{1 - R^2}}$$

pooled correlation between responses

where t_1 and t_2 are individual t-statistics for each response, R is the correlation between responses.

Under the ideal model,

$$F = \frac{n_1 + n_2 - 3}{2(n_1 + n_2 - 2)} T^2$$

has an F distribution with

2 and n_1+n_2 - 3 degrees of freedom

We have convincing evidence against

the difference in mean long term memory between treated and control monkeys **and**

the difference in mean short term memory between treated and control monkeys

both being zero.

(Alternate wording) We have convincing evidence at least one of

the difference in mean long term memory between treated and control monkeys **and** $= 0$

the difference in mean short term memory between treated and control monkeys

is not zero. $\neq 0$

Display 16.6 p. 476

Hotelling's T^2 calculations for the memory example

	short-term	long-term
1 (Control-Treatment) Average Differences	$17.18 = (80.36 - 63.18)$	$-0.71 = (67.62 - 68.33)$
2 Standard Errors	$3.35 = 6.93\sqrt{\frac{1}{7} + \frac{1}{11}}$	$3.22 = 6.66\sqrt{\frac{1}{7} + \frac{1}{11}}$
3 t-statistics for individual hypotheses of no difference	$5.12 = \frac{17.18}{3.35}$	$-0.22 = \frac{-0.71}{3.22}$

4 Hotelling's T^2

$$26.28 = \frac{(5.12)^2 + (-0.22)^2 - 2(5.12)(-0.22)(-.0765)}{1 - (-.0765)^2}$$

5 F-statistic

$$12.32 = \frac{(15)(26.28)}{(2)\ (16)}, \qquad df = 2, 15$$

1 - pf(12.32, 2, 15) = 0.00068

Joint confidence intervals

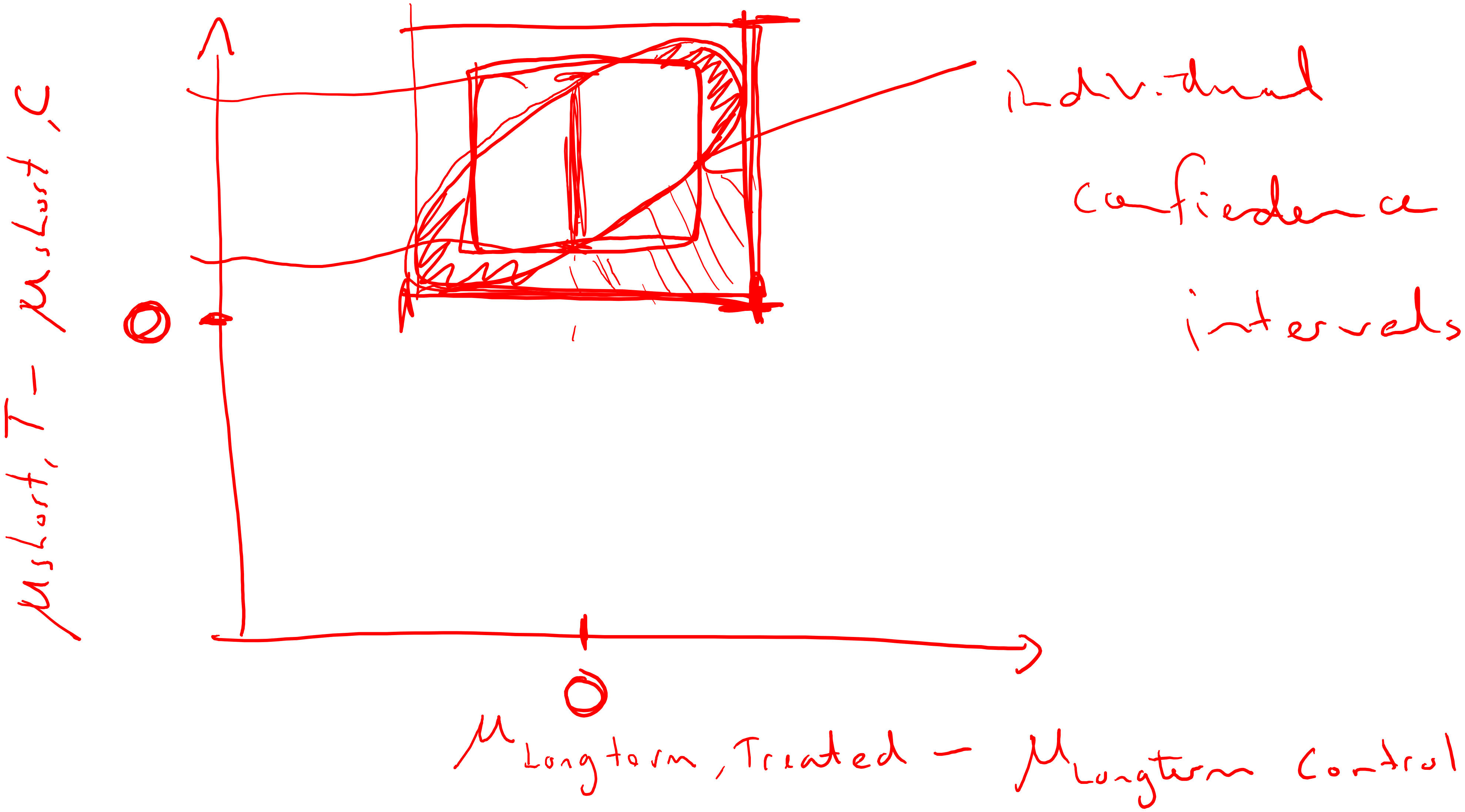
An ellipse is the best description of our joint confidence, but

Hotelling's adjusted confidence intervals guarantee at least 95% confidence

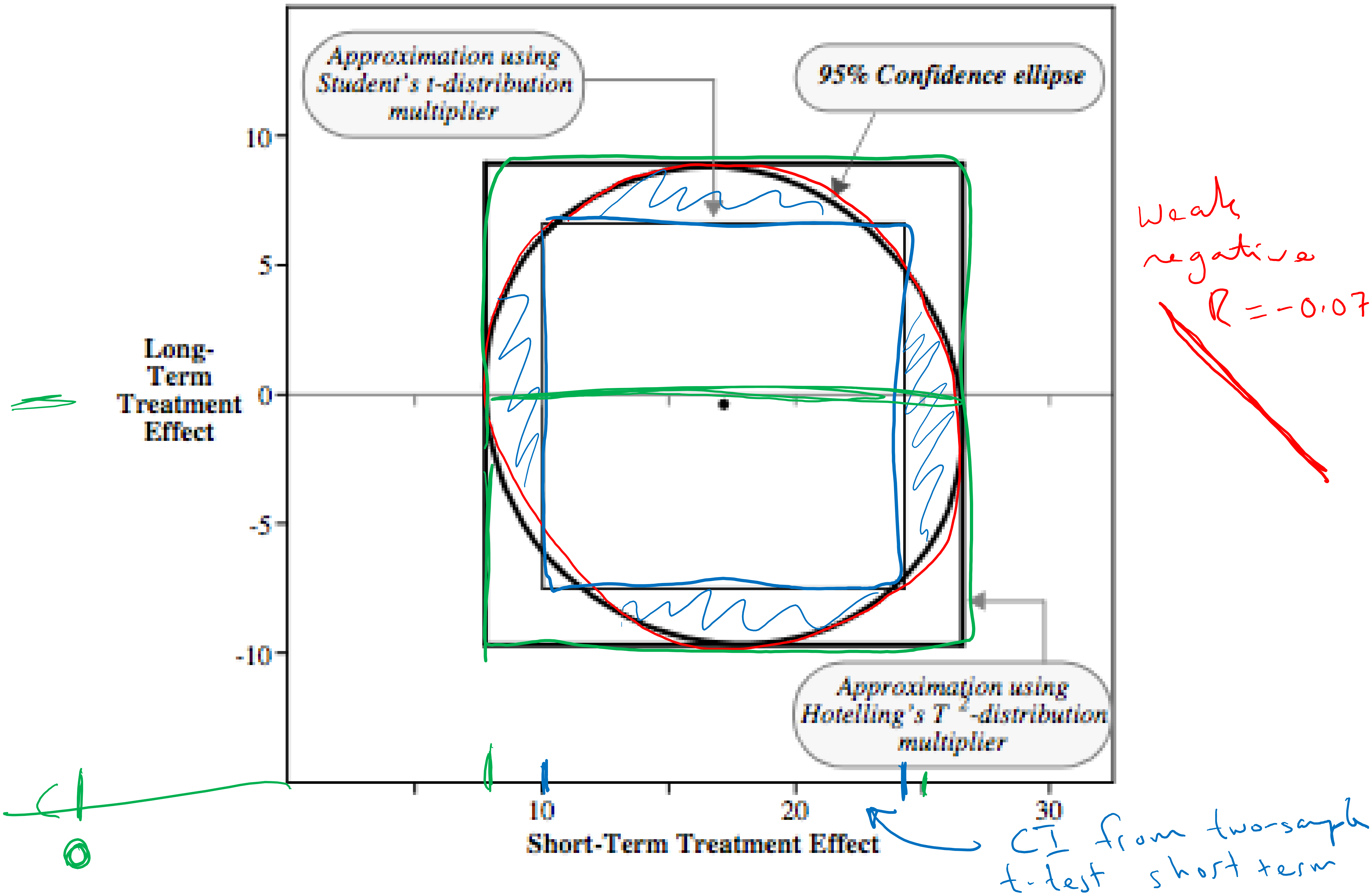
estimate ± multiplier × SE

multiplier= $$\sqrt{\frac{2(n_1+n_2-2}{n_1+n_2-3}F_{2,n_1+n_2-3}(1-\alpha)}$$

2.803 for the monkey study
(compare to 2.120 for the usual t CIs)


individual confidence intervals
0
0
μ Longterm, Treated − μ Longterm Control
μ short, T − μ short, C

95% confidence ellipse for differences in long-term and short-term memory means, and approximate confidence rectangles, constructed with T^2 and t multipliers



Weak negative $R = -0.07$

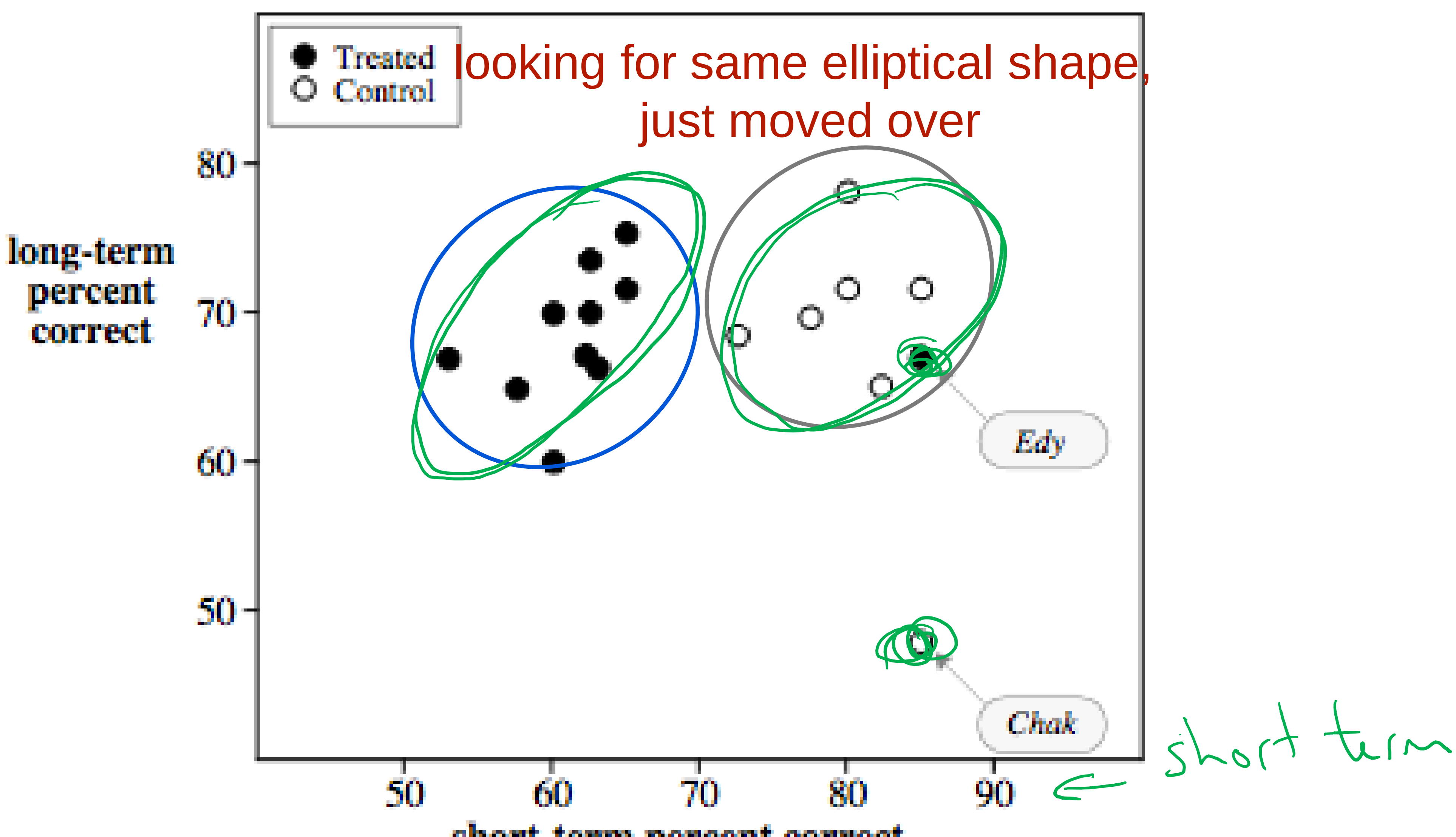
Checking the ideal model

Use usual tools for checking equal SDs and normality.

Also check ...

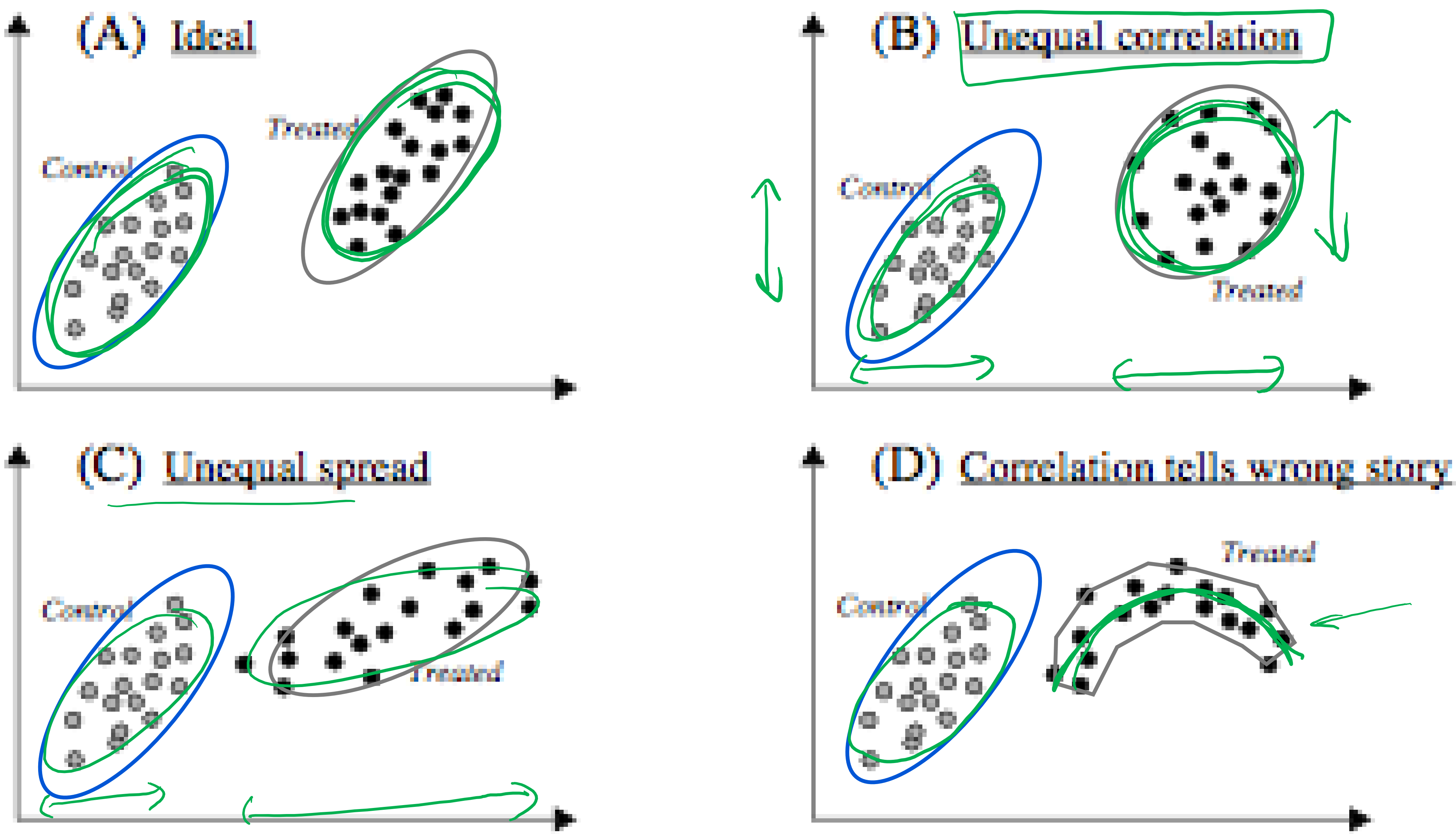
Display 16.7 **p. 477**

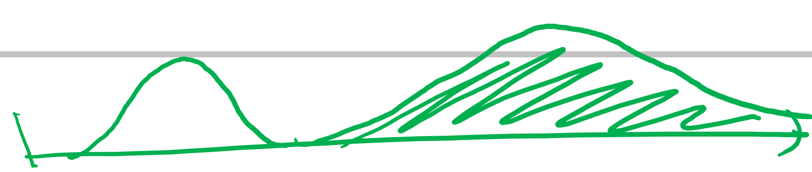
Scatter plot of short- and long-term memory scores, by treatment group



Display 16.13

Diagnostic patterns for correlated responses in the two-sample problem. (Hypothetical data)





Strategies

1. A single univariate analysis on a summary of the multivariate response.
2. Separate univariate analyses on several summaries.
3. Multivariate analysis on several summaries. Hotelling's T^2
4. Treat subject as a factor.