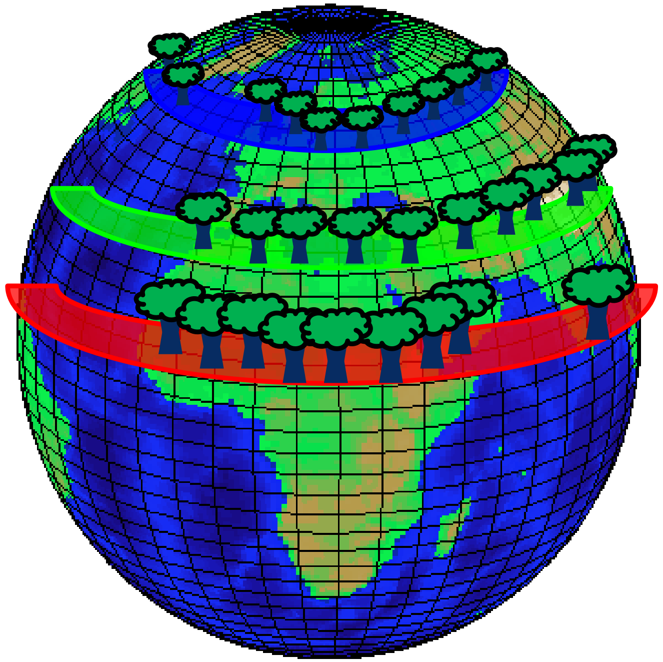
UNDERSTANDING STATISTICS & EXPERIMENTAL DESIGN

Content

Statistics & Data Reduction

Tree Heights

- Suppose we want to compare tree heights over the three regions shown on the right in the red, green, and blue bands.
- Are the tree heights all the same, or does at least one region contain trees whose heights differ from those in the other regions?



How to avoid false discovery

During m independent statistic test with α significance level, the probability of at least one false discovery should be

$$1-(1-\alpha)^m < 0.05$$

$$\alpha = 1-(1-0.05)^{1/m} \approx \frac{0.05}{m}$$

Bonferroni correction: during m independent statistic tests only those results are significant, for which

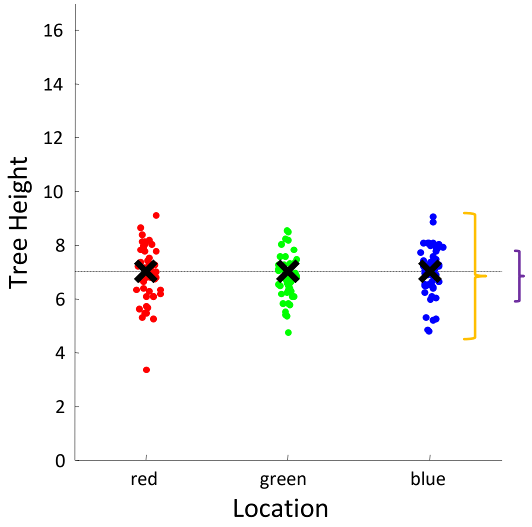
$$p < \frac{0.05}{m}$$

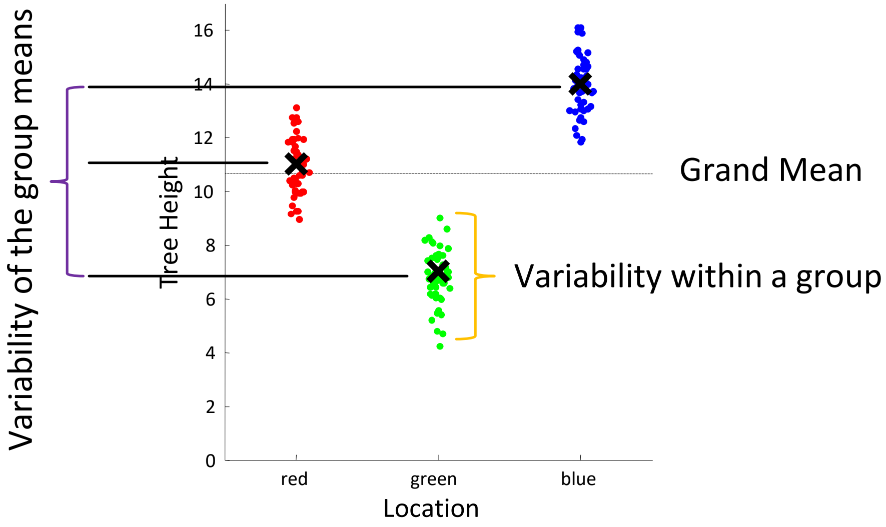
ANOVA

$$F = \frac{\text{Variability between treatments}}{\text{Variability within treaments}}$$

$$F = \frac{\text{Small Number}}{\text{Large Number}} = \text{Small Number}$$

$$F = \frac{\text{Large Number}}{\text{Small Number}} = \text{Large Number}$$

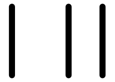


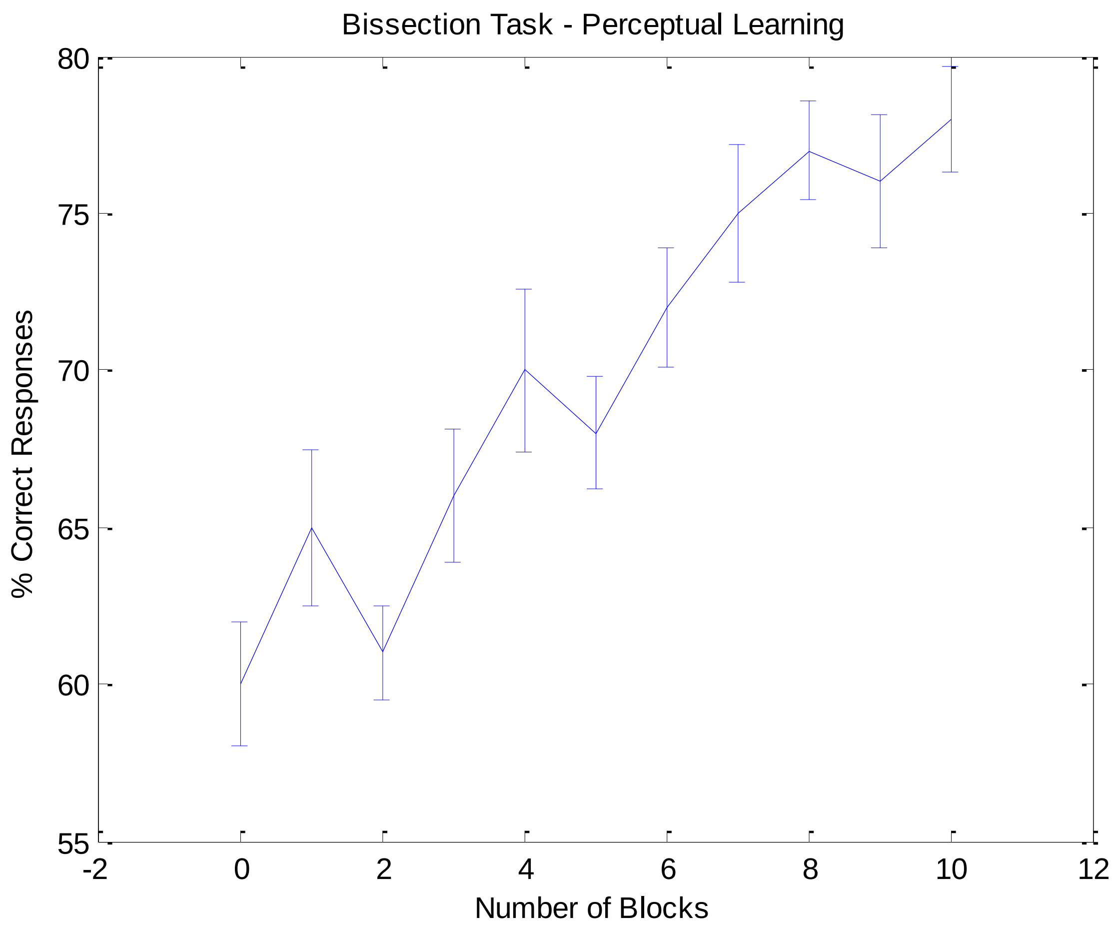


Reduce data I
Model Fitting

Why is ANOVA a bad idea?

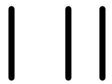
Bisection Task

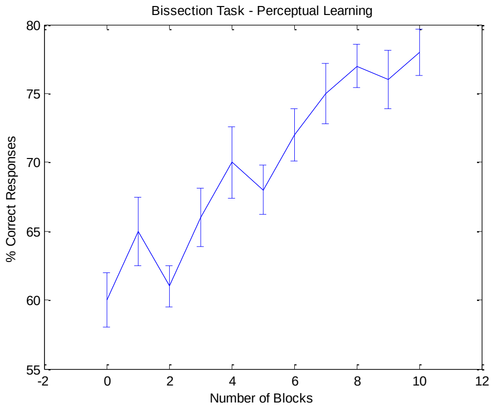




Statistics after data reduction

Bisection Task

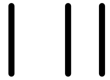


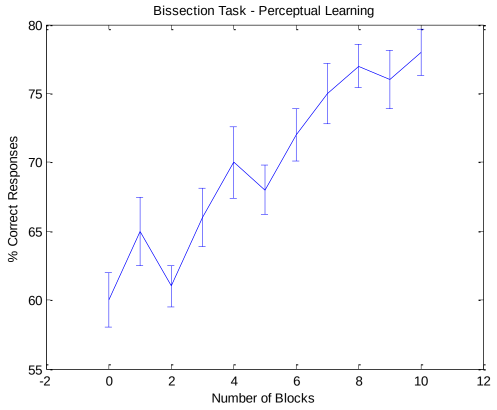


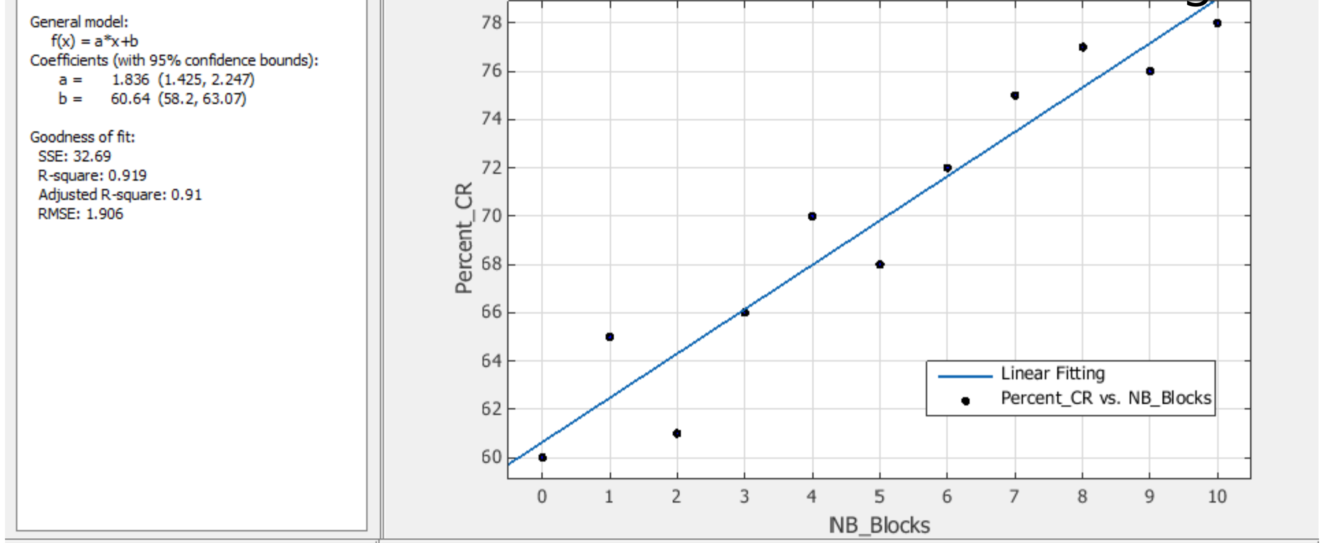
Statistics after data reduction

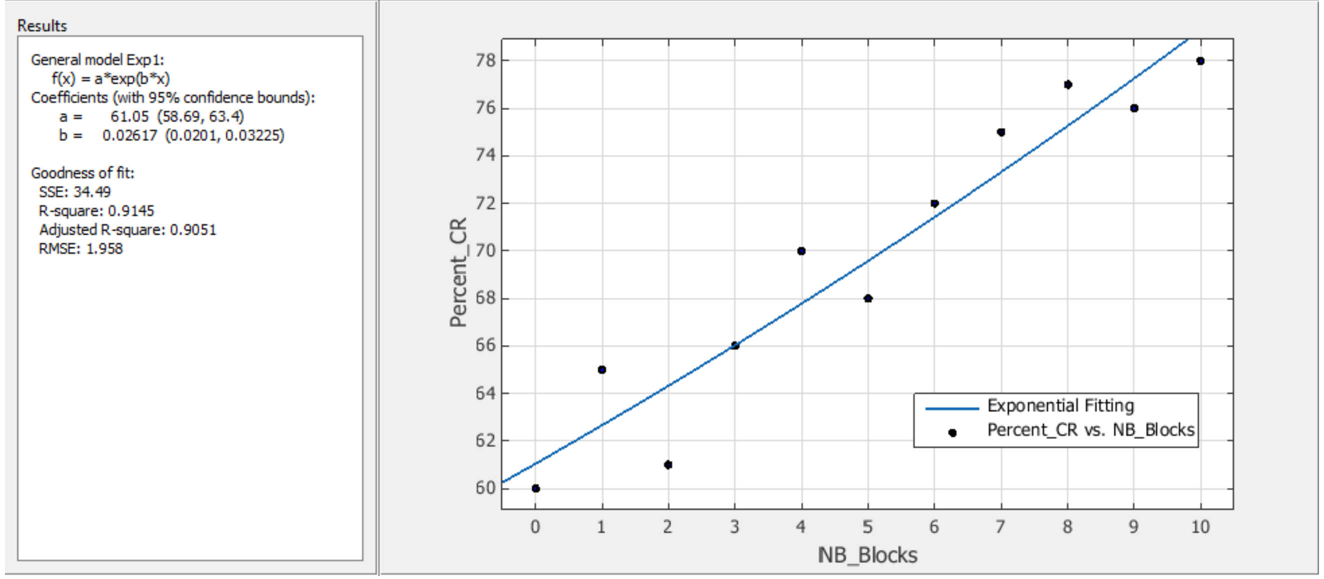
Linear Fitting: R-square= 0.91

Bisection Task



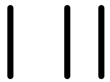


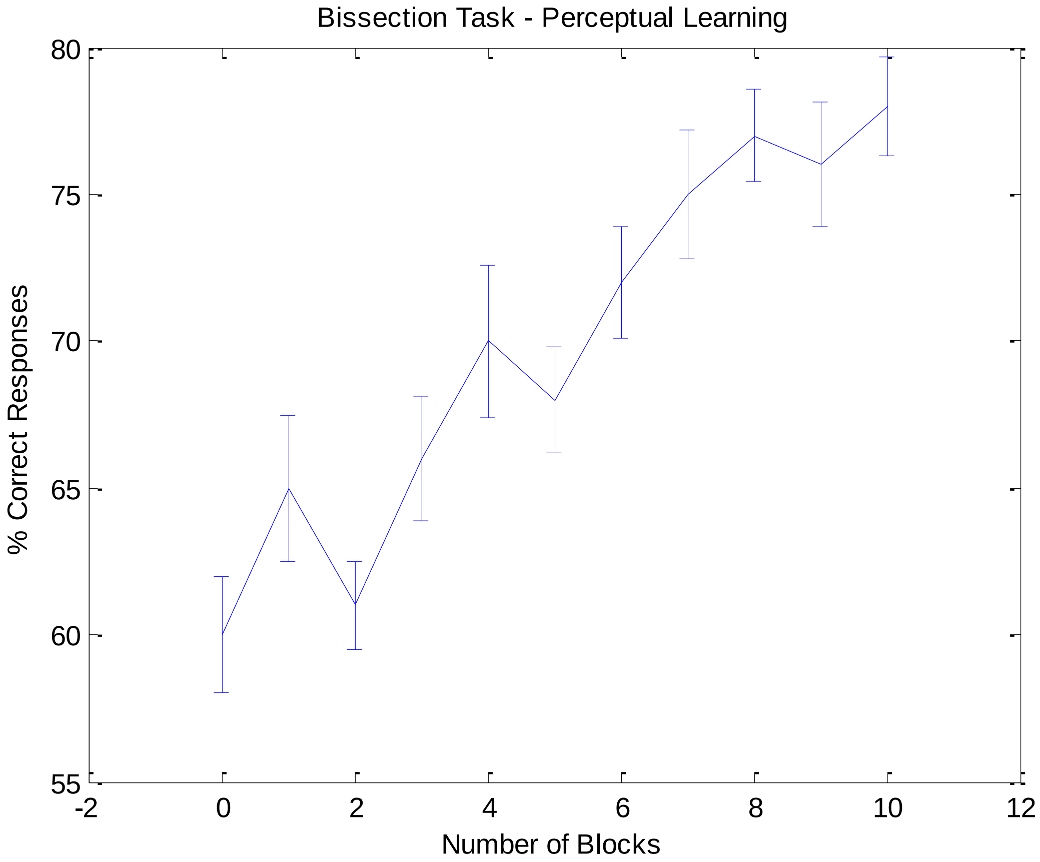




Scales are an issue

Bisection Task





Two sample *t*-test

$$t_n = \frac{|\bar{x}_{North} - \bar{x}_{South}|}{\hat{\sigma}/\sqrt{n}}$$

One sample *t*-test

$$t = \frac{|\bar{x} - \mu|}{\hat{\sigma}/\sqrt{n}}$$

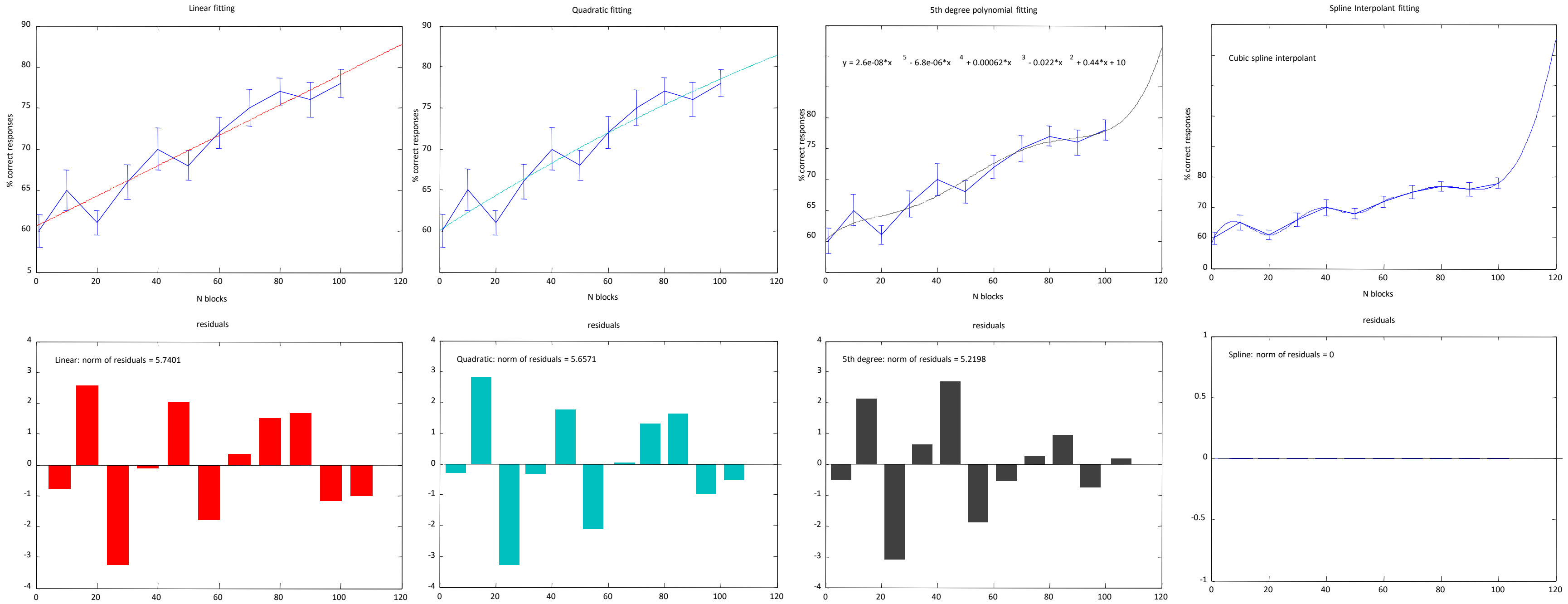
Repeated measures *t*-test

The two sample *t*-test is in fact a one-way t-test with $\mu = 0$

Statistics by data reduction

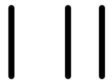
$$y = 0.18*x + 11$$

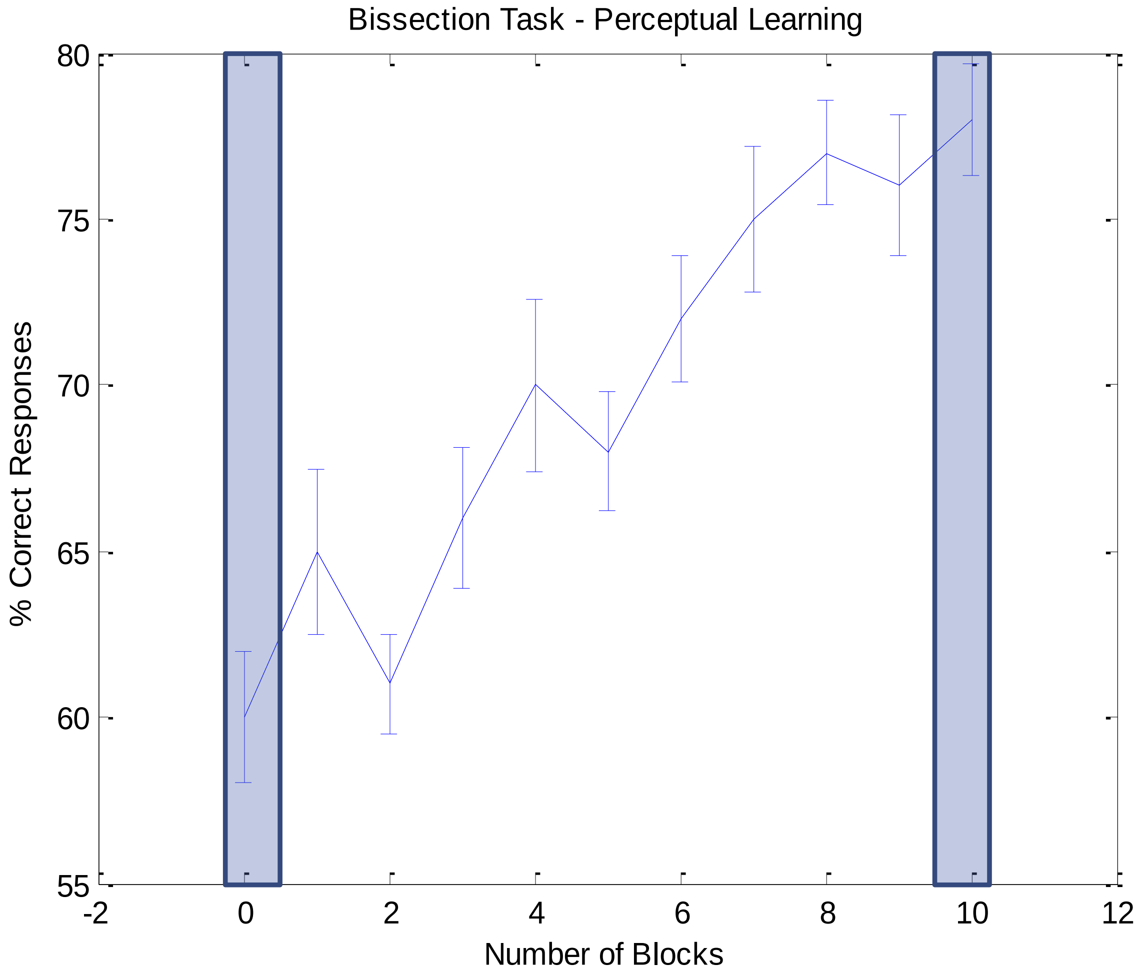
$$y = -0.00034*x^2 + 0.22*x + 10$$



Pre-post testing

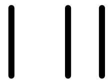
Bisection Task

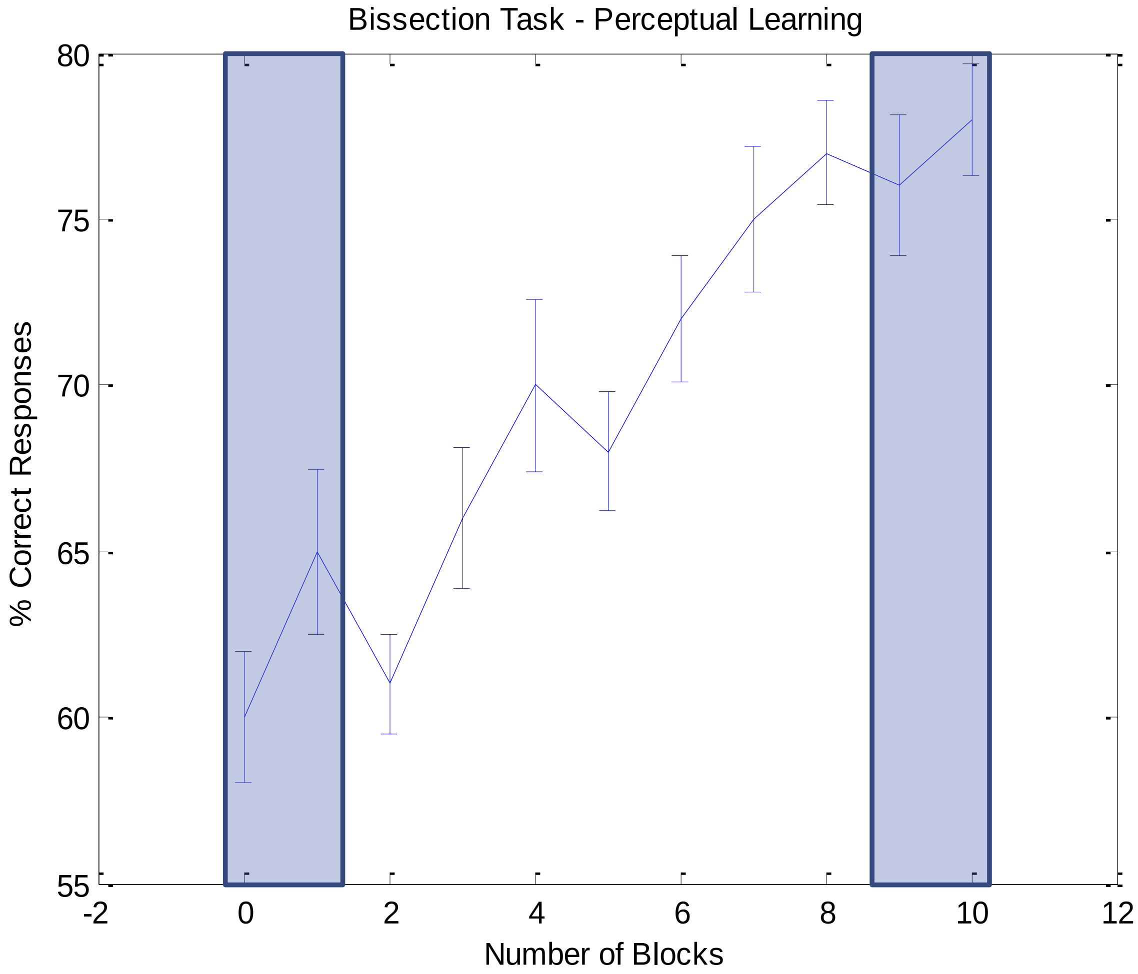




Pre-post testing

Bisection Task





Two sample *t*-test

$$t_n = \frac{|\bar{x}_{North} - \bar{x}_{South}|}{\hat{\sigma}/\sqrt{n}}$$

One sample *t*-test

$$t = \frac{|\bar{x} - \mu|}{\hat{\sigma}/\sqrt{n}}$$

Repeated measures *t*-test

The two sample *t*-test is in fact a one-way t-test with $\mu = 0$

Reduce data IIa: Method of constant stimuli (MCS)

A brain enhancing program proposes to increase not only cognition but also perception. We like to test this proposal using an orientation discrimination task before and after training.

A brain enhancing program proposes to increase not only cognition but also perception. We like to test this proposal using an orientation discrimination task before and after training.

Before training  

Brain Training

After training  

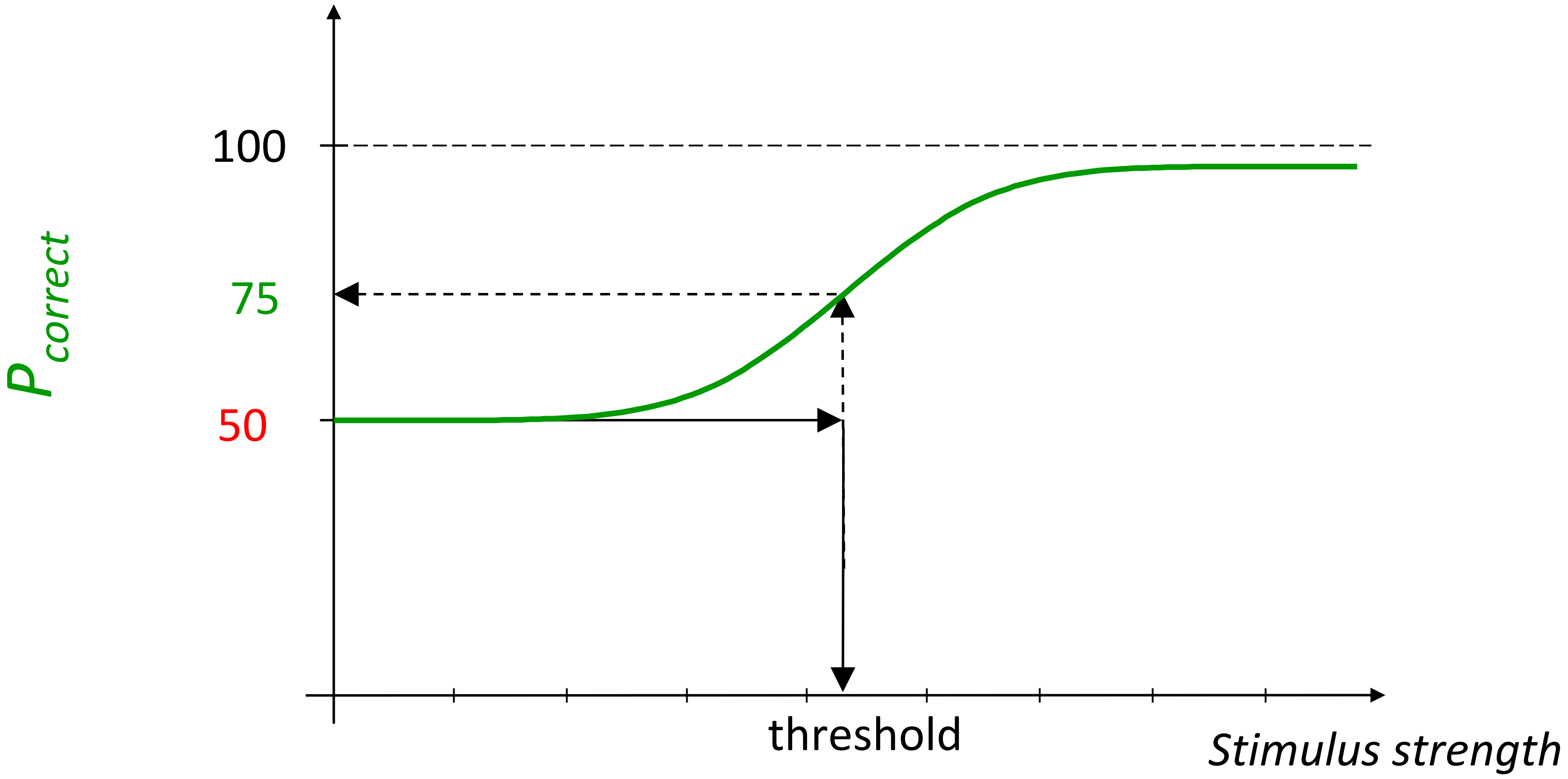
A brain enhancing program proposes to increase not only cognition but also perception. We like to test this proposal using an orientation discrimination task before and after training.

We need to think about floor and ceiling effects and therefore offer a range of stimuli:

Train: Brain enhancing

Compute ANOVA?

Method of constant stimuli (MCS)

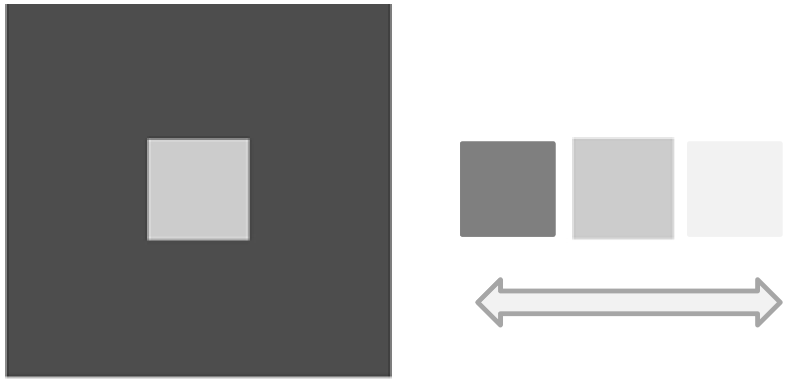


Reduce data IIc: Adjustment Tasks

Adjustment Task

Orientation adjustment

Color adjustment



Adapted from: http://www.frontiersin.org/files/Articles/9032/fpsyg-02-00119-HTML/image_m/fpsyg-02-00119-g001.jpg

Reduce data collection II

Adjustment Tasks

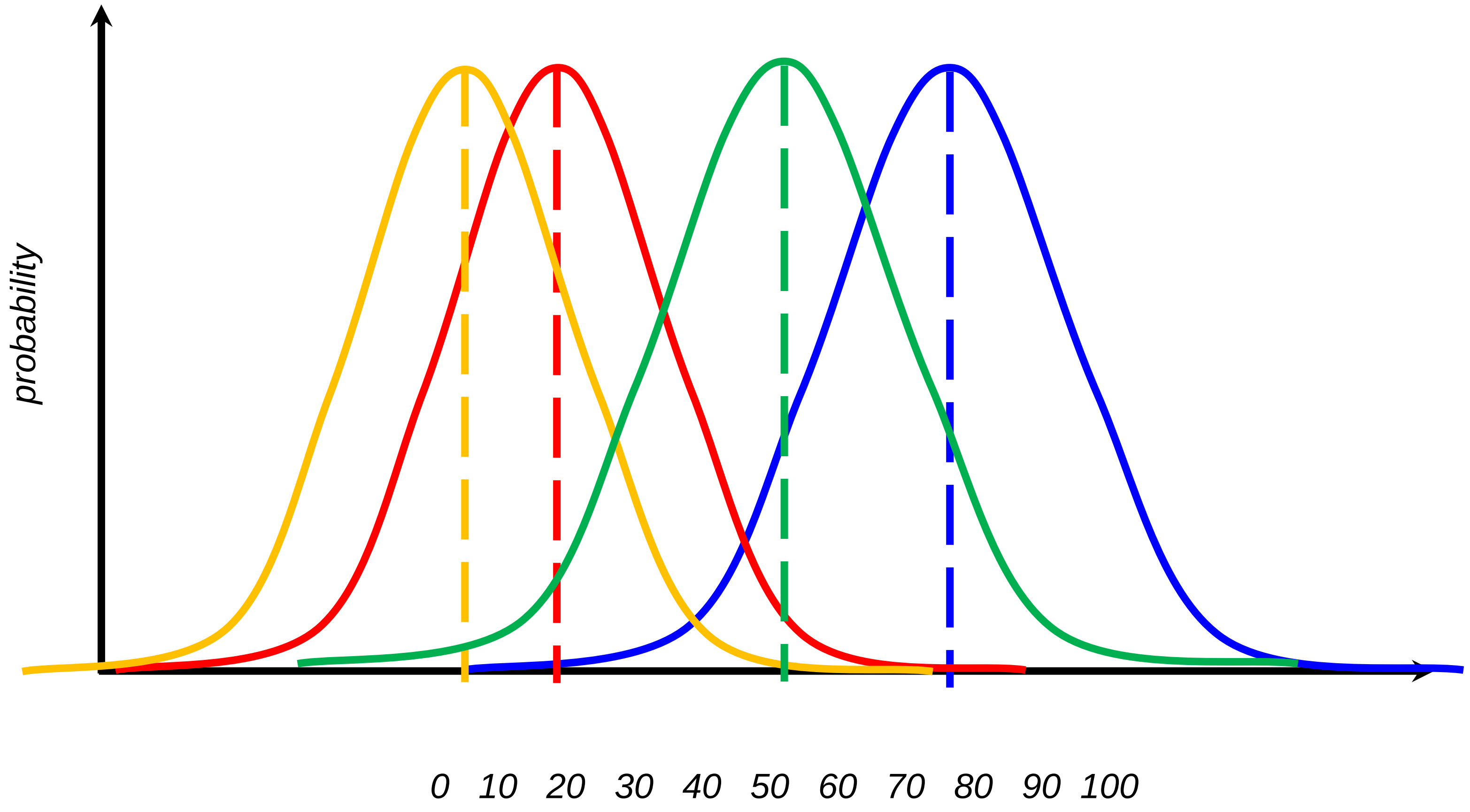
Main Problem: outlier detection

Why not identification task?

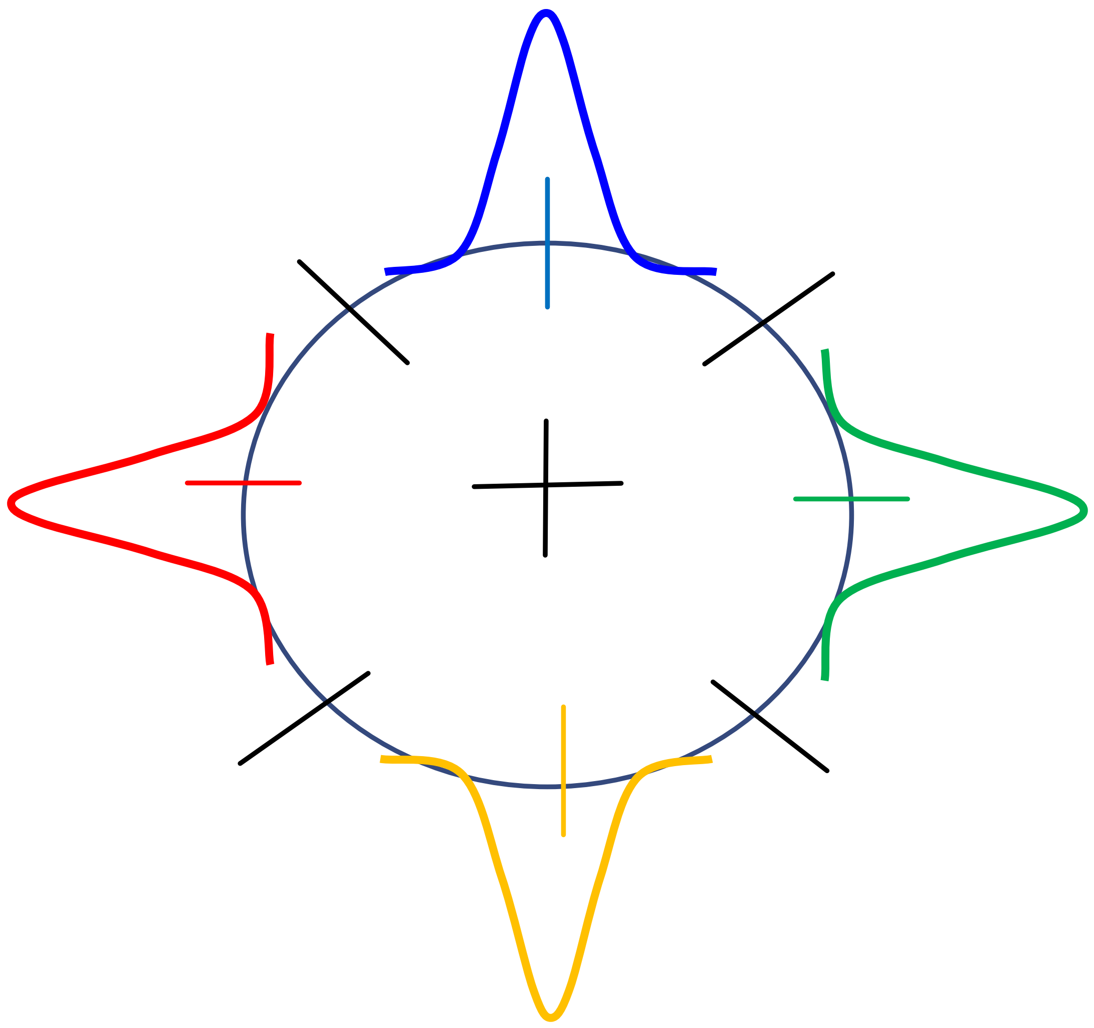
A brain enhancing program proposes to increase not only cognition but also perception. We like to test this proposal using an orientation discrimination task before and after training.

Training

Discriminability



Experiment with 4 stimulus and response alternatives



Reduce data collection III
Number of trials

Number of trials

Resources are always limited in science

For example, you can screen only

- 1000 cells with an anti-body
- Ask 1000 people in a poll
- Collect 5 stones from moon
- Check 10% of a product for quality control

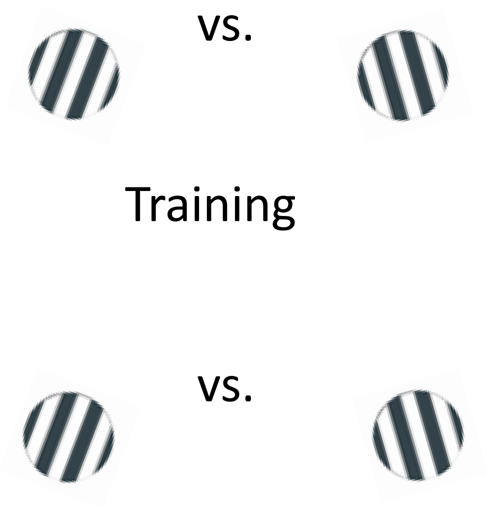
because of budget, time etc. limitations

Q: How to spend trials best?

- It is more useful to study one animal for 1000 hours than to study 1000 animals for one hour” — B. F. Skinner (quoted in Kerlinger and Lee, 1999)

Example

A brain enhancing program proposes to increase not only cognition but also perception. We like to test this proposal using an orientation discrimination task, tested before and after training.



We determine performance in terms of correct responses. We can measure only 1.000 trials (line clockwise/counter-clockwise?) because whatever reason.

Q: How many participants (N) with how many trials (M)?

n = N x M = 1.000

Suppose you have N subjects who each run m trials. The score for the ith trial for the j th subject is defined by

$$X_{ij} = \mu + \alpha_j + e_{ij} \qquad (1)$$

where μ is the overall grand population mean across subjects, α_j is the deviation from the grand mean for subject j , and e_{ij} is the noise for a particular trial for a particular subject.

Suppose you have n subjects who each run m trials. The score for the ith trial for the j th subject is defined by

$$X_{ij} = \mu + \alpha_j + e_{ij} \qquad (1)$$

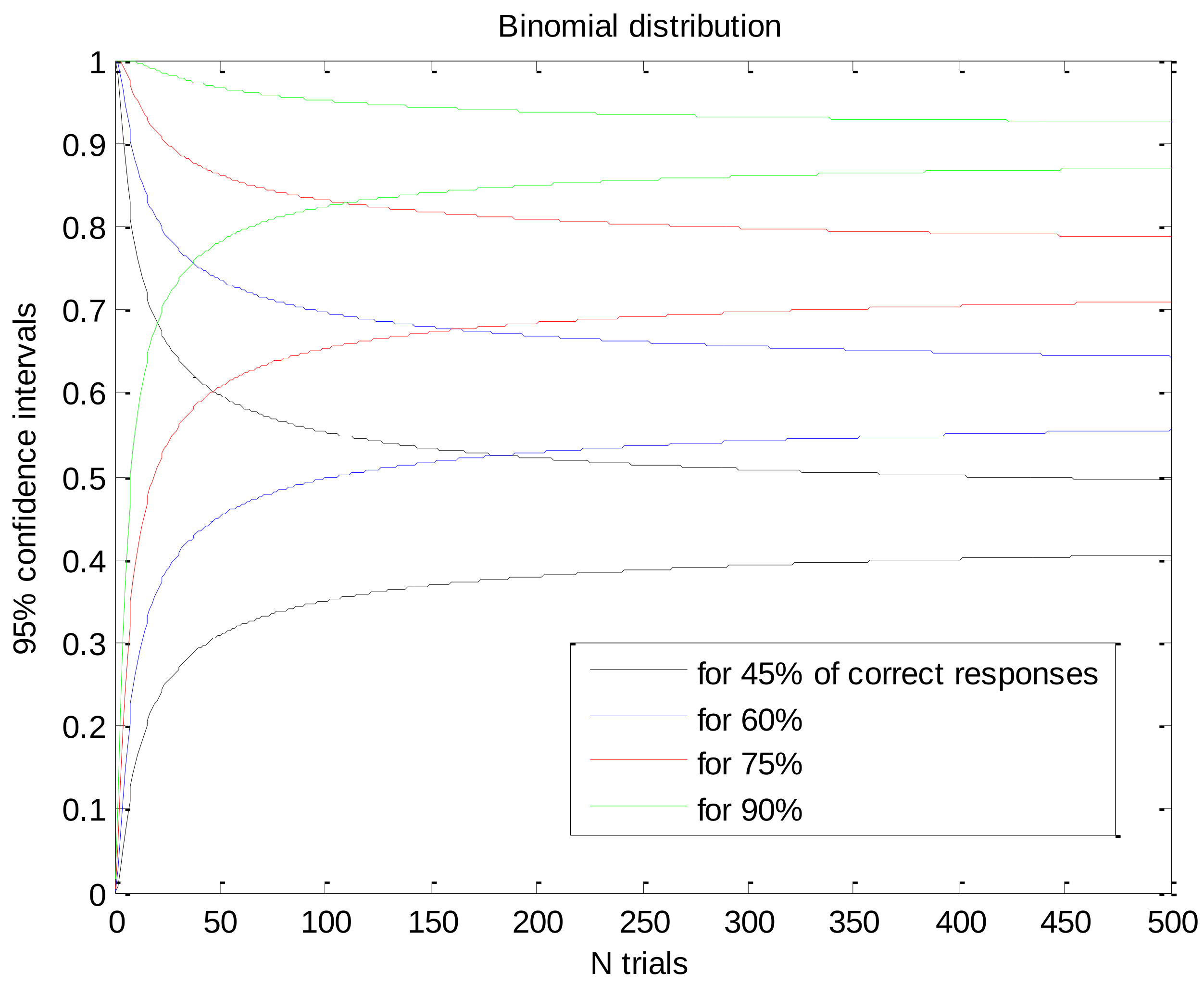
where μ is the overall grand population mean across subjects, α_j is the deviation from the grand mean for subject j , and e_{ij} is the noise for a particular trial for a particular subject.

$$\sigma_{\overline{X}} = \sqrt{\frac{\sigma^2}{N}} = \sqrt{\frac{\sigma_\alpha^2}{N} + \frac{\sigma_\epsilon^2}{MN}}$$

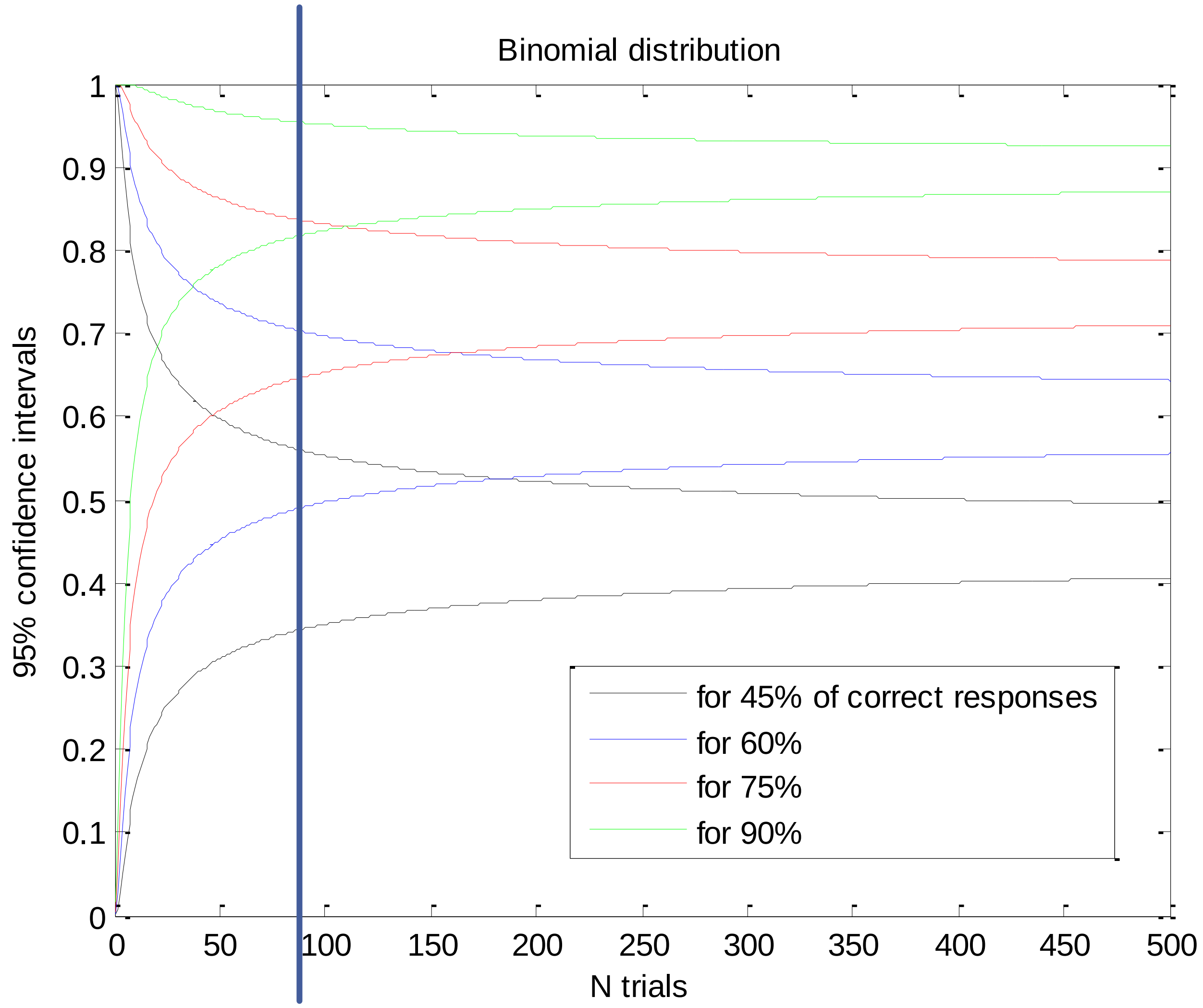
σ_ϵ	N	M	Estimated Power
0.25	100	1	1
0.25	50	2	0.93
0.25	20	5	0.57
0.25	10	10	0.29
0.25	2	50	0.06

σ_ϵ	N	M	Estimated Power
0.25	100	1	1
0.25	50	2	0.93
0.25	20	5	0.57
0.25	10	10	0.29
0.25	2	50	0.06
0.5	100	1	0.99
0.5	50	2	0.91
0.5	20	5	0.56
0.5	10	10	0.29
0.5	2	50	0.06
1	100	1	0.94
1	50	2	0.81
1	20	5	0.49
1	10	10	0.27
1	2	50	0.06
2	100	1	0.61
2	50	2	0.52
2	20	5	0.36
2	10	10	0.23
2	2	50	0.06
4	100	1	0.23
4	50	2	0.21
4	20	5	0.18
4	10	10	0.15
4	2	50	0.06

Percent correct & Binomial Distribution



Percent correct & Binomial Distribution



Take home messages

Take Home Messages

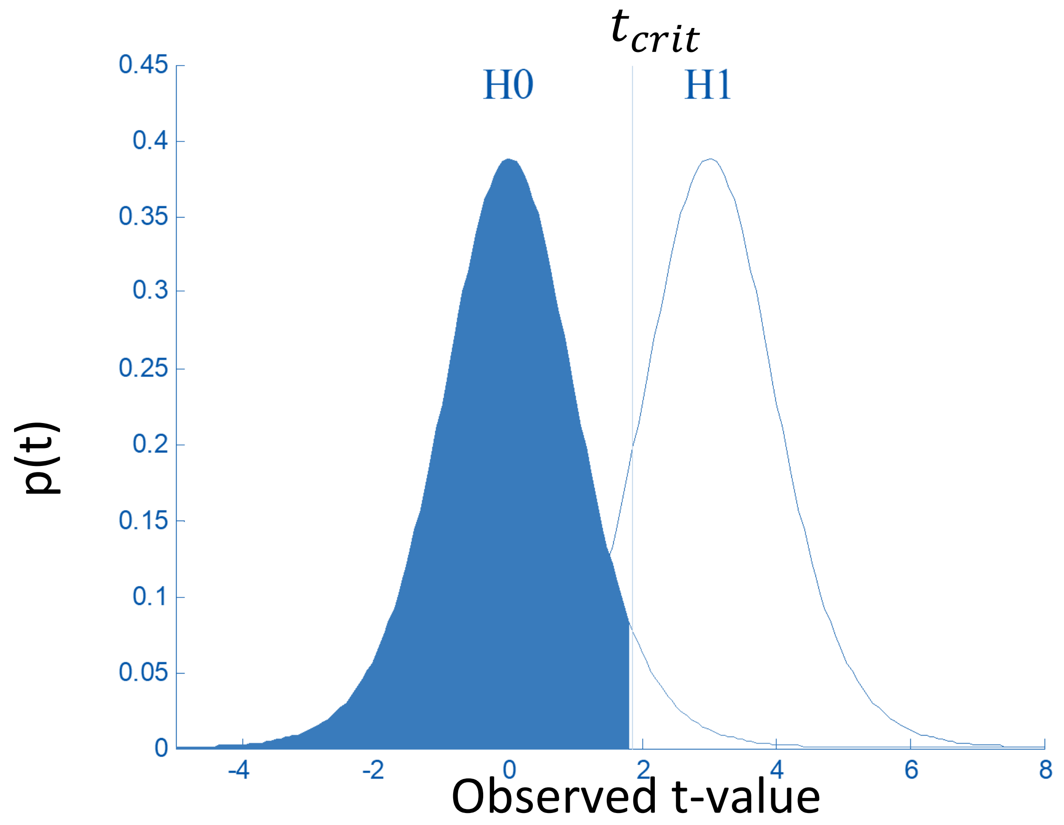
1. Keep your design simple: consider compressing raw data into intermediate variables, which then are subjected to statistical analysis.
2. Compute a power analysis before you do your experiment to check whether there is a real chance that it may show an existing effect.
3. Keep your design simple: if a theory presupposes both significant and null results your power may be strongly reduced.

END Class 7

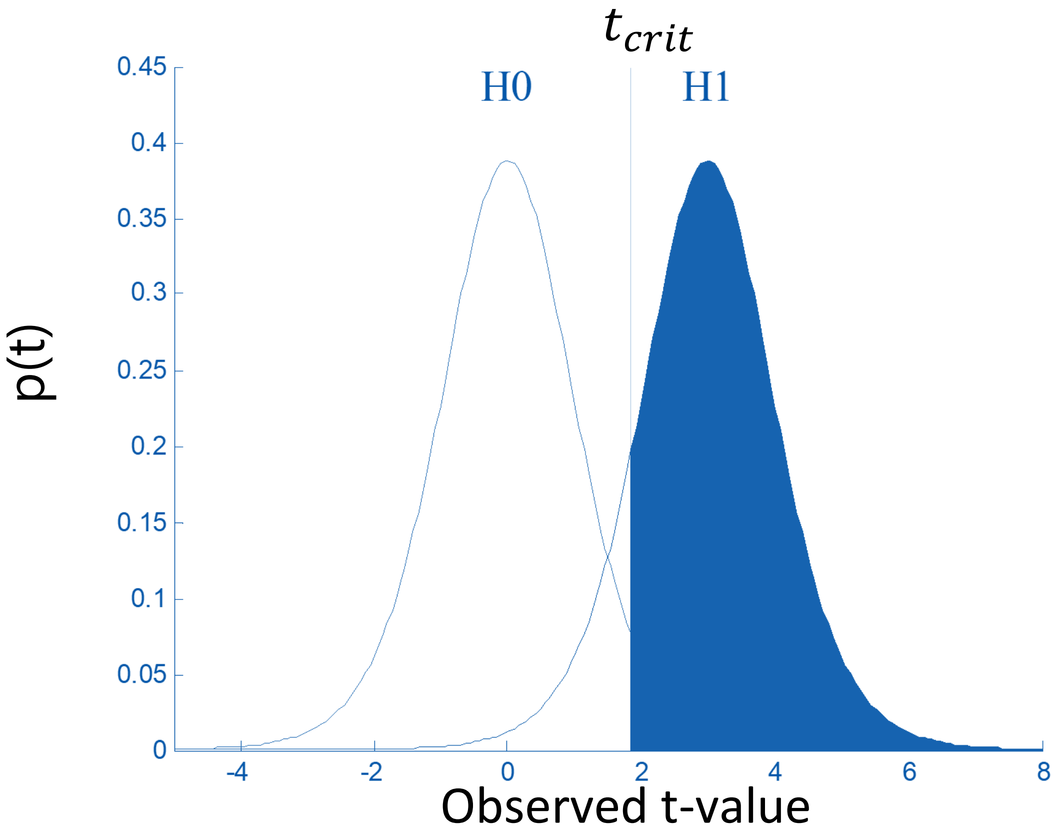
Power Analysis

	H_0 is false	H_0 is true
Decide there is a significant difference	Hit	False Alarm (Type I error)
Do not decide there is a significant difference	Miss (Type II error)	Correct Rejection

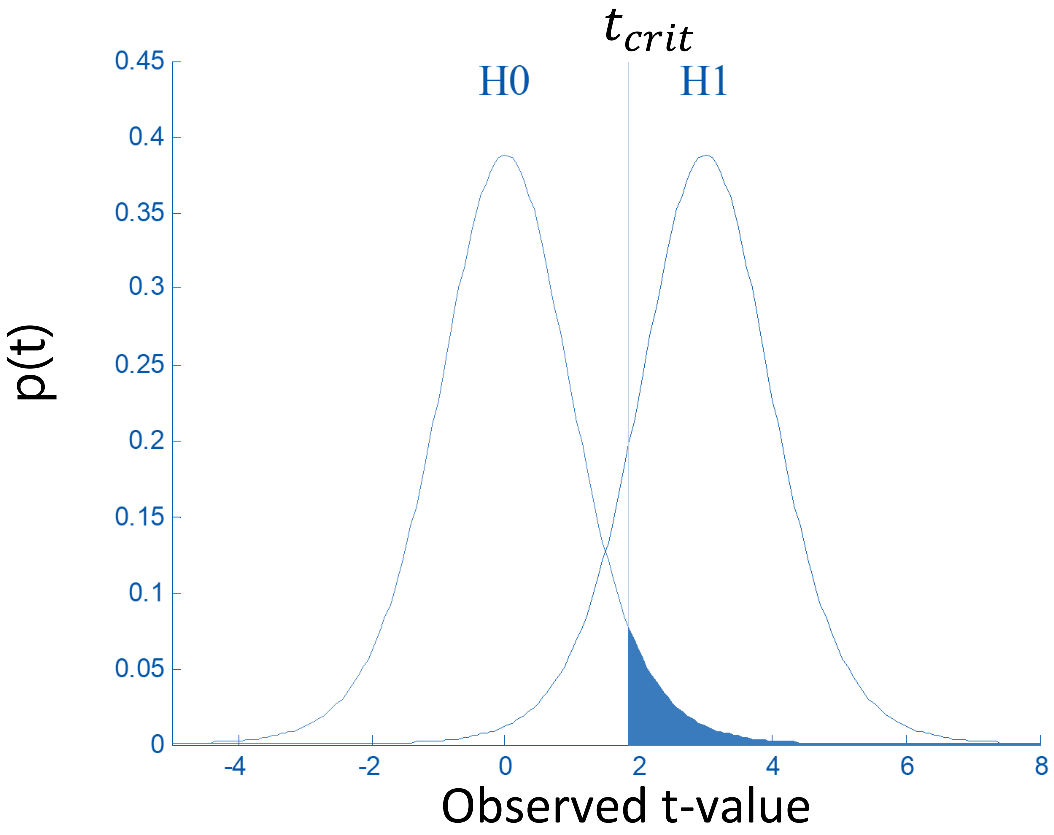
	H_0 is false	H_0 is true
Decide there is a significant difference	Hit **Power**	False Alarm (Type I error)
Do not decide there is a significant difference	Miss (Type II error)	Correct Rejection



	H0 True	H1 True
We Say H0 True	1-α	β
We Say H1 True	α	Power (1-β)



	H0 True	H1 True
We Say H0 True	1-α	Β
We Say H1 True	α	Power (1-β)



	H0 True	H1 True
We Say H0 True	1-α	B
We Say H1 True	α	Power (1-β)