

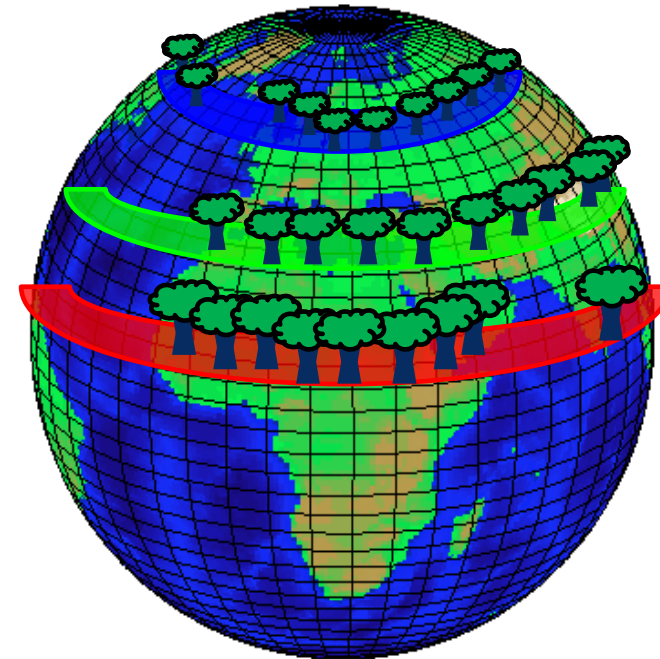
UNDERSTANDING STATISTICS & EXPERIMENTAL DESIGN

No class next week and 4.12

1. Basic Probability Theory
2. Signal Detection Theory (SDT)
3. SDT and Statistics I and II
4. Statistics in a nutshell
5. Multiple Testing
6. ANOVA
7. Experimental Design & Statistics
8. Correlations & PCA
9. Meta-Statistics: Basics
10. Meta-Statistics: Too good to be true
11. Meta-Statistics: How big a problem is publication bias?
12. Meta-Statistics: What do we do now?

Statistics & Data Reduction

- 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

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

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

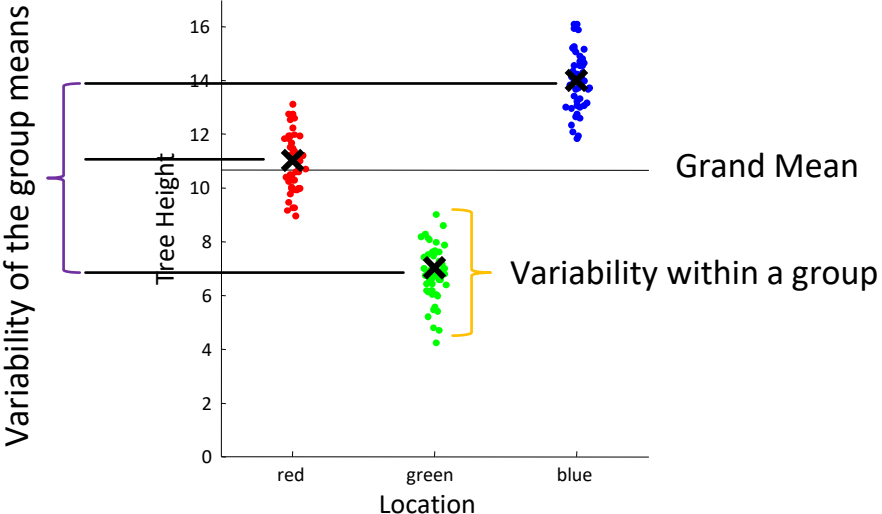
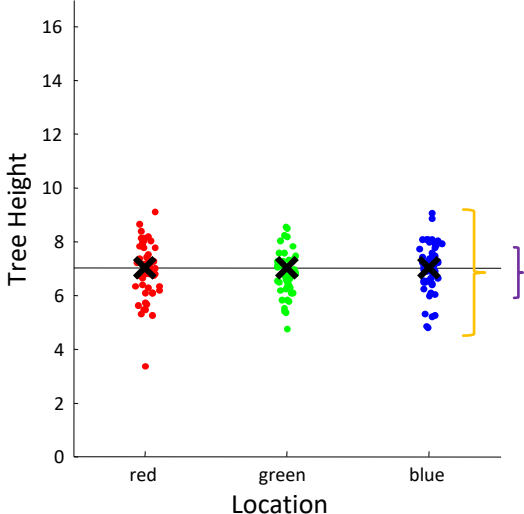
Bonferroni correction: during m independent statistic tests only those results are significant, for which

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

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

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

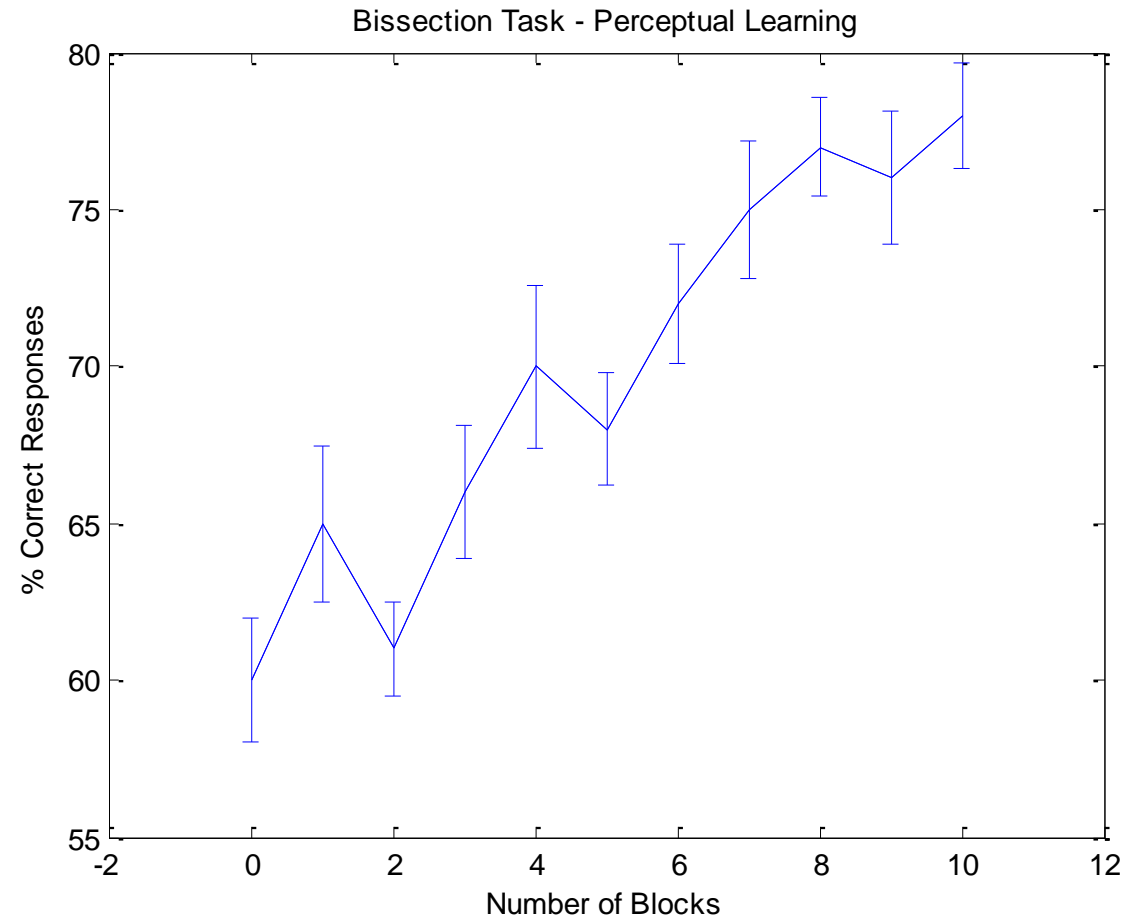
$$F = \frac{\text{Large Number}}{\text{Small Number}} = \text{Large Number}$$



Reduce data I

Model Fitting

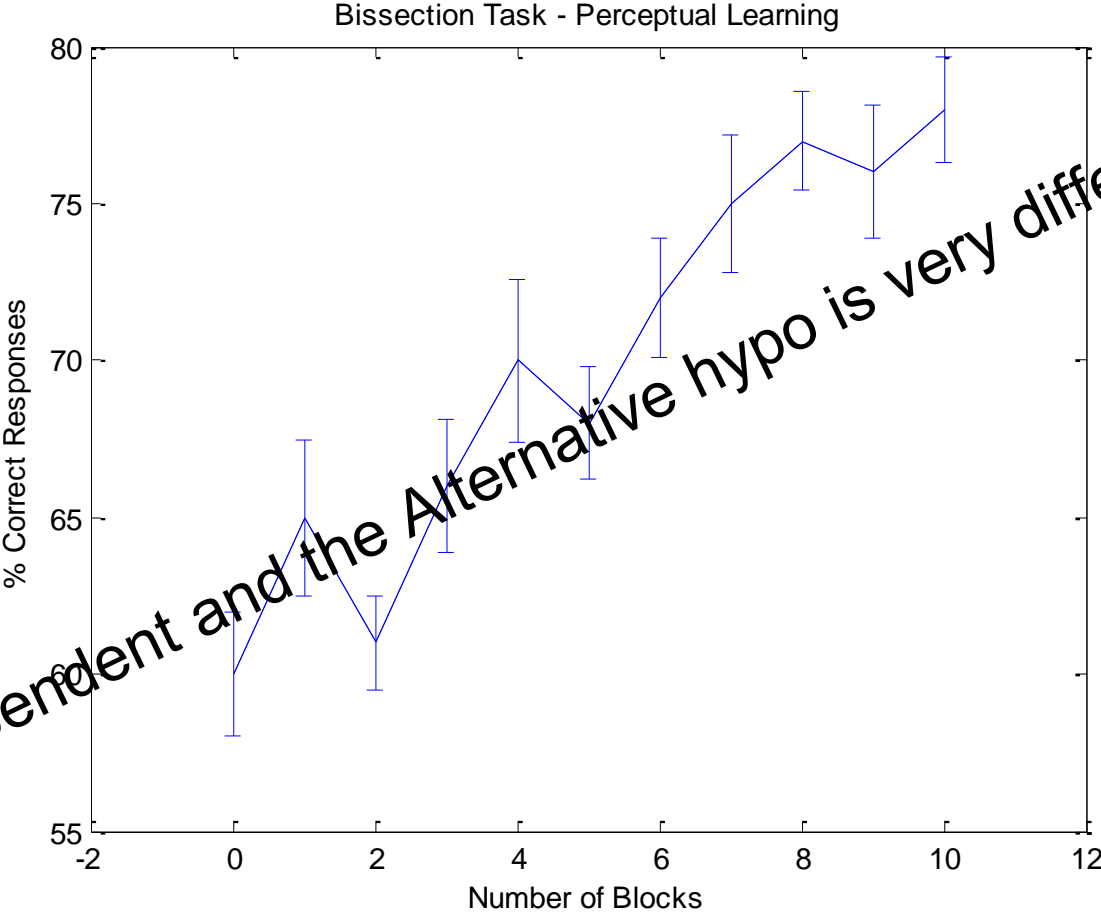
Bisection Task



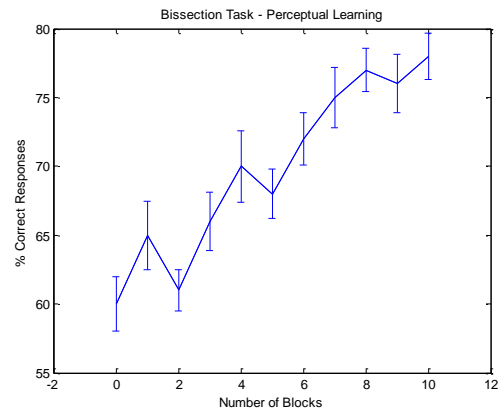
Bisection Task



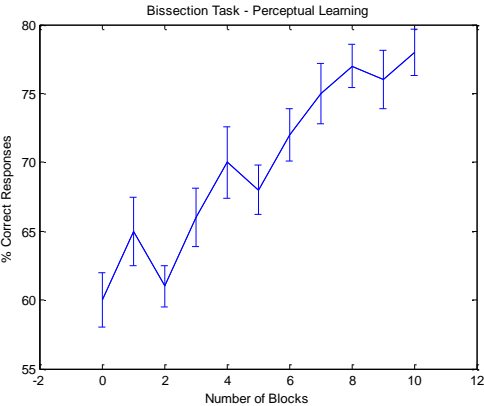
Data are not independent and the Alternative hypo is very different than in Anova



Bisection Task



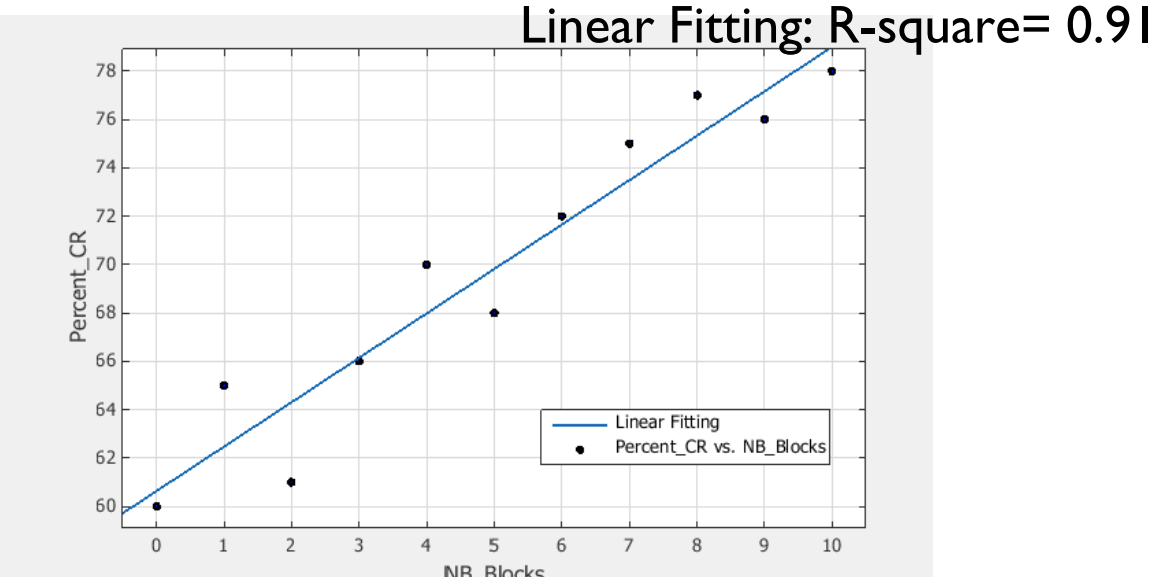
Bisection Task



Results

General model:
 $f(x) = a*x+b$
 Coefficients (with 95% confidence bounds):
 a = 1.836 (1.425, 2.247)
 b = 60.64 (58.2, 63.07)

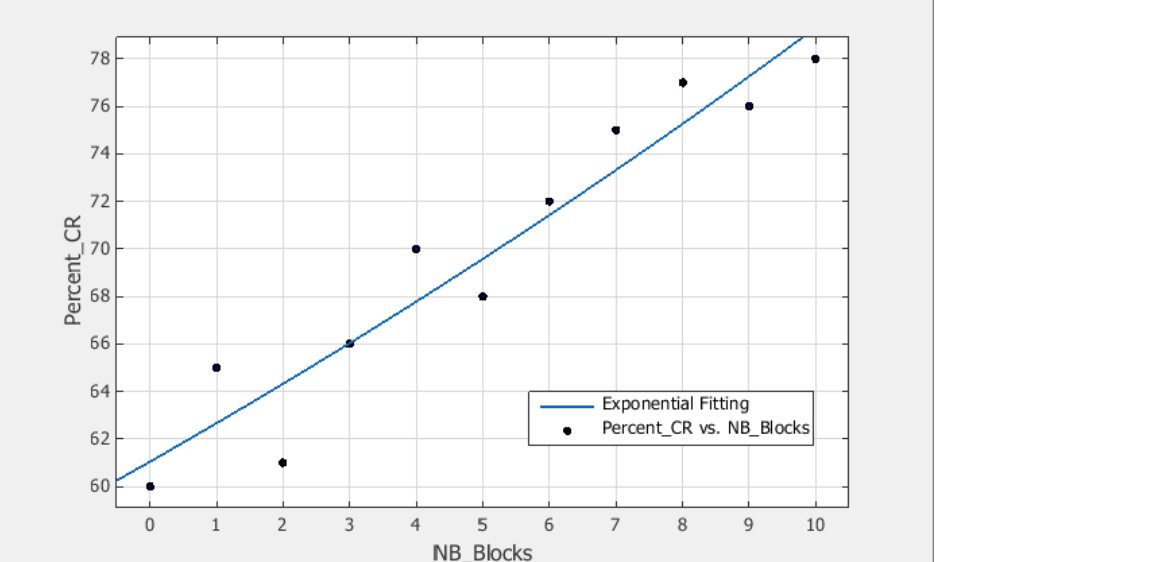
Goodness of fit:
 SSE: 32.69
 R-square: 0.919
 Adjusted R-square: 0.91
 RMSE: 1.906



Results

General model Exp1:
 $f(x) = a*\exp(b*x)$
 Coefficients (with 95% confidence bounds):
 a = 61.05 (58.69, 63.4)
 b = 0.02617 (0.0201, 0.03225)

Goodness of fit:
 SSE: 34.49
 R-square: 0.9145
 Adjusted R-square: 0.9051
 RMSE: 1.958

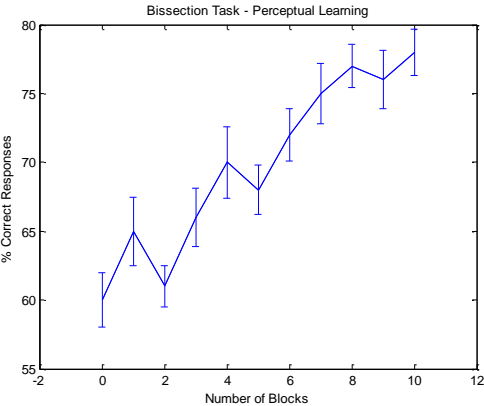


**One sample t -test with
the slopes a**

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

**which serves, now, as our Repeated measures
 t -test**

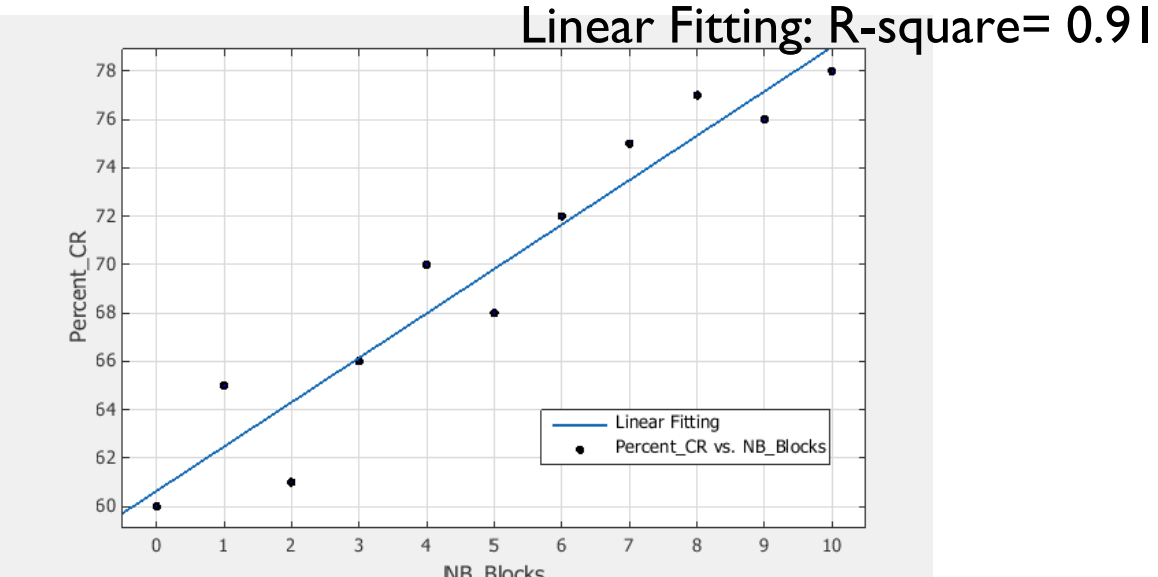
Bisection Task



Results

General model:
 $f(x) = a*x+b$
Coefficients (with 95% confidence bounds):
a = 1.836 (1.425, 2.247)
b = 60.64 (58.2, 63.07)

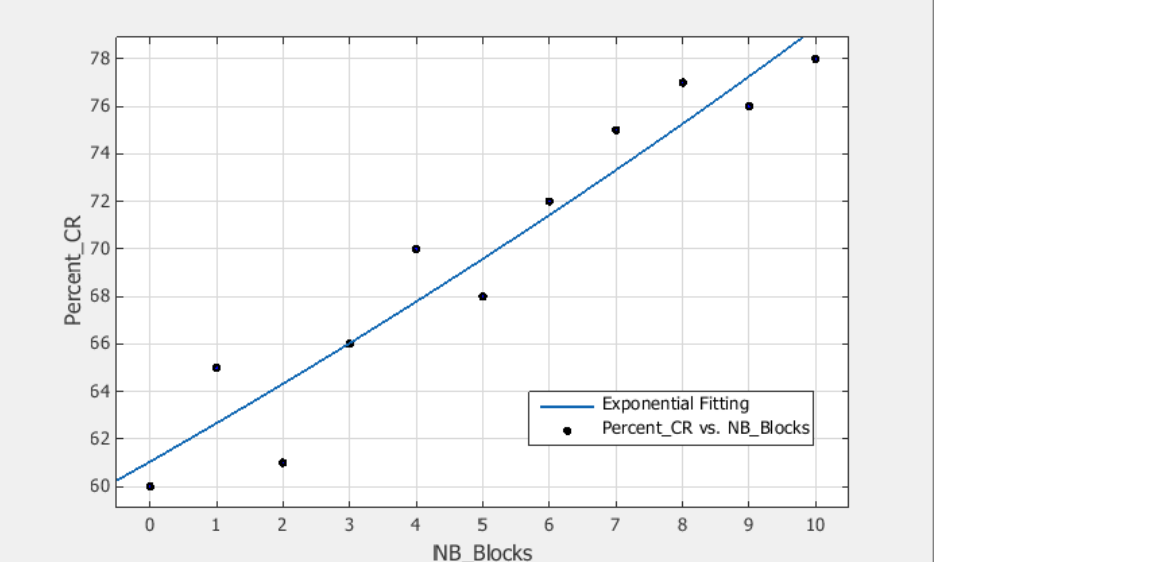
Goodness of fit:
SSE: 32.69
R-square: 0.919
Adjusted R-square: 0.91
RMSE: 1.906



Results

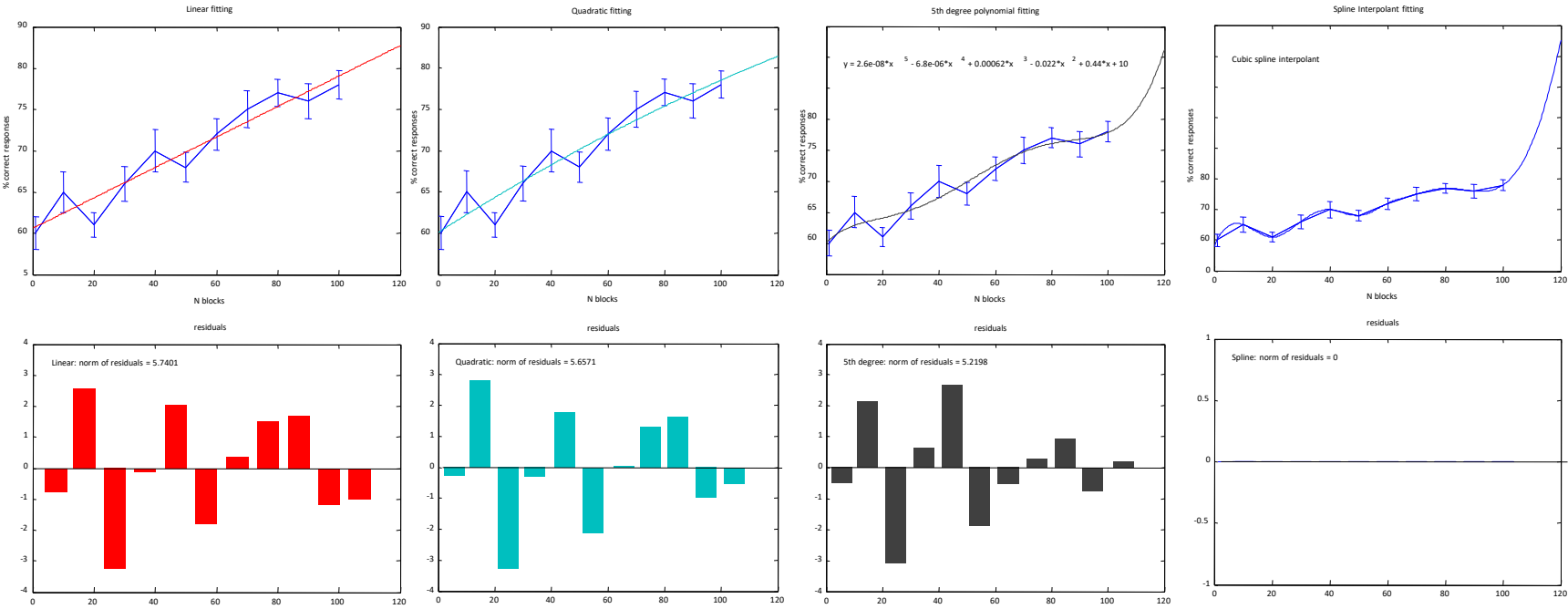
General model Exp1:
 $f(x) = a*\exp(b*x)$
Coefficients (with 95% confidence bounds):
a = 61.05 (58.69, 63.4)
b = 0.02617 (0.0201, 0.03225)

Goodness of fit:
SSE: 34.49
R-square: 0.9145
Adjusted R-square: 0.9051
RMSE: 1.958

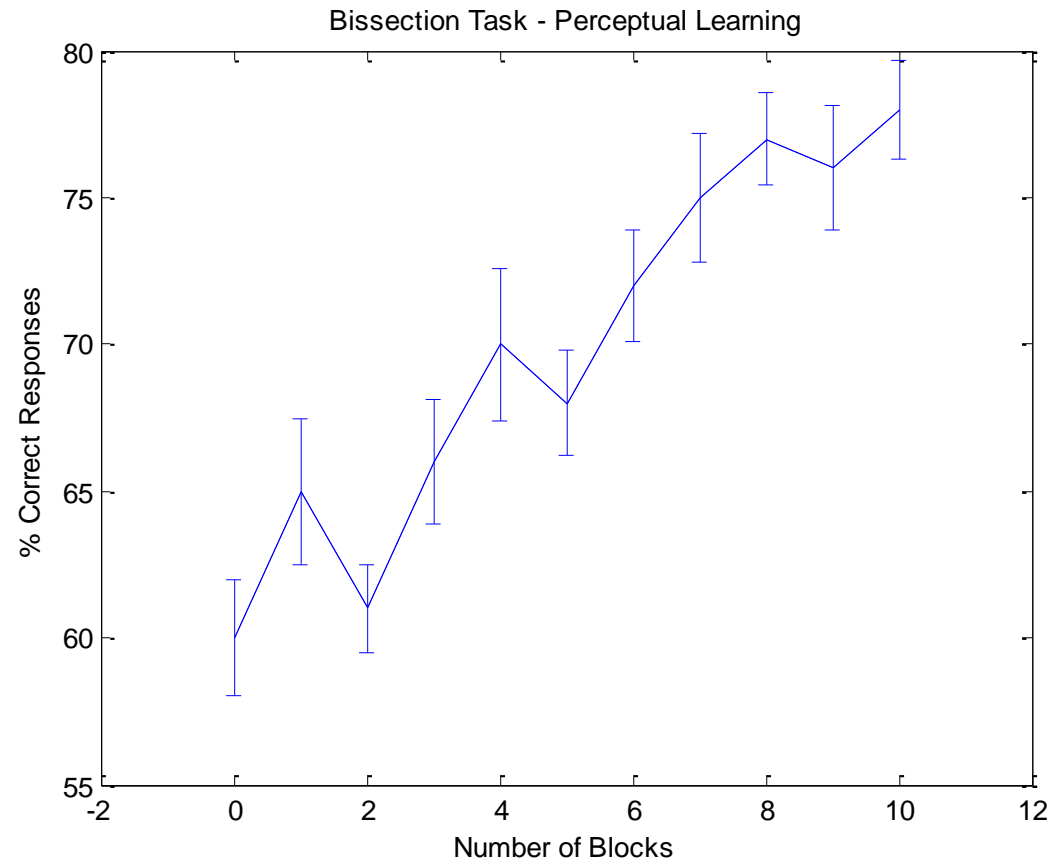


$$y = 0.18 * x + 11$$

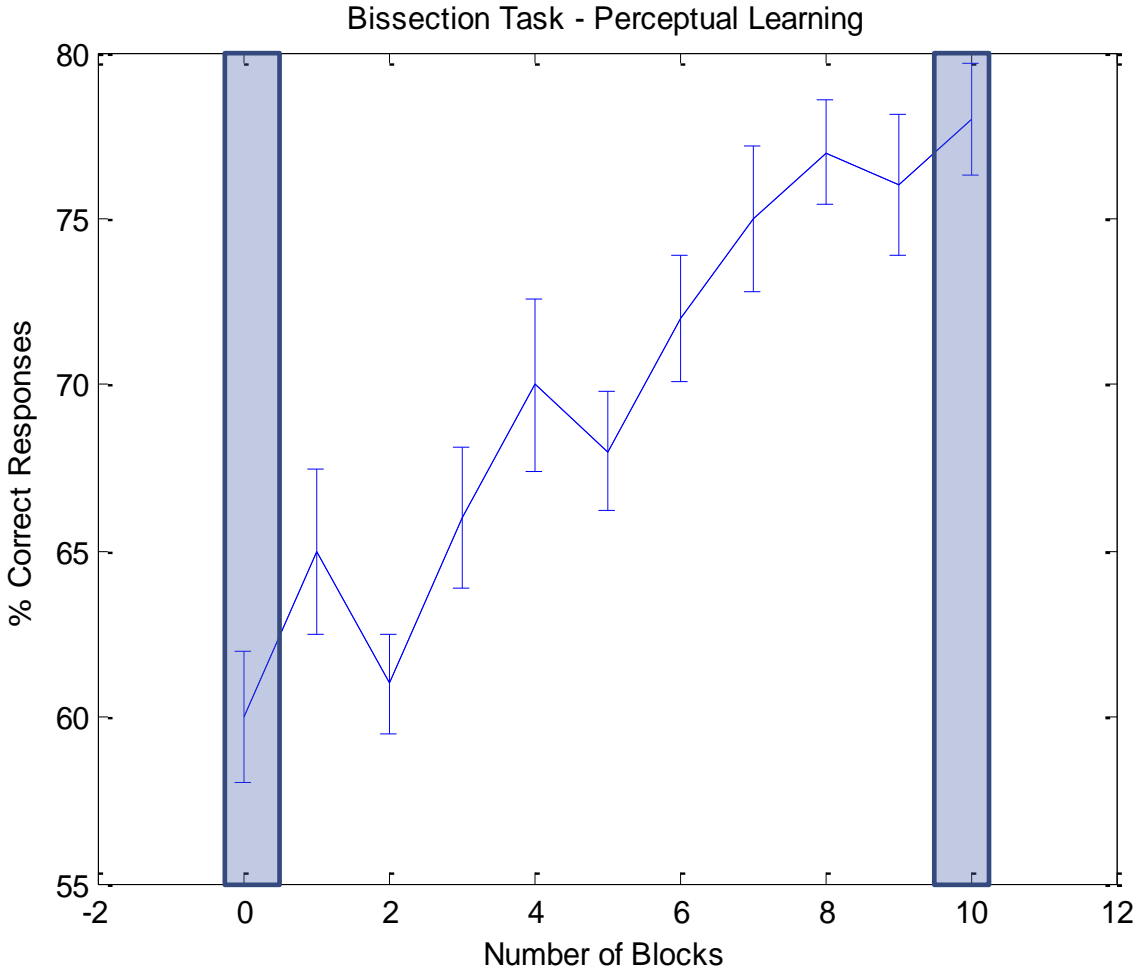
$$y = - 0.00034 * x^2 + 0.22 * x + 10$$



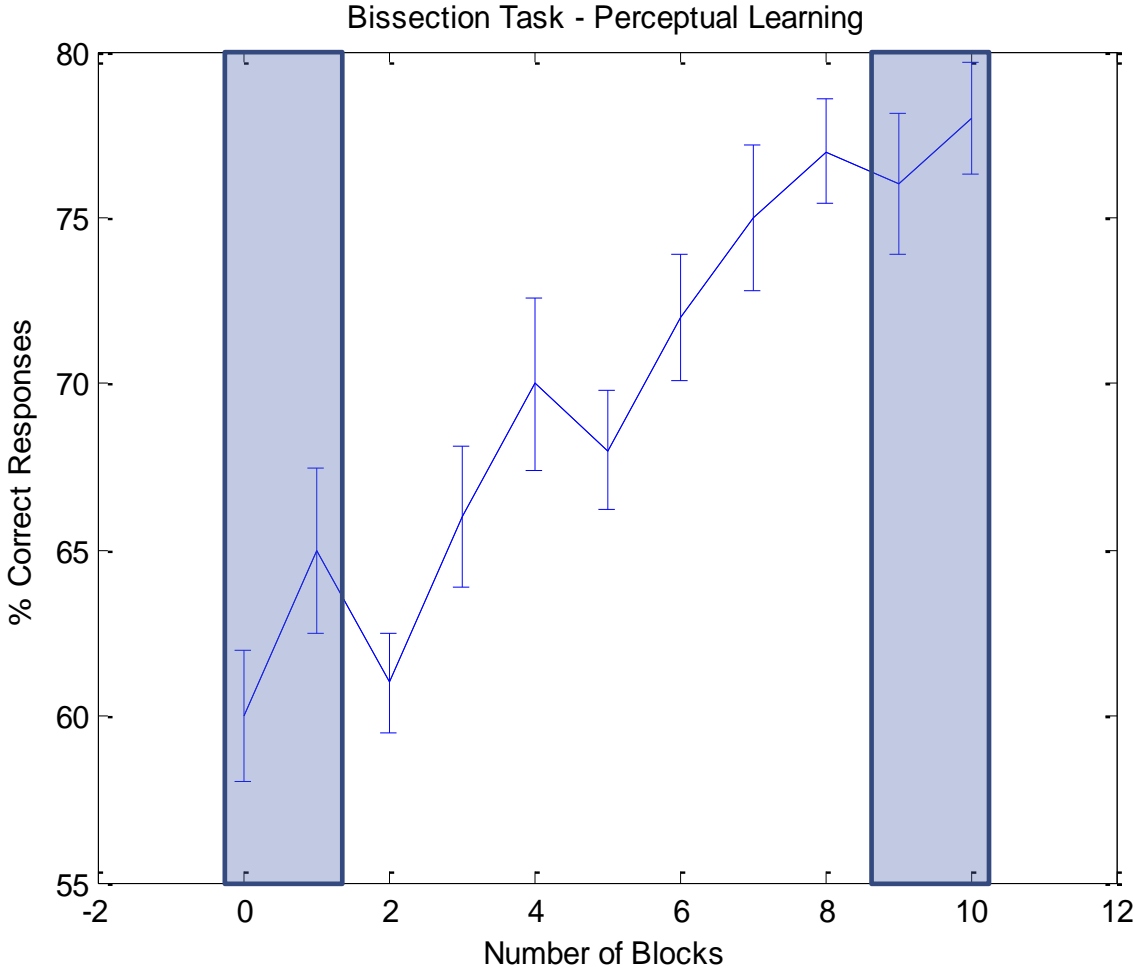
Bisection Task



Bisection Task



Bisection Task



Two sample t -test

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

One sample t -test

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

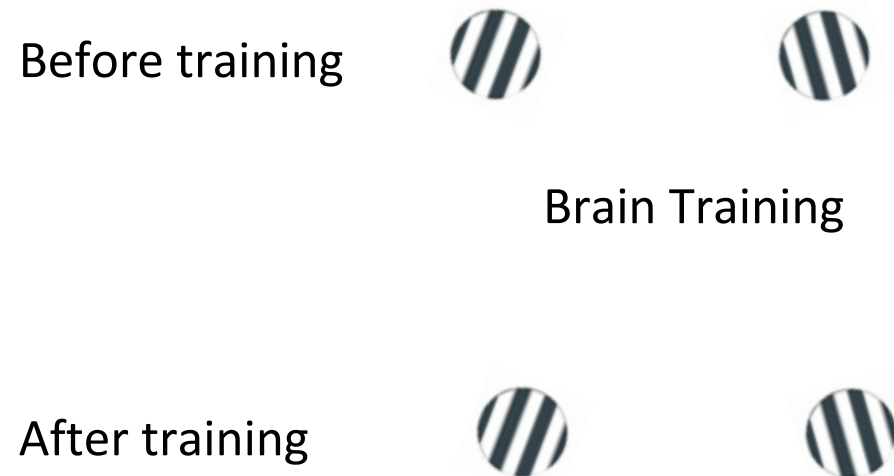
Repeated measures t -test

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.



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.

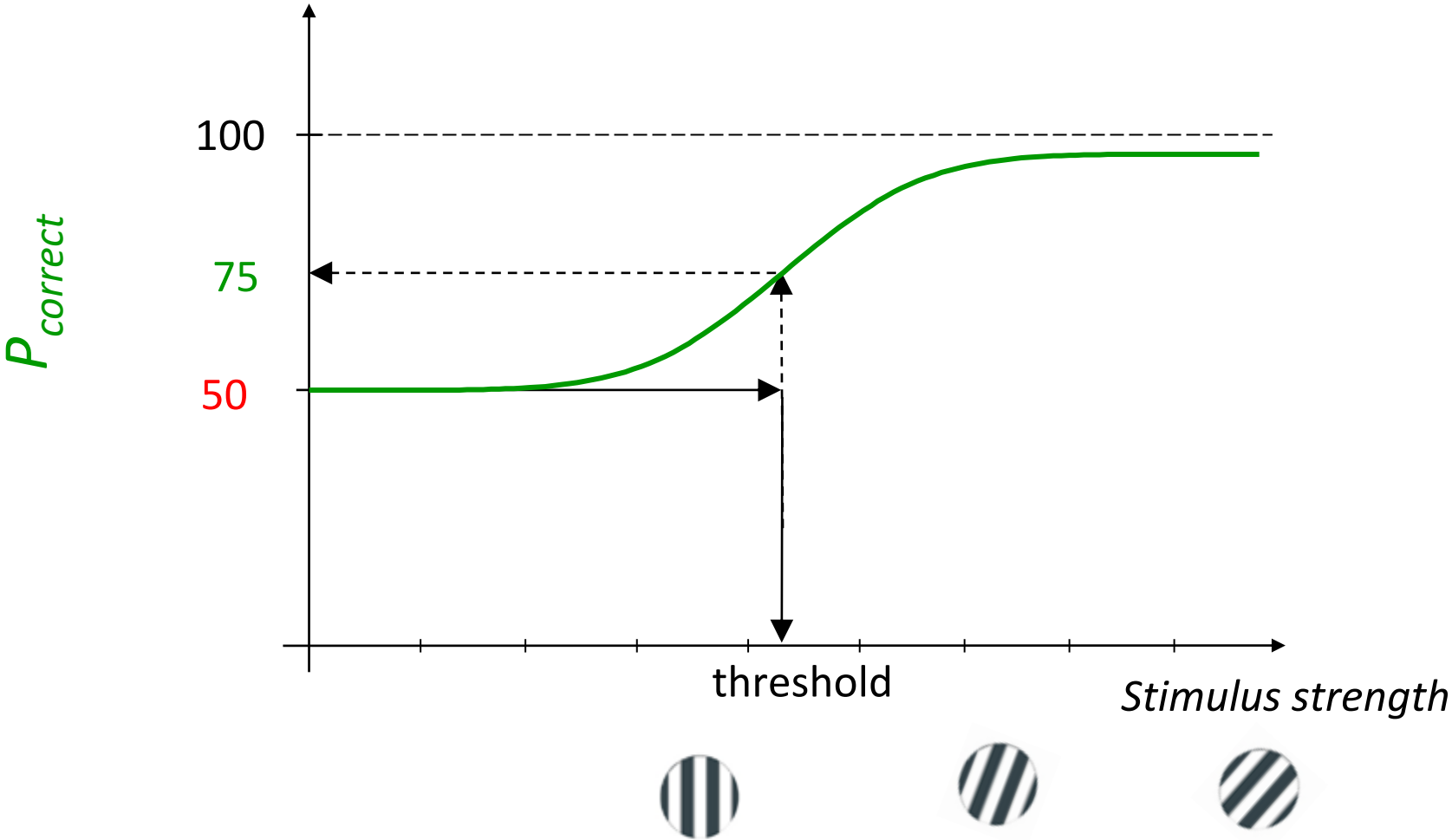
We need to think about floor and ceiling effects and therefore offer a range of stimuli:



Train: Brain enhancing



Compute ANOVA?

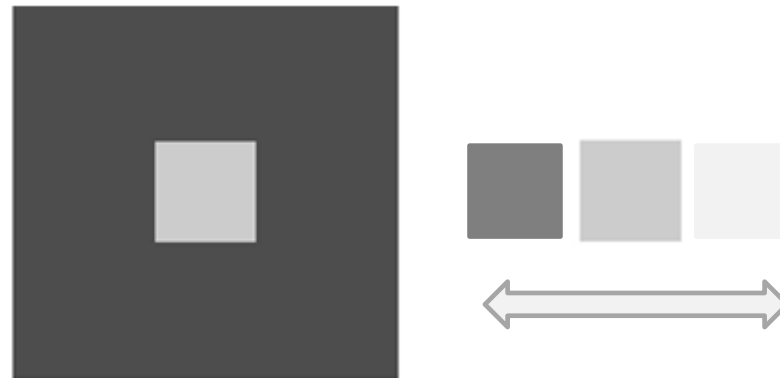


Reduce data IIb: Adjustment Tasks

Orientation adjustment



Color adjustment



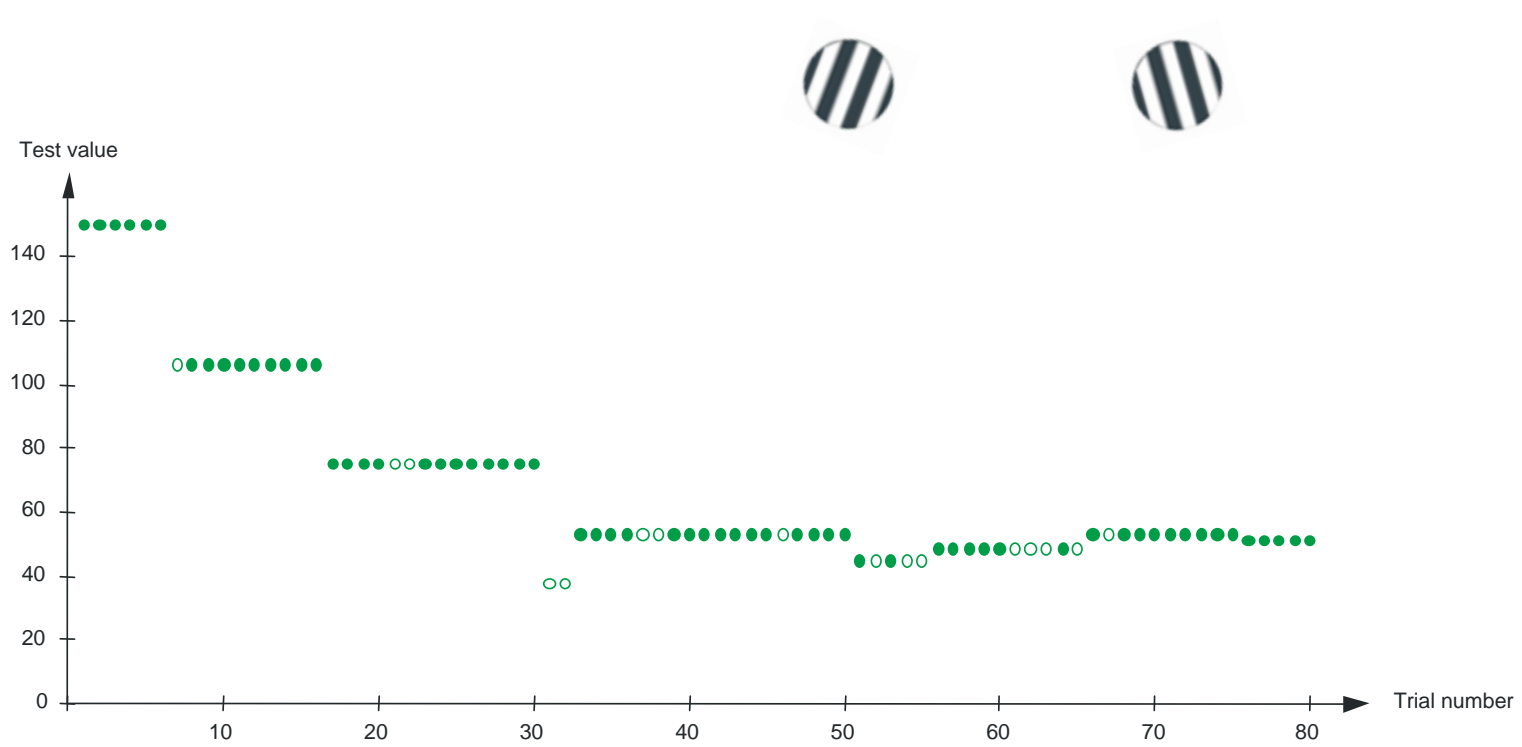
Adapted from: http://www.frontiersin.org/files/Articles/9032/fpsyg-02-00119-HTML/image_m/fpsyg-02-00119-g001.jpg

Reduce data collection II

Adjustment Tasks

Main Problem: outlier detection

Reduce data IIc: Adaptive Methods



Datenliste: "Beide Versatzrichtungen Task1":

PL	LEVEL	ALL	HIT	EP	R-TI_A (se)	R-TI_H (se)	R-TI_F (se)
1	37.7	2	0	0.0%	530 (50)	-	530 (50)
-	44.8	5	2	40.0%	453 (7)	445 (12)	458 (9)
-	48.8	10	6	60.0%	477 (23)	486 (38)	463 (3)
-	51.0	5	5	100.0%	525 (40)	525 (40)	-
-	53.2	28	24	85.7%	532 (21)	533 (21)	528 (83)
-	75.2	14	12	85.7%	485 (17)	478 (18)	529 (36)
-	106.2	10	9	90.0%	468 (29)	471 (32)	437 (0)
-	150.0	6	6	100.0%	431 (20)	431 (20)	-

====

80

Probitanalyse (t=LOG, Pg=50%, Pl=2.0%, Pa=99.0%, SIGMOID, % OHNE RateWK):

50%-Punkt (μ): **50.6** (-6.19, +7.05)

84-50%-Delta (σ): 28.2 (-19.30, +25.56)

korr. Chi: 9.94 (6 Freiheitsgrad(e))

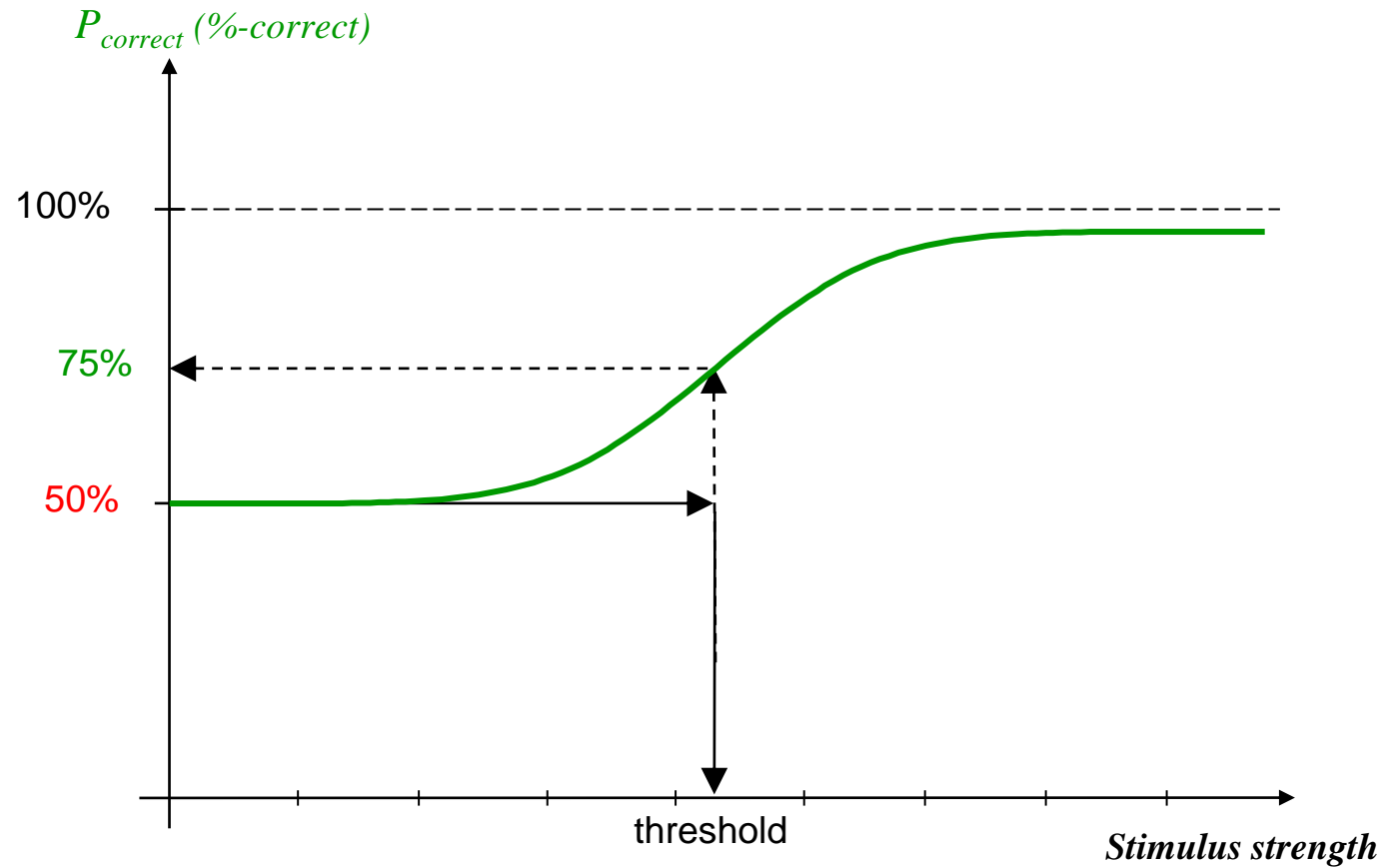
LIKELI-Analyse:

50%-Punkt (μ): **48.8** (-5.76, +6.53)

84-50%-Delta (σ): 26.4 (-17.99, +23.65)

korr. Chi: 10.41 (6 Freiheitsgrad(e))

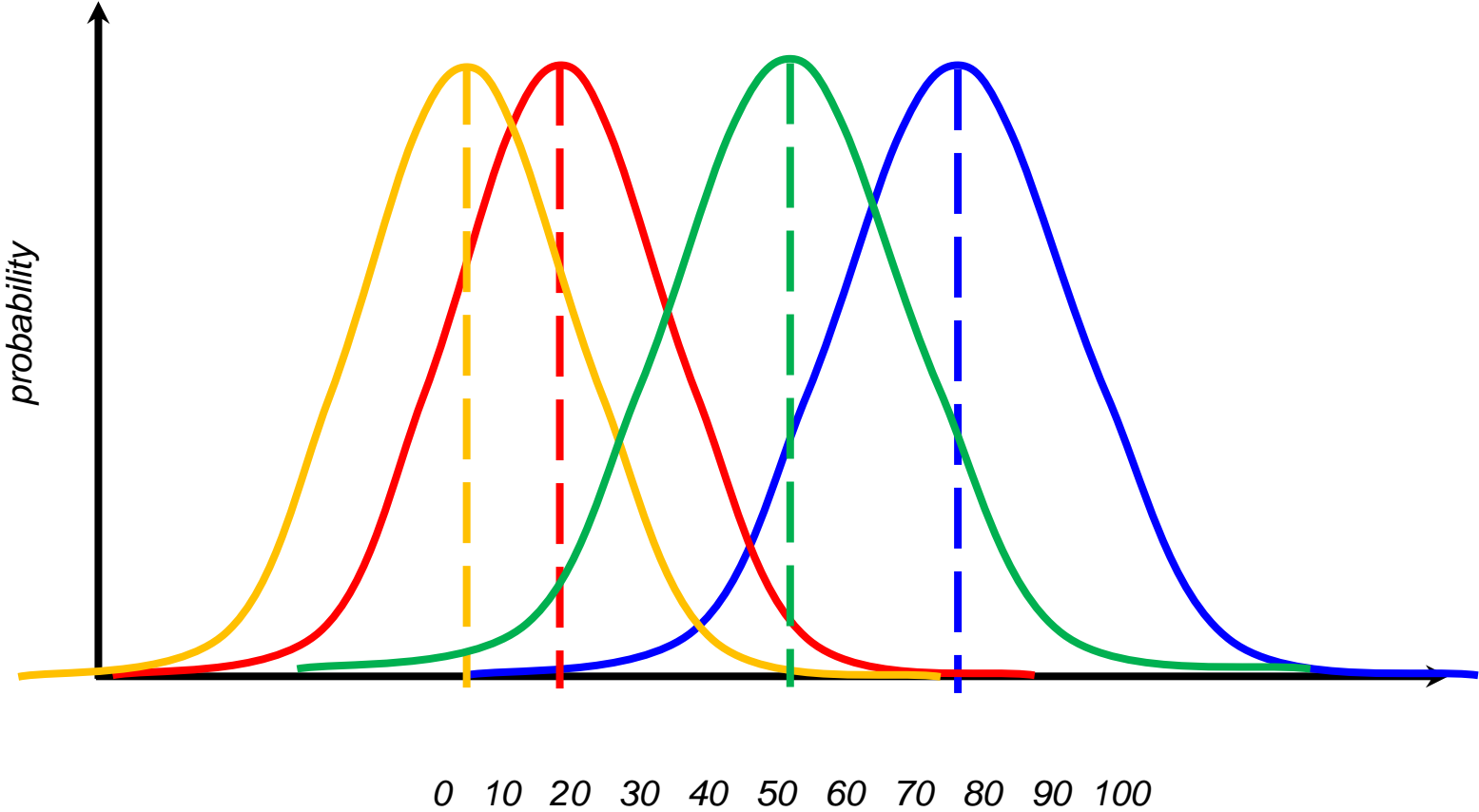
Response & psychometric function (2AFC)

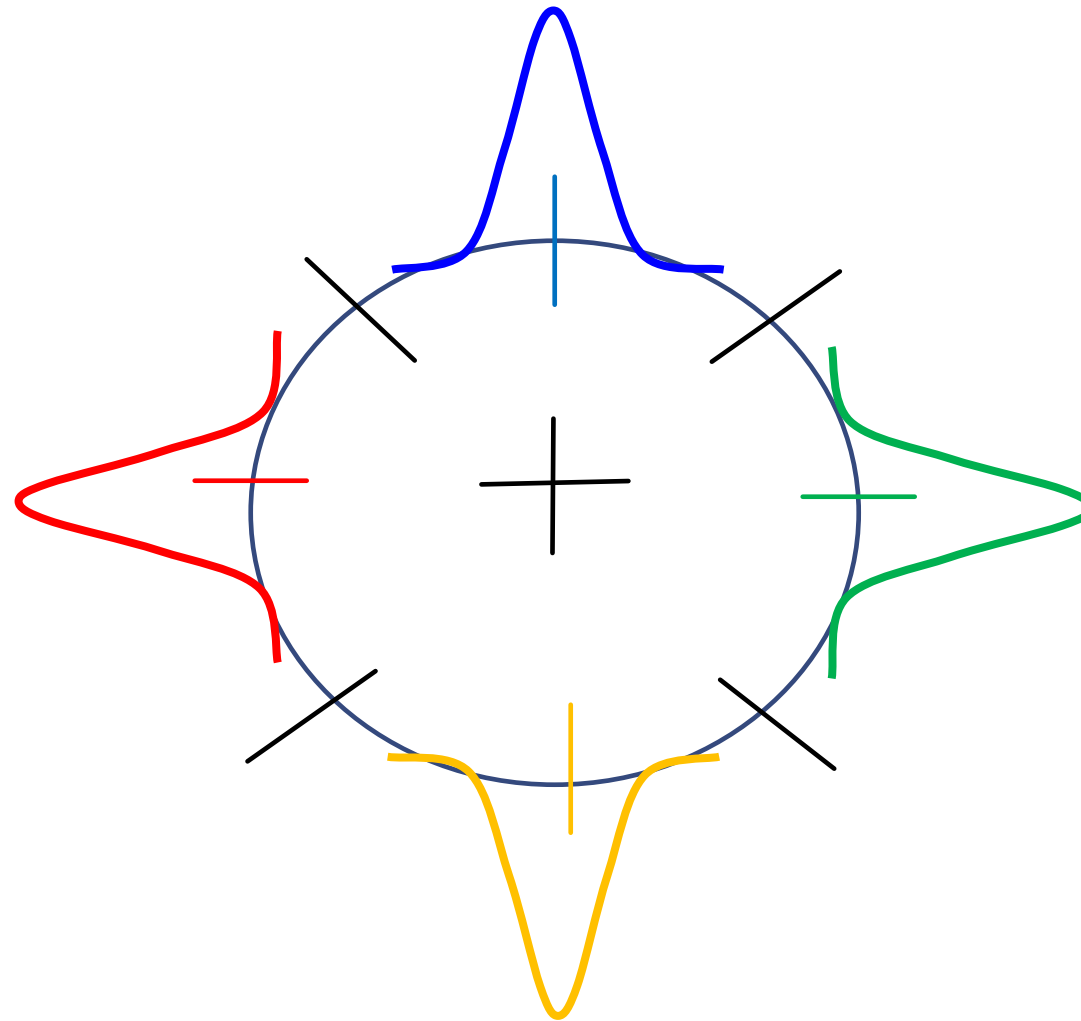


Why not identification task?

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.







Reduce data collection III

Number of trials

Resources are always limited in science

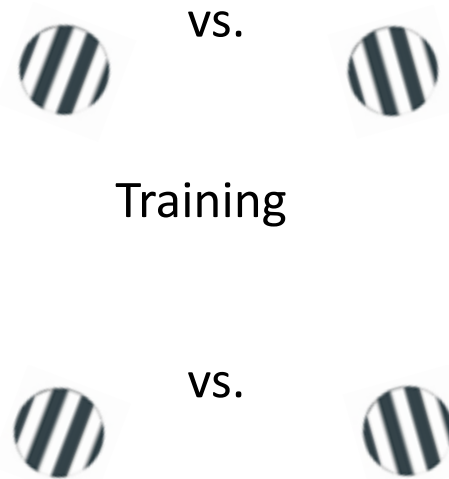
For example, you can screen only

- 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)

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)?

$$n = N \times M = 1.000$$

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

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

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 i th trial for the j th subject is defined by

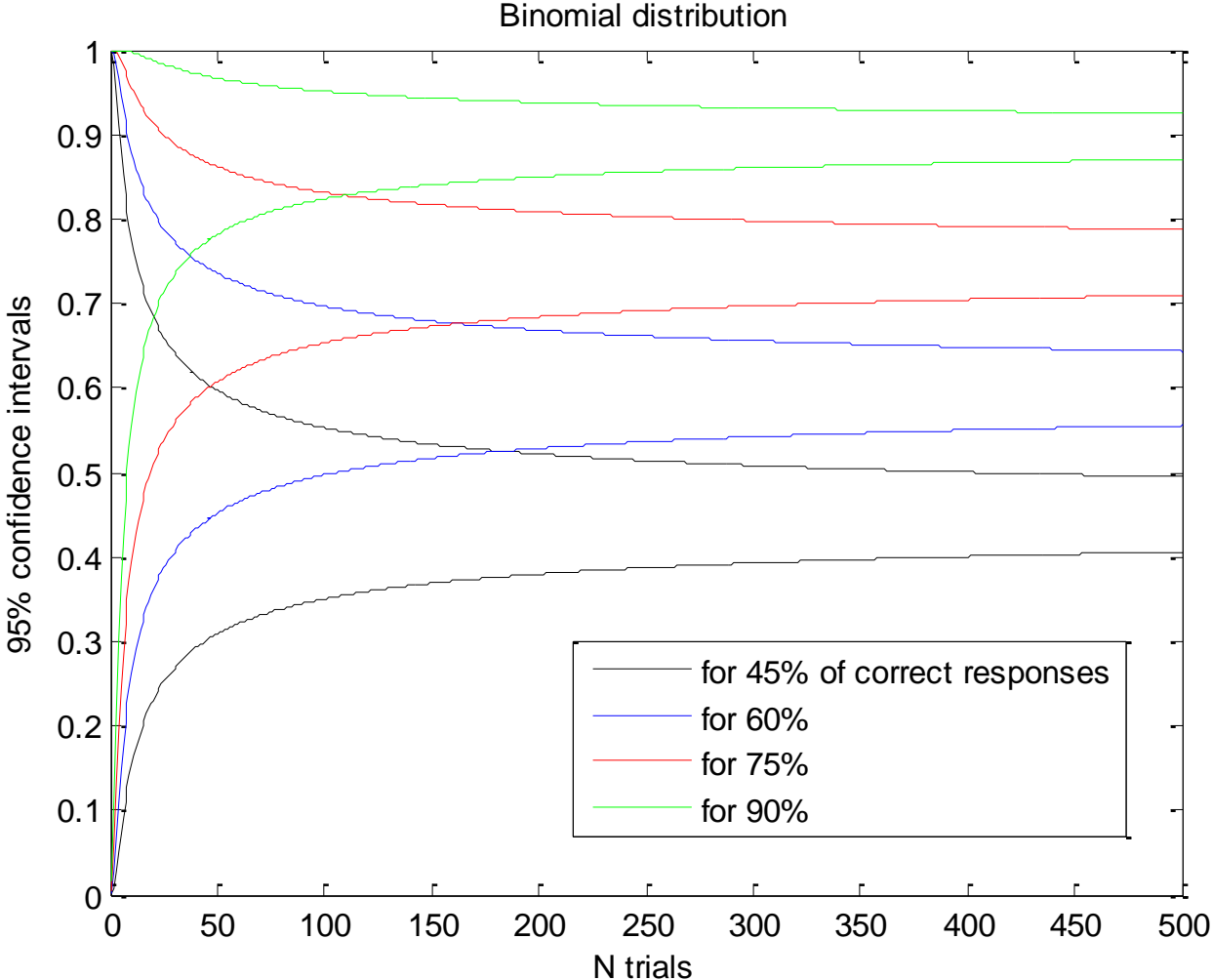
$$X_{ij} = \mu + \alpha_j + e_{ij} \quad (1)$$

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.

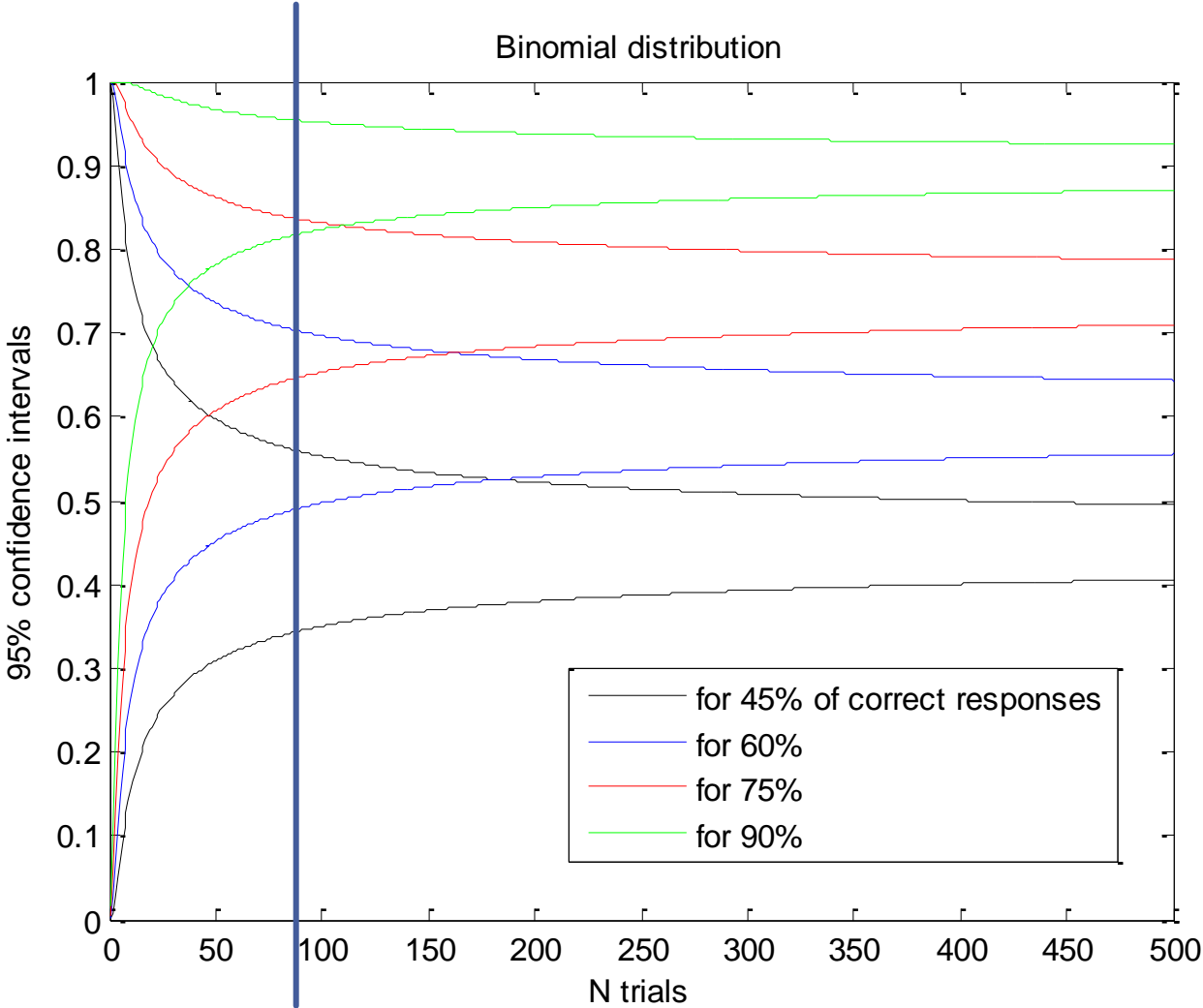
$$\sigma_{\bar{X}} = \sqrt{\frac{\sigma^2}{N}} = \sqrt{\frac{\sigma_{\alpha}^2}{N} + \frac{\sigma_{\epsilon}^2}{MN}}$$

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



Percent correct & Binomial Distribution



Question 10 The table below shows the outcome of an experiment, which had something to do with eye itches. How many participants joined the experiment?

Source	Sum of Squares (SS)	Degrees of freedom (df)	Mean of the Squares (MS)	F-value	p-value
Between	546.50	5	109.30	5.43	< 0.05
Within	483.12	24	20.13		



Question 10 The table below shows the outcome of an experiment, which had something to do with eye itches. How many participants joined the experiment?

Source	Sum of Squares (SS)	Degrees of freedom (df)	Mean of the Squares (MS)	F-value	p-value
Between	546.50	5	109.30	5.43	< 0.05
Within	483.12	24	20.13		

- 29
- 28
- 24
- 21
- 30
- 4

Question 20 A study wants to investigate whether peak eye itch intensity, which change in a sinusoidal fashion, are stronger during a statistics lecture than in leisure time. As a marker for the intensity of the eye itches, pupil size was measured in 100 students 6 times during the class of 90min and 6 times during 90min of leisure time. How to best compute statistics?

1. For each participant, fit a sine function to the class and leisure time conditions, extract the phases , take the difference of phases, and use a one sample t-test to test whether this value is significantly different from zero.
2. For each participant, fit a sine function to the class and leisure time conditions, extract the amplitudes, and compute a dependent t-test to test whether there is significant difference.
3. None of them
4. For each participant, fit a cosine function to the class and leisure time conditions, extract the amplitudes, take the difference of the amplitudes, and use a one sample t-test to test whether this value is significantly different from zero.

Question 20 A study wants to investigate whether peak eye itch intensity, which change in a sinusoidal fashion, are stronger during a statistics lecture than in leisure time. As a marker for the intensity of the eye itches, pupil size was measured in 100 students 6 times during the class of 90min and 6 times during 90min of leisure time. How to best compute statistics?

1. For each participant, fit a sine function to the class and leisure time conditions, extract the phases , take the difference of phases, and use a one sample t-test to test whether this value is significantly different from zero.
2. For each participant, fit a sine function to the class and leisure time conditions, extract the amplitudes, and compute a dependent t-test to test whether there is significant difference.
3. None of them
4. For each participant, fit a cosine function to the class and leisure time conditions, extract the amplitudes, take the difference of the amplitudes, and use a one sample t-test to test whether this value is significantly different from zero.

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