Group analysis of fMRI data

Roel Willems,
FC Donders Centre Nijmegen
roel.willems@fcdonders.ru.nl

Reminder: Analysis of data from 1 person

- Robert:

Why group analysis?

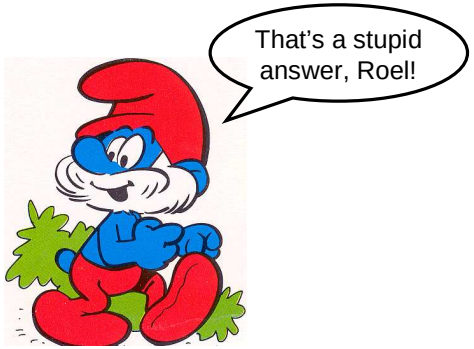
- Results from Robert only apply to Robert
- → Generalise to population level
- [the boring stuff again]
- Big deal: if you want to generalise you measure several people from the population

What is a population?

- All men whose first name starts with ‘R’
- Butterfly collectors
- All women below 1.70 m length
- All people with a horse called ‘Blacky’
- Speakers of German
- All humans
- All inhabitants of Dorpsstraat in Ons Dorp

Sample from a population

- Sometimes it's possible to measure complete population, most of the time this is not possible
- → Sample from population
- Should be representative
  - Random
  - Not biased
- We don't do this
  - A) Why not?
  - B) Is this a problem?
    - Not if sample is representative with respect to factors that matter for the behaviour under study



Sample

- The criteria that matter?
  - Age?
  - Gender?
  - SES?
  - Smoking?
- Example: Gender

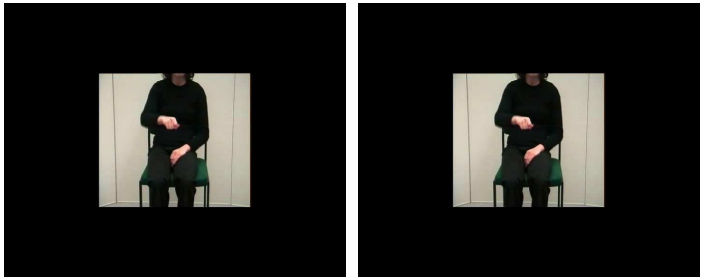
Gender and fMRI

- More variability in BOLD signal across women as compared to across men:
  - Variability in hormonal levels due to menstrual cycle
  - Variability across subjects is bad
  - → Solution: only scan men
- When is this perhaps a good idea?
- When is this perhaps a bad idea?

Conclusion: sample

- Sample from population you're interested in (within practical limits)
- Which criteria matter is often based on 'common sense' criteria

Our experiment



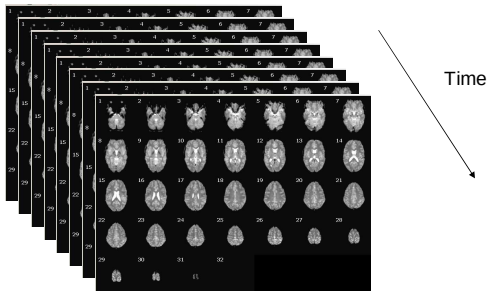
2 Conditions **1)Match**
2) Mismatch

Methods

- N = 16 participants (11 female; mean age = 22.3 years, range = 19.3 - 27.4 years)
- Right-handed, no hearing problems, normal or corrected-to-normal vision
- No known neurological history; not on medication
- Data acquisition was performed using a Siemens 'Trio' MR-scanner with 3 Tesla magnetic field strength. Whole-brain echo-planar images (EPIs) were acquired using a standard bird-cage head coil with single pulse excitation with ascending slice order (TR=2130 ms, TE=30 ms, flip angle=80 degrees, 32 slices, slice thickness=3mm, 0.5 mm gap between slices, voxel size 3.5x3.5x3.5 mm).

Our data

- Subject 1:

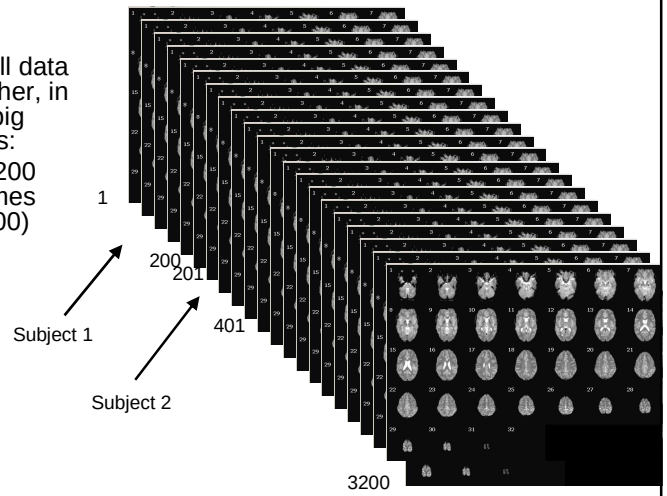


- ……
- Subject 16

200 volumes / scans per subject
(how many TRs?)
(How many slices?)

Analysis 1: Fixed effects (FFX)

- Put all data together, in one big series:
- 16 * 200 volumes (=3200)



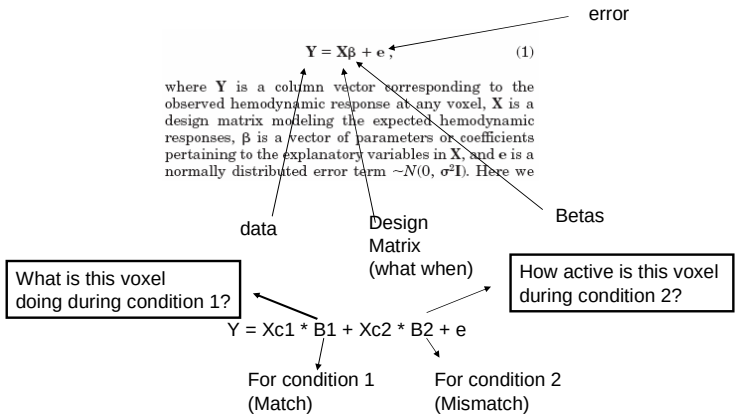
FFX

- Design matrix: what happened when?
- 1 200 401 3200
- c1 c2 c2 c1 c1 c2 c2 c1 c1 c2 c2 c1 c1 c2 c2 c1 c2 c2 c2 c1 c1

Subject 1 | Subject 2

(Convolve with HRF)

Estimate beta weights



Results for one voxel

- 1 beta per condition:
- B1: 3.75
- B2: 4.75
- Condition 2 > condition 1?
- T= B2-B1 / std

Results

	Cond1	Cond2	
S1	1	3	
S2	3	7	
S3	3	8	
S4	2	4	
S5	6	9	
S6	3	5	
S7	5	5	
S8	8	7	
S9	2	4	
S10	3	6	
S11	5	4	
S12	4	7	
	3.75	**4.75**	

$$T = \frac{B2 - B1}{se}$$

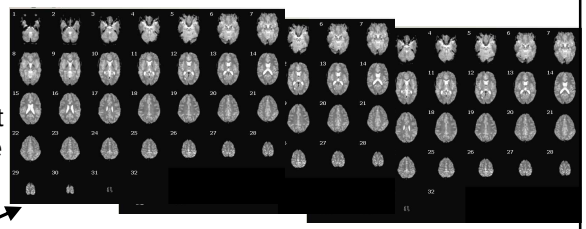
Results

	Cond1	Cond2	
S1	1	3	
S2	3	7	
S3	3	8	
S4	9	2	
S5	6	9	
S6	3	5	
S7	5	5	
S8	8	7	
S9	10	0	
S10	3	6	
S11	5	4	
S12	4	7	
	5.00	**5.25**	

$$T = \frac{B2 - B1}{se}$$

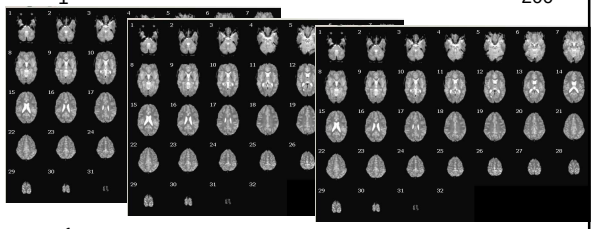
Analysis 2: Random effects (RFX)

- First, look at effect in each subject
- Second, see if effect is stable over subjects



1 200

Subject 1



Subject 2

1 200

RFX

- Design matrix: what happened when?

- 1 200
- c1 c2 c2 c1 c1 c2 c2 c1 c1 c2 c2 c1 c1 c2 c2 c1 c2 c2 c2 c1 c1

Subject 1

- 1 200
- c1 c1 c2 c1 c2 c2 c2 c1 c1 c2 c2 c2 c1 c2 c2 c1 c2 c2 c1 c1 c2

Subject 2

FFX

- Design matrix: what happened when?

- 1 200 400 3200
- c1 c2 c2 c1 c1 c2 c2 c1 c1 c2 c2 c1 c1 c2 c2 c1 c2 c2 c2 c1 c1

Subject 1 | Subject 2

Results for one voxel

- 1 **difference** per subject:
- Subject 1: B2 – B1: 4-2
- Subject 2: B2 – B1: 7-3

- Subject 16: B2 – B1: 3 - 5

Results

	Cond1	Cond2	Difference c2-c1
S1	1	3	2
S2	3	7	4
S3	3	8	5
S4	2	4	2
S5	6	9	3
S6	3	5	2
S7	5	5	0
S8	8	7	-1
S9	2	4	2
S10	3	6	3
S11	5	4	-1
S12	4	7	3

$$T = \frac{\text{difference}}{\text{se(difference)}}$$

Are differences > 0?

Results: FFX

	Cond1	Cond2	
S1	1	3	
S2	3	7	
S3	3	8	
S4	2	4	
S5	6	9	
S6	3	5	
S7	5	5	
S8	8	7	
S9	2	4	
S10	3	6	
S11	5	4	
S12	4	7	
	3.75	**4.75**	

$$T = \frac{B2 - B1}{se}$$

Results: RFX

	Cond1	Cond2	Difference c2-c1
S1	1	3	2
S2	3	7	4
S3	3	8	5
S4	2	4	2
S5	6	9	3
S6	3	5	2
S7	5	5	0
S8	8	7	-1
S9	2	4	2
S10	3	6	3
S11	5	4	-1
S12	4	7	3

$$T = \frac{difference}{se}$$

Are differences > 0?

- **FFX**
- df: N*number of conditions
- Effect size: mean from all subjects
- Variance: divided by N*number of conditions
- SE: Var/N*M

As if all observations come from one subject!

- **RFX**
- df: N-1
- Effect size: mean from all subjects
- Variance: variance across subjects
- SE: Var / N

When is an effect consistent?

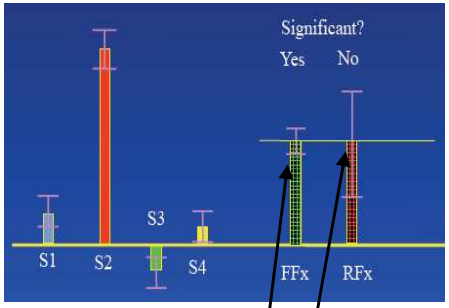
- Degrees of freedom are different in FFX versus RFX
- T(degrees of freedom)= effect size / se
- -> T: the bigger the better
- -> df: the more the better
- -> effect size: the bigger the better
- -> se: the smaller the better

But: it has to make sense!

- **FFX**
- df: N*number of conditions (M)
- Effect size: mean from all subjects
- Variance: divided by N*number of conditions
- SE: Var/N*M

As if all observations come from one subject!

- **RFX**
- df: N-1
- Effect size: mean from all subjects
- Variance: variance across subjects
- SE: Var / N



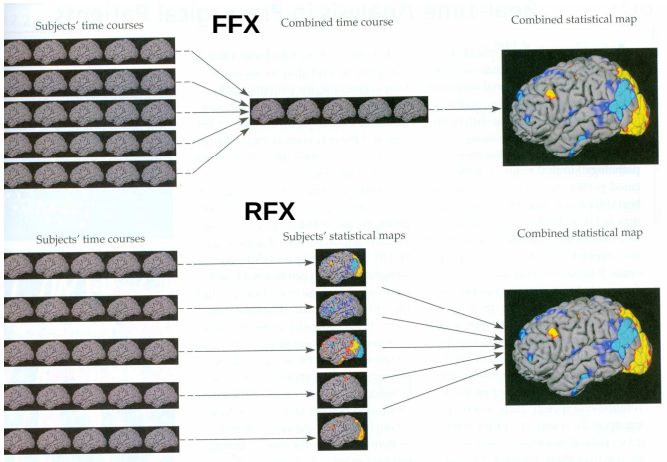
Note: Average is the same!

Figure From D. Greve

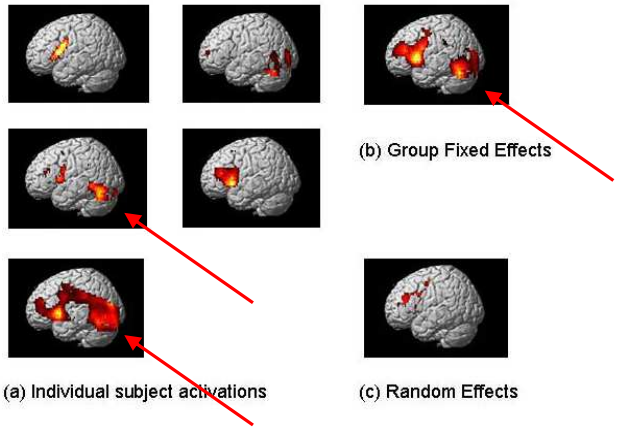
Why FFX / RFX?

- FFX: Is there a difference in the sample?
- RFX: Is there a difference in the sample, *which is consistent across participants* ?

- FFX: Apply results to **sample**
- RFX: Apply results to **population**
- ('draw inference at population level')
- Would a new participant from the same population show the same effect?



From Huetell, Song & McCarthy



From Penny & Holmes, 2003

Finally

- You need more subjects to do RFX. Sometimes, if you have a specific population and can only get few subject, one can do FFX / case study
- Check what your software is doing!

Break

- (after break easier stuff)

T vs. F

- T and F statistic have the same underlying principle:
  - Systematic variance / Unsystematic variance
- Significance of an effect size is dependent upon:
  - The number of measurements ('trust')
  - The effect size ('magnitude')
  - The variance of the effect size ('stability')

T vs. F

– T (dof) = ((diff A-B) – 0) / SE
– (Difference between conditions / variance)
– F (dof, dof) = SSmodel / SSres
– (Explained variance / Unexplained variance)

T

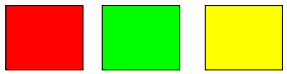
- Which areas are more activated to ‘Red’ compared to ‘Green’?

C1 (red) C2 (green) C3(yellow)

- Red > Green
- Green > Yellow
- Red > Yellow

F

- ‘Is there *an* effect of this factor?’
- E.g. Factor ‘Colour’:

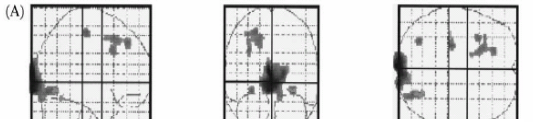


C1 (red) C2 (green) C3(yellow)

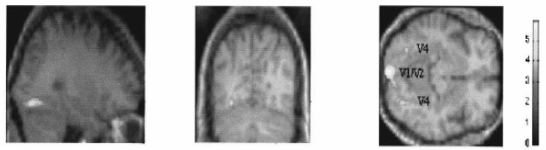
- F tells us whether this factor significantly explains variance in our data

Example

Effect of factor colour:



There is a difference between the three colours



McKeefrey and Zeki, 1997

Post-Hoc testing

- If effect of ‘Colour’, we want to know which colour is responsible for the effect and / or whether area is more sensitive to ‘Red’ as compared to ‘Yellow’

One-sided vs. Two-sided

- One sided: Is A > B?
  - Positive T
- One sided: Is B < A?
  - Negative T
- Two-sided: Is A > B or B > A ?
  - Positive or Negative T

Sensitivity

- With one-sided test you’re statistically *more sensitive*
- One-sided:
  - t(15) = 1.75, p = 0.05
- Two-sided:
  - t (15) = 2.13, p = 0.05
- BUT: maybe you’re missing something interesting

Dependent / Independent

- Independent: measurements come from independent samples
- Dependent: measurements come from same sample
- **Between / within subjects**

Within subjects

- +: effects are not due to differences between subjects
- -: effects of repetition

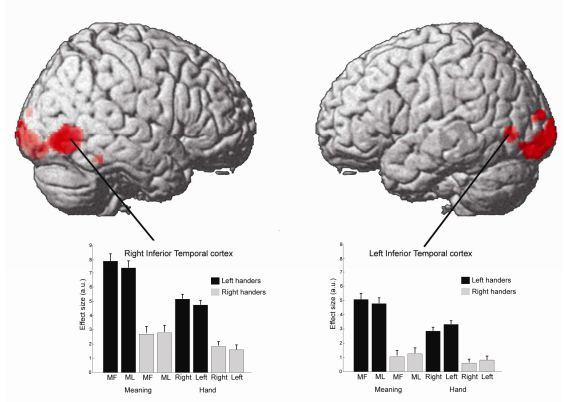
Between subjects

- +: Ability to compare groups on factors you cannot induce experimentally
  - Patient groups
  - Handedness
- -: Hard to control between subject variables that may drive your effect

Example

- 16 left-handers
- 16 right-handers
- Defined with Edinburgh Handedness index
- Controlled for:
  - Gender
  - Age
- Action observation

Example: Main effect of Group



Conclusion

- Interesting explanation:
  - Left-handers perceive actions / body parts from others differently
- Boring (but more parsimonious) explanation:
  - Left-handers have been less often in the scanner (but no global effect)
  - Left-handers were more artistic and watch other people better
  - Left-handers were more social and watch other people better
  - Left-handers were more x
- Correct explanation:
  - We don't know, let's find out in new, better controlled experiment

The new experiment

- **Question**: Do left-handers perceive motion of others / body parts differently from right-handers (because they grow up in a 'right-handed world')

Design

- Can we test this in a within-subjects design?
- Between-subjects: How?
  - Match groups
    - On which factors?
  - Look at differential effects between groups: Group x Condition interaction

Design

- **Question**: Do left-handers *specifically* perceive motion of others / body parts differently from right-handers (because they grow up in a 'right-handed world')
- Do groups differ on this, but not on perception of e.g. houses?

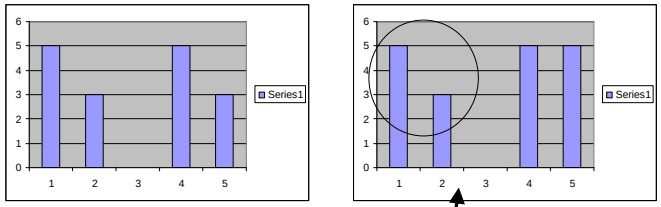
Design

- Factor 1: ‘Group’:
  - left-handers, right-handers
- Factor 2: ‘Stimulus type’:
  - Body parts, houses
- Interaction Group x Stimulus type
- / Use functional localisers

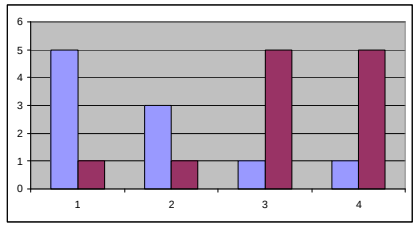
Localisers

- Often used in visual system:
  - Some regions respond to perception of faces, motion, colour etc. (degree of modularity?)
  - Two localisers: body parts, houses
  - Look in these areas whether we see A) a global difference between groups

  B) Specific increase for left-handers to body parts in ‘body-part area’



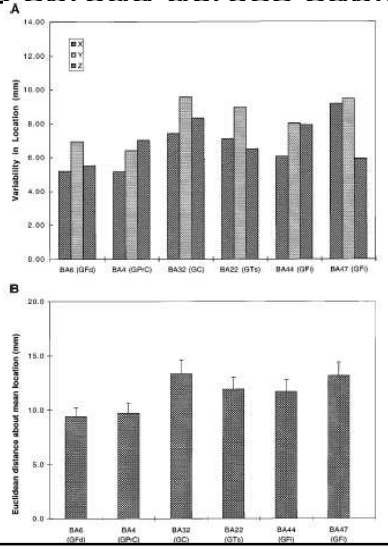
How strong would our conclusion be?



Not all brains are the same

- People's brains differ: shape, size etc.
- → Normalisation
- → Smoothing

What is spread across subjects?



Xiong et al., 2000

Summary

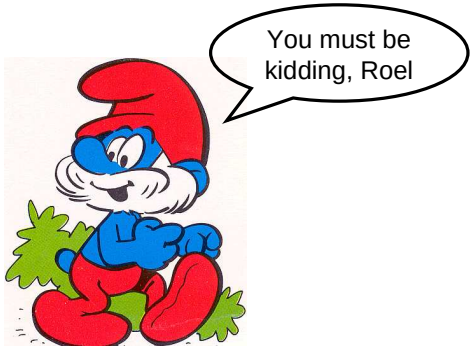
- Population / Sample
- FFX / RFX
  - Degrees of freedom
- Between / Within
- T / F
- Not all brain are the same

Final note

- Stop clicking, start thinking
- BUT: Try around with tests: try FFX, try RFX

  (e.g. check how df change)

  (do a test by hand)



Final note

- Stop clicking, start thinking
- BUT: Try around with tests: try FFX, try RFX
  - (e.g. check how df change)
  - (do a test by hand)
- Check what your software does