C84ANM

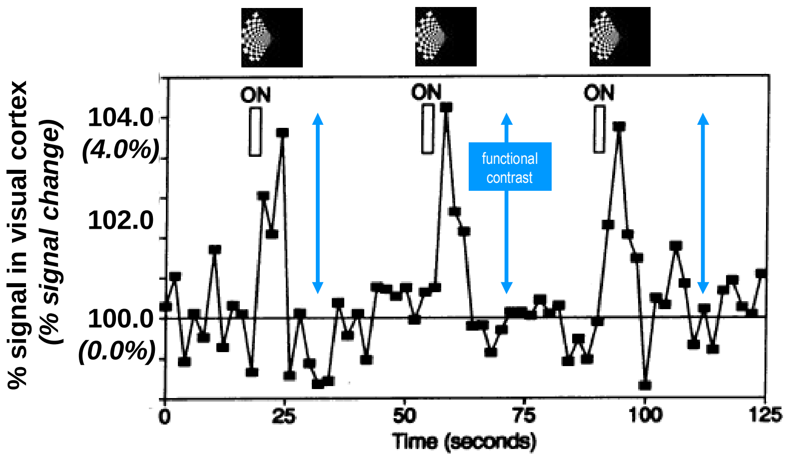
Statistics in neuroimaging and electrophysiology: examples from recent research articles

Martin Schürmann
School of Psychology
University of Nottingham

see links: http://www.psychology.nottingham.ac.uk/staff/mxs/neuroweb/
in particular http://onlinestatbook.com/rvls/ and http://onlinestatbook.com/stat_sim/index.html
free online tools: http://www.davidmlane.com/hyperstat/analysisf.html

Software for this course: SPSS (installed on classroom computers)

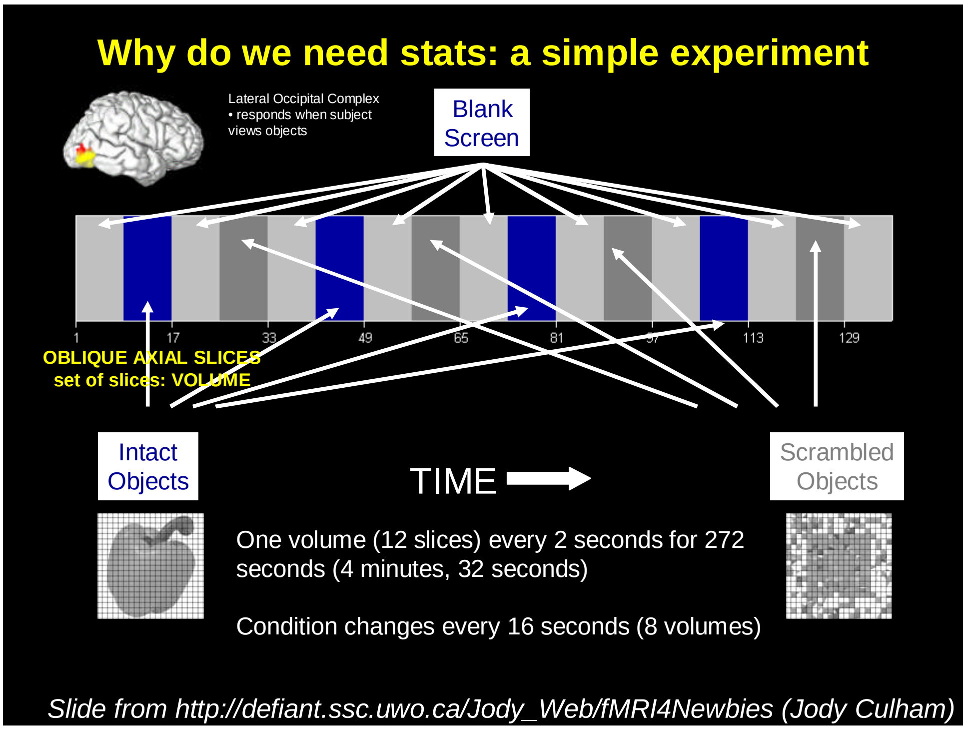
Functional contrast based on BOLD: timecourse

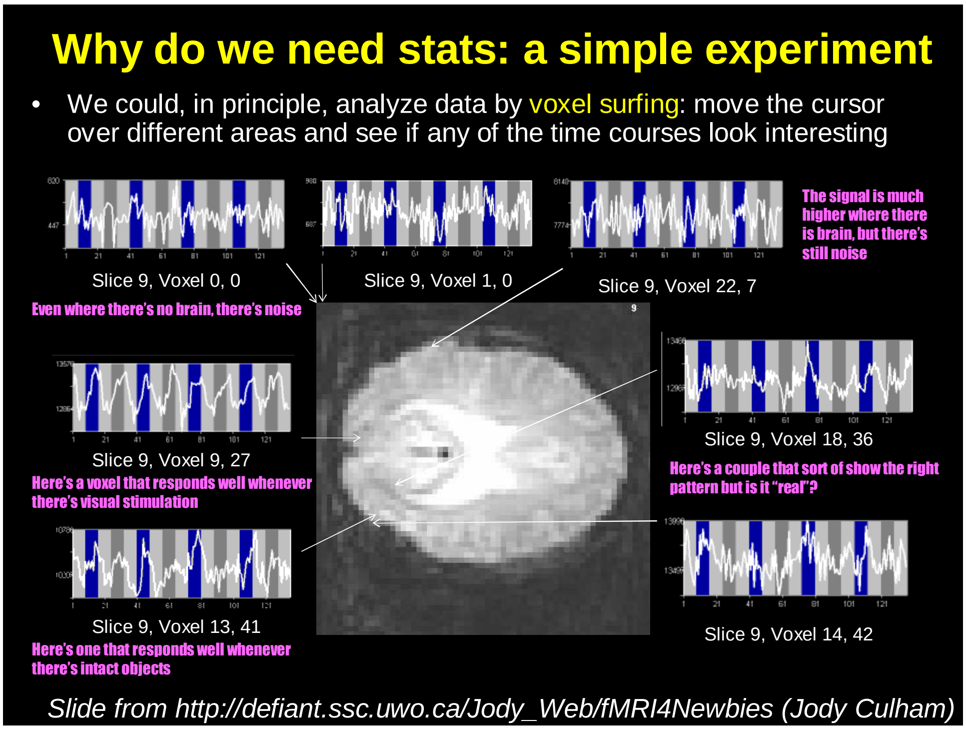


Blamire, Ogawa, Ugurbil et al. PNAS 1992

100 % = summed prestimulation signal intensity over 10 voxels in visual cortex

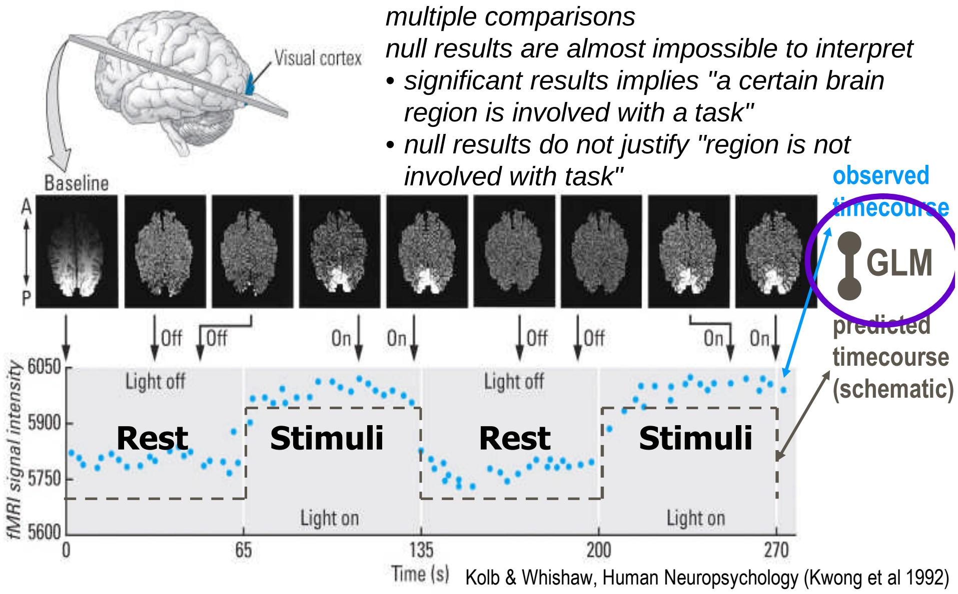
NOTE: statistical results depend on contrast AND noise


Why do we need stats: a simple experiment
Lateral Occipital Complex
• responds when subject views objects
Blank Screen
1
17
33
49
65
81
97
113
129
OBLIQUE AXIAL SLICES
set of slices: VOLUME
Intact Objects
TIME
Scrambled Objects
One volume (12 slices) every 2 seconds for 272 seconds (4 minutes, 32 seconds)
Condition changes every 16 seconds (8 volumes)
Slide from http://defiant.ssc.uwo.ca/Jody_Web/fMRI4Newbies (Jody Culham)


Why do we need stats: a simple experiment
• We could, in principle, analyze data by voxel surfing: move the cursor over different areas and see if any of the time courses look interesting
The signal is much higher where there is brain, but there's still noise
Slice 9, Voxel 0, 0
Slice 9, Voxel 1, 0
Slice 9, Voxel 22, 7
Even where there's no brain, there's noise
Slice 9, Voxel 9, 27
Here's a voxel that responds well whenever there's visual stimulation
Slice 9, Voxel 18, 36
Here's a couple that sort of show the right pattern but is it "real"?
Slice 9, Voxel 13, 41
Here's one that responds well whenever there's intact objects
Slice 9, Voxel 14, 42
Slide from http://defiant.ssc.uwo.ca/Jody_Web/fMRI4Newbies (Jody Culham)

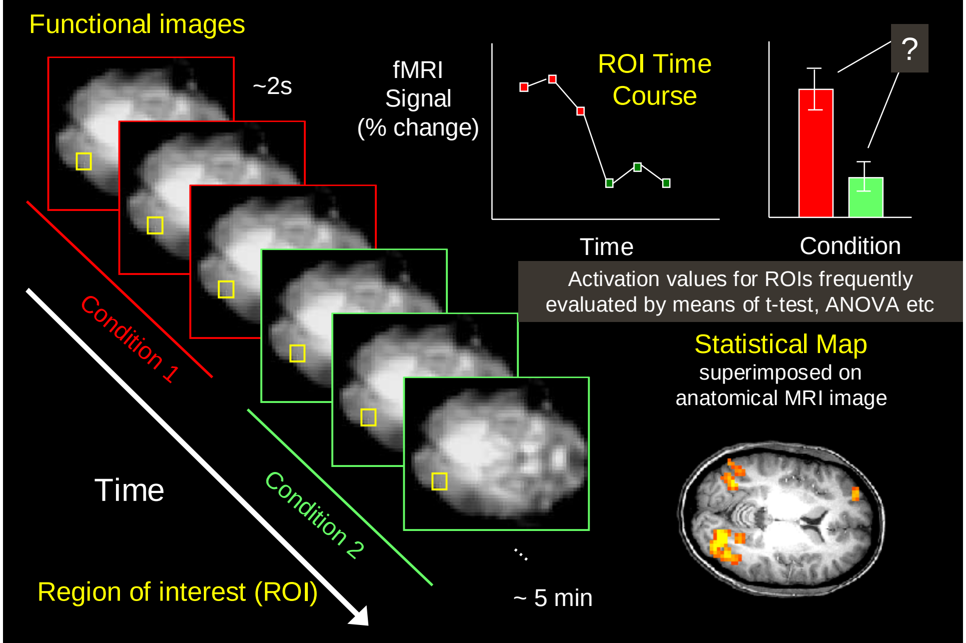
fMRI statistics: GLM for block design

NOTE: problems ...
multiple comparisons
null results are almost impossible to interpret
- *significant results implies "a certain brain region is involved with a task"*
- *null results do not justify "region is not involved with task"*



Kolb & Whishaw, Human Neuropsychology (Kwong et al 1992)

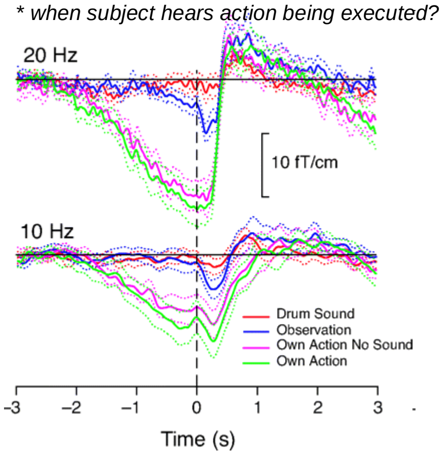
Activation statistics for regions of interest (ROIs)



Slide from http://defiant.ssc.uwo.ca/Jody_Web/fMRI4Newbies (Jody Culham)

From fMRI statistics to MEG, EEG, ERP

Shared representations in the brain
Motor cortex is activated
* when subject executes action
* when subject sees action being executed („mirror neurons")
* *when subject hears action being executed?*



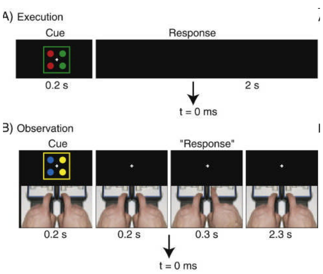
Caetano, Jousmäki, Hari PNAS 2007

13 healthy subjects performing, seeing, or hearing the tapping of a drum membrane with the right index finger. *Parameter of interest: suppression and rebound of 20-Hz activity*

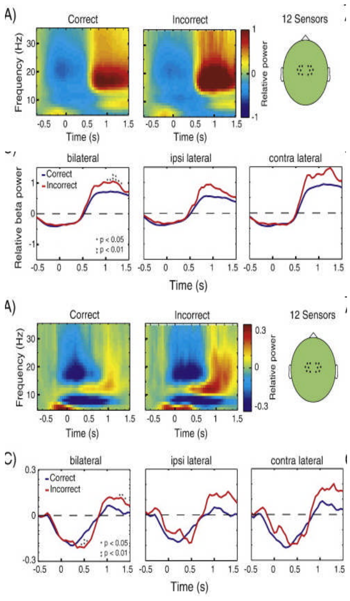
Observation condition: maximum suppression of the 20-Hz activity was only 42 ± 9% of that during *Own Action* **($P < 0.005$, two-tailed paired *t* test, $n = 13$).** The ≈20-Hz rebounds, measured as the mean values from 500 to 900 ms, were, in all four conditions, **statistically significant ($P < 0.05$) above the baseline** (defined as the mean level from −2.9 to −2.4 s), without any systematic differences in the peak amplitudes.
For the ≈10-Hz oscillations, only tiny rebounds were visible. **During *Observation*, the suppression was only 46 ± 16% ($P < 0.05$) of that observed during *Own Action*.** Strikingly, the ≈10-Hz level **returned to the baseline 580 ± 195 ms ($P < 0.05$, $n = 10$) later** during Own Action than during Observation.
NB selection of baseline

Categorical comparisons: MEG, EEG, ERP examples

Shared representations in the brain
Motor cortex is activated
* when subject executes action
* when subject sees action being executed („mirror neurons")
* *when subject observes a wrong vs correct action being executed?*



Koelewijn, van Schie, Bekkering, Oostenveld, Jensen: Neuroimage 2008



Action execution. (A) Grand average time–frequency power spectra time-locked to the button press (t = 0 s). (C) Time-resolved beta power averaged over 12 sensors. Significant difference (correct vs incorrect executed actions) at 1.05–1.3 s.

A repeated measurements ANOVA on mean power over a single time period of 0.8–1.3 s with laterality (ipsi or contra) and correctness (correct or incorrect) as factors showed a main effect for laterality ($F(1,11) = 14.835$, $p < 0.005$), a strong trend for correctness ($F(1,11) = 3.530$, $p = 0.087$), and no interaction effect ($F < 1$).

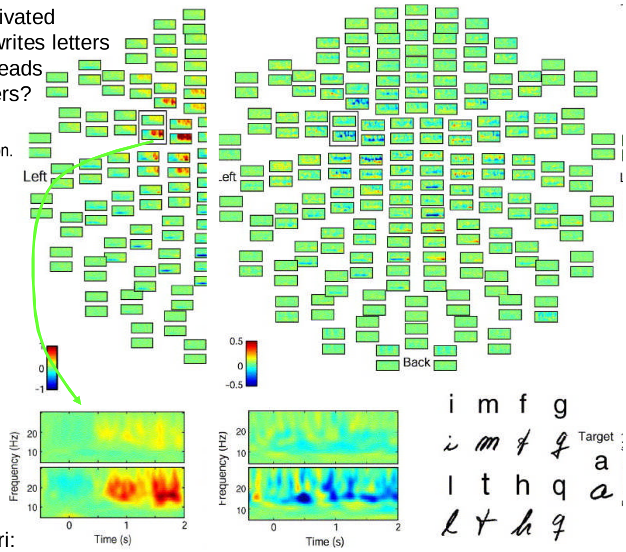
Action observation. (A) Grand average time–frequency power spectra time-locked to the button press (t = 0 s). (C) Time-resolved beta power averaged over 12 sensors. Significant difference (correct vs incorrect executed actions) during beta desynchronization and during rebound

A repeated measurements ANOVA over a time period of 1–1.5 s with laterality (ipsi or contra) and correctness (correct or incorrect) as factors revealed a significant laterality ($F(1,11) = 4.860$, $p = 0.050$), a significant effect for correctness ($F(1,11) = 6.846$, $p < 0.05$), and no interaction effect ($F < 1$).

Categorical comparisons: MEG, EEG, ERP examples

Shared representations in the brain
Motor cortex activated
* when subject writes letters
* when subject reads handwritten letters?

Right finger lift condition. TFRs wrt baseline − 500 to − 300 ms

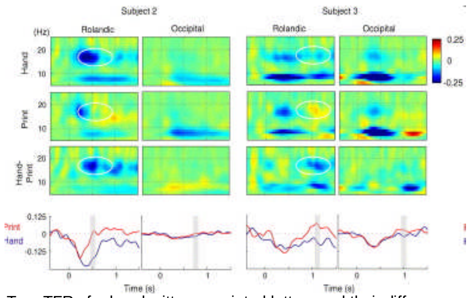


TFRs (handwr'n letters) – TFRs (printed letters). Time 0: onset of visual stimulus. Baseline: − 400 to − 20 ms.

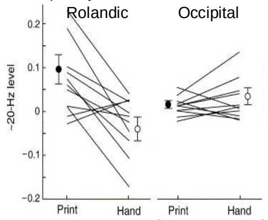
Data for location, frequency and timing need to be reduced to obtain statistically testable values. Therefore: Focused analysis on primary motor cortex using the results of the finger lift task to define -- for each individual -- sensors and frequency bands of interest.

Longcamp, Tanskanen, Hari: Neuroimage 2006

Categorical comparisons: MEG, EEG, ERP examples



Top: TFRs for handwritten vs printed letters and their difference. Bottom: Level of the ~20-Hz activity in the two visual conditions (printed letters: red, handwritten letters: blue). Shaded areas: maximum difference between the 2 conditions; these values were subsequently used for the ANOVA.



~20-Hz level at the time of maximum difference as a function of condition and region. Individual values as lines and mean ± SEM as circles with error bars

Longcamp, Tanskanen, Hari: Neuroimage 2006

Rolandic region **~ 20-Hz level** higher for printed letters in 9 subjects out of 11 ($p = 0.065$, sign test); occipital region: only in 4 subjects out of 11 (n.s.).

ANOVA: main effect of condition ($F(1,10) = 6.89$; $p < 0.026$); interaction between condition and region ($F(1,10) = 13.76$; $p < 0.005$), with no main effect of area ($F(1,10) = 0.01$, n.s.).

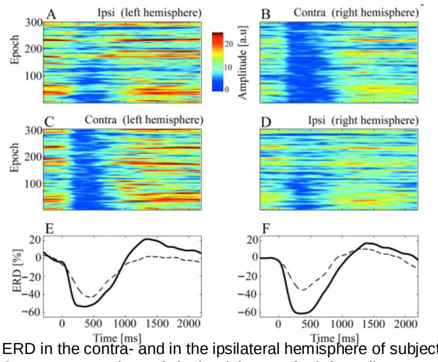
In the rolandic region, the level was positive for the printed letters and negative for the handwritten letters (planned comparison $F(1,10) = 12.08$; $p < 0.006$).

Results due to the time window selection? ANOVA for ~ 20-Hz level averaged over the whole (0–1500 ms) time window: main effect of condition ($F(1,10) = 5.84$; $p < 0.04$), same interaction between condition and region ($F(1,10) = 10.13$; $p < 0.01$), and regional differences (planned comparisons $F(1,10) = 8.53$ $p < 0.02$ for the rolandic region and $F(1,10) = 1.64$, n.s. for the occipital region).

Control ANOVA on **~ 10-Hz levels** at the time of maximum difference between the two conditions failed to show any statistically significant main effect (condition: $F(1,10) = 1.92$, n.s.; region: $F(1,10) = 1.77$, n.s.) nor interaction between the two factors ($F(1,10) = 2.24$, n.s.). This result was not due to the selection of the time window because also the analysis of the ~ 10-Hz level averaged over the whole (0–1500 ms) time window revealed neither main effects (region: $F(1,10) = 0.32$, n.s.; condition: $F(1,10) = 0.48$, n.s.) nor interaction between the two factors ($F(1,10) = 2.87$, n.s.).

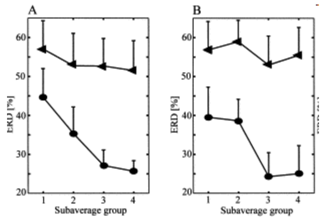
Categorical comparisons: MEG, EEG, ERP examples

Compare contra- vs ipsilateral reactivity of somatosensory cortex to median nerve stimuli: lateralization of ERD in MEG, 6 subjects



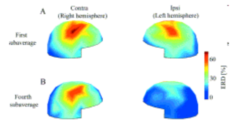
ERD in the contra- and in the ipsilateral hemisphere of subject S5. Representations of single-trial ERDs for left median nerve stimulation (A,B), and for right median nerve stimulation (C,D). The corresponding averaged waveforms are presented in (E,F). Solid and dashed lines show ERD for the contra- and for the ipsilateral hemisphere, respectively. The amplitude of the single trials is color-coded and expressed in arbitrary units.

Nikouline, Linkenkaer-Hansen, Wikström, Kesäniemi, Antonova, Ilmoniemi, Huttunen, Neurosci Lett 2000



ERD values of four consecutive subaverage groups in the course of the experiment. (A) Stimulation of the left median nerve. (B) Stimulation of the right median nerve. Ellipsoids and triangles denote ipsi- and contralateral hemispheres, respectively.

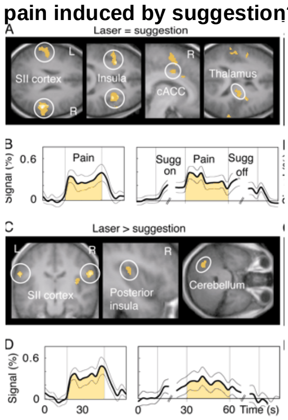
Ipsilat' hemispheres: Friedman tests, $p < 0.001$ (non-parametric replacement for one-way ANOVA)



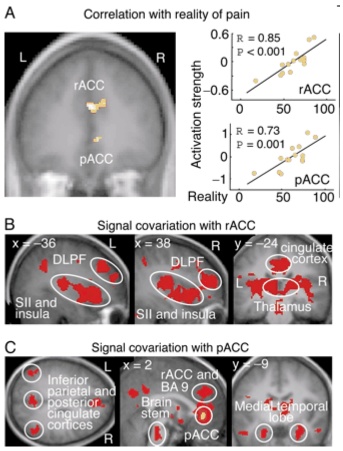
Spatial distribution of ERD in subject S2. ERD to the left median nerve stimulation for the first (A) and fourth (B) subaverage group.

Beyond categorical comparisons: correlational analysis

Brain correlates of subjective reality of pain induced by suggestion?



(A, B) Brain regions activated during both suggestion- and laser-induced pain and mean (± SEM across subjects) time courses of signal changes. Only the first three and the last three time points of the suggestion periods are illustrated, because time to pain appearance and to pain relief varied among subjects. (C, D) Areas where activation was stronger during laser- than suggestion-induced pain and related signal time courses.

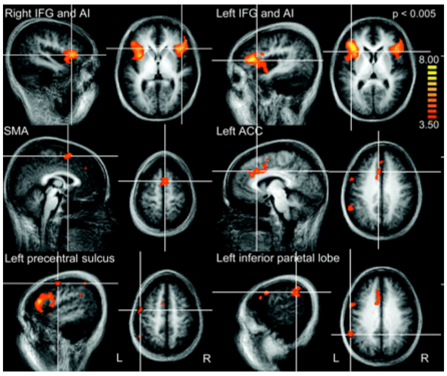


(A) Regions where activation correlated with the subjective estimate of reality of pain (B) Brain areas where fMRI signal covaried with that of rACC. (C) Brain areas where fMRI signal covaried with that of pACC.

Raij, Numminen, Närvänen, Hiltunen, Hari, PNAS 2005

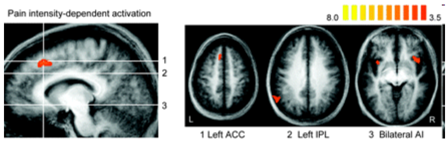
Beyond categorical comparisons: correlational analysis

The compassionate brain: Regions activated when observing another person who is in pain?

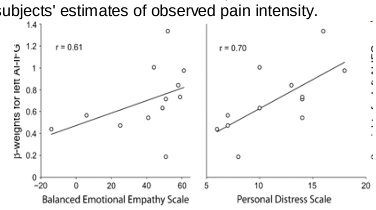


Statistical maps in the provoked pain – chronic pain contrast (threshold: P < 0.005). Color bar indicates t-value.

Saarela, Hlushchuk, Williams, Schürmann, Kelso, Hari, Cereb Cortex 2007



Brain areas where the BOLD response correlated with the subjects' estimates of observed pain intensity.



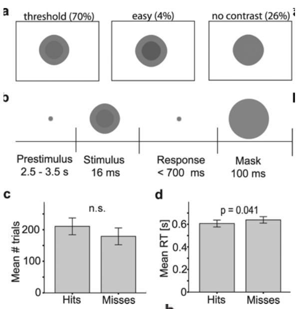
Activation strengths as beta weights vs. individual scores in personal distress scale and BEES. Lines represent the linear best fits. The correlation coefficients for both scales differed from zero at P < 0.05.

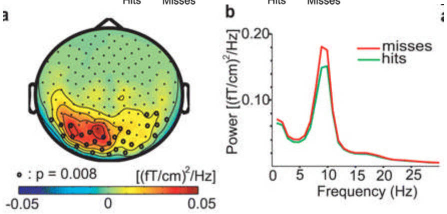
•Balanced Emotional Empathy Scale (BEES) assesses the capacity to experience another's emotions (including pain).

* Interpersonal Reactivity Index (IRI) with subscales: "empathic concern" relating to others' hardships, "fantasy" taking the place of characters in books or movies, "perspective taking" in relation to interpersonal conflicts,"personal distress" concerning the subjects' tendency to become anxious when witnessing others' suffering or need for help.

Advanced statístics: randomization methods for MEG

Relationship between prestimulus alpha power (as indicator of „brain state", Basar 1980) and responses to stimuli (reaction time, hit rates)?





Compare two sets of stimuli, hits vs misses

Problem:multiple comparison over 151 MEG channels

Solution: „nonparametric randomization method identifying clusters of sensors with significant changes corrects for multiple comparisons over sensors, in within subject comparisons

- Clusters are defined as spatially contiguous sensors where the t statistics exceed a previous threshold ($p < 0.05$) with respect to two conditions.
- The cluster-level test statistic was defined as the sum of the t statistics of the sensors in a cluster.
- In a nonparametric statistical test, the type-I error rate for the complete set of 151 sensors was controlled by evaluating the cluster-level test statistic under the randomization null distribution of the maximum cluster-level test statistic.
- This was obtained by **randomly permuting** the data between the two experimental conditions within every participant.
- By creating a reference distribution from 1000 random sets of permutations, the p value was estimated as the proportion of the elements in the randomization null distribution exceeding the observed maximum cluster-level test statistic."

van Dijk, Schoffelen, Oostenveld, Jensen: J Neurosci 2008