

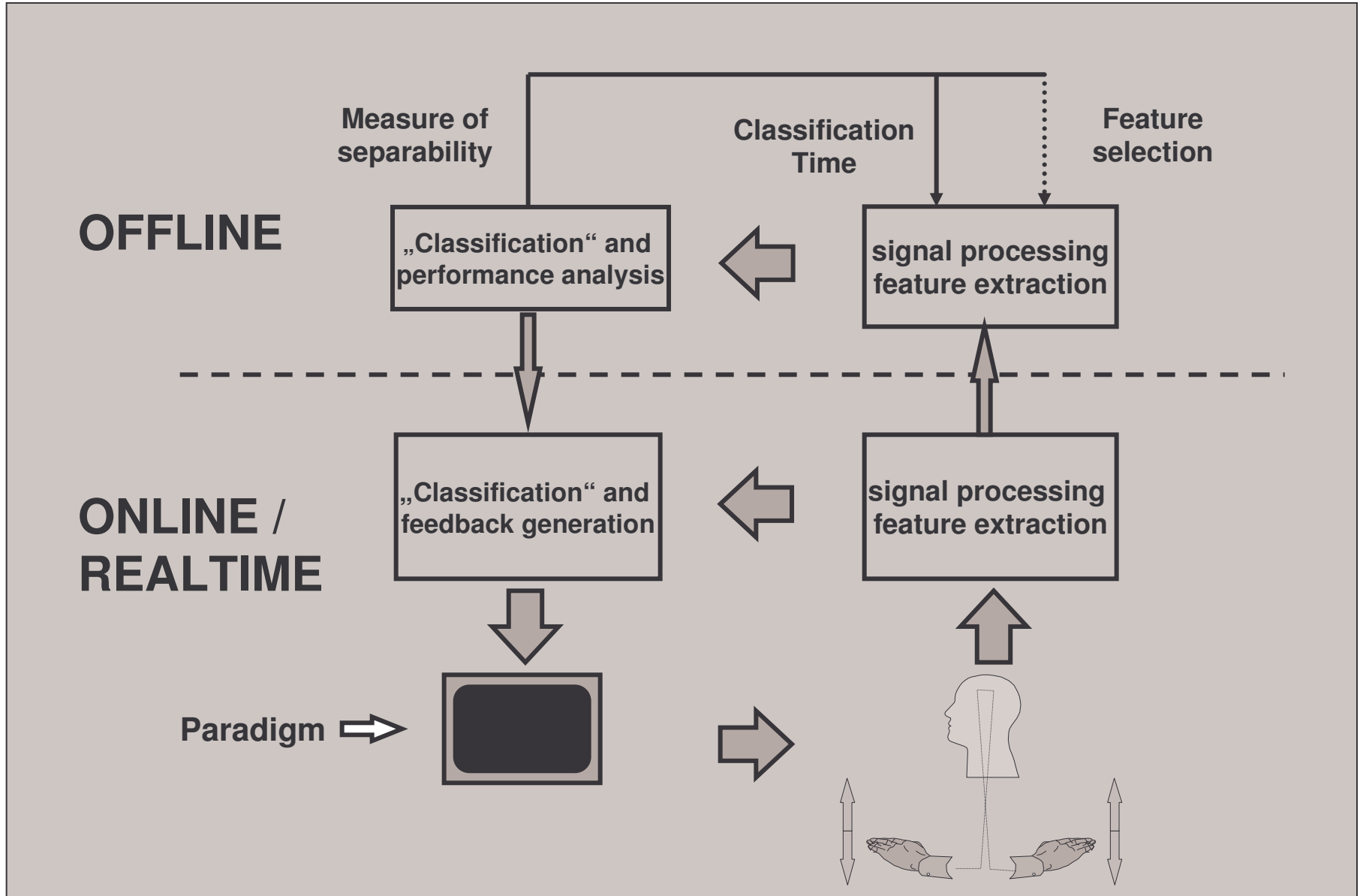
Evaluation criteria
or
Quantifying the information content
of the BCI feedback

BCI 2005 meeting
Rensselaerville, NY, USA

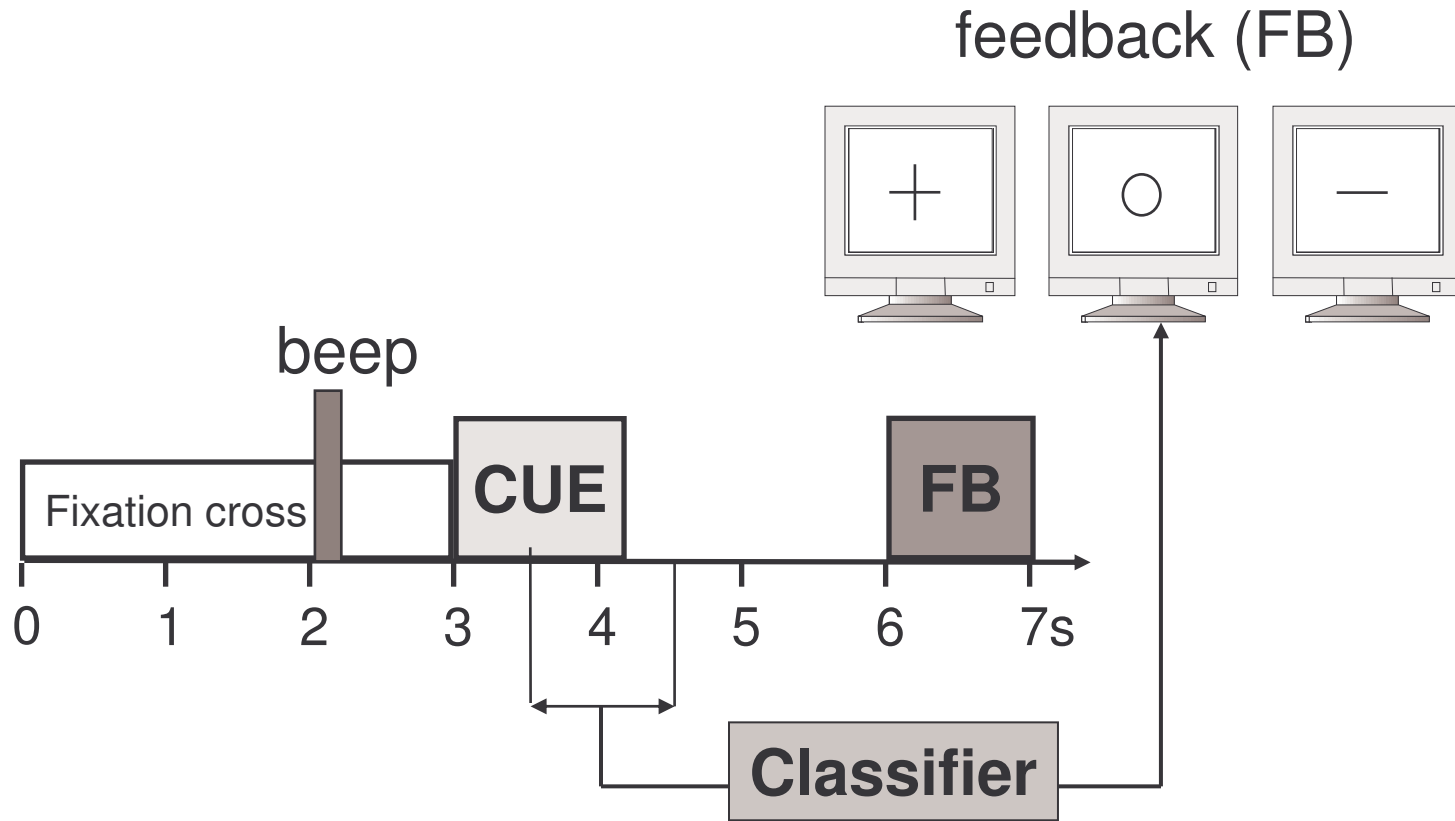


Alois Schlögl
<alois.schloegl@tugraz.at>

Scheme of a BCI



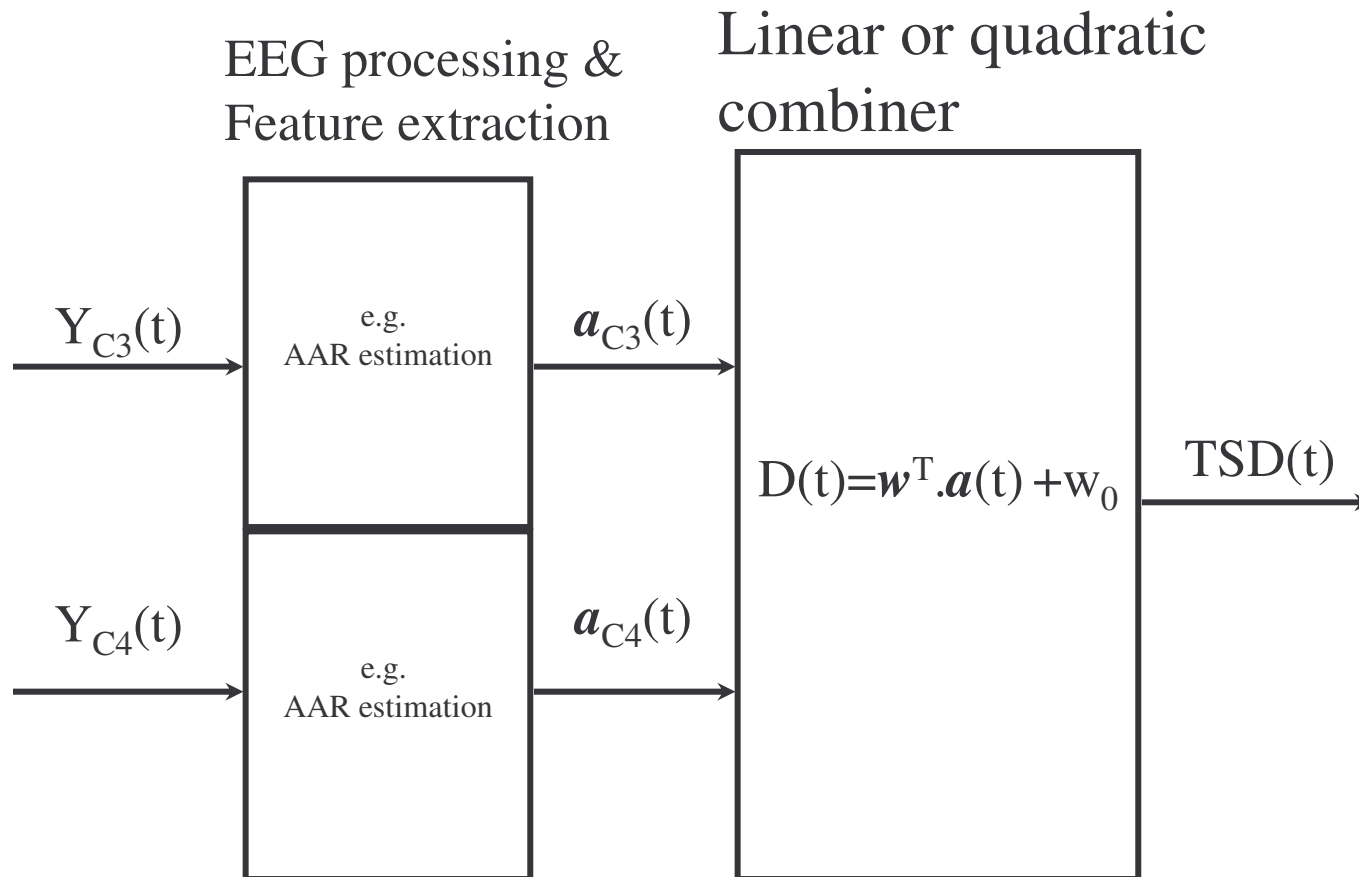
Paradigm with discrete feedback



band power and LVQ

Pfurtscheller et al., 1997

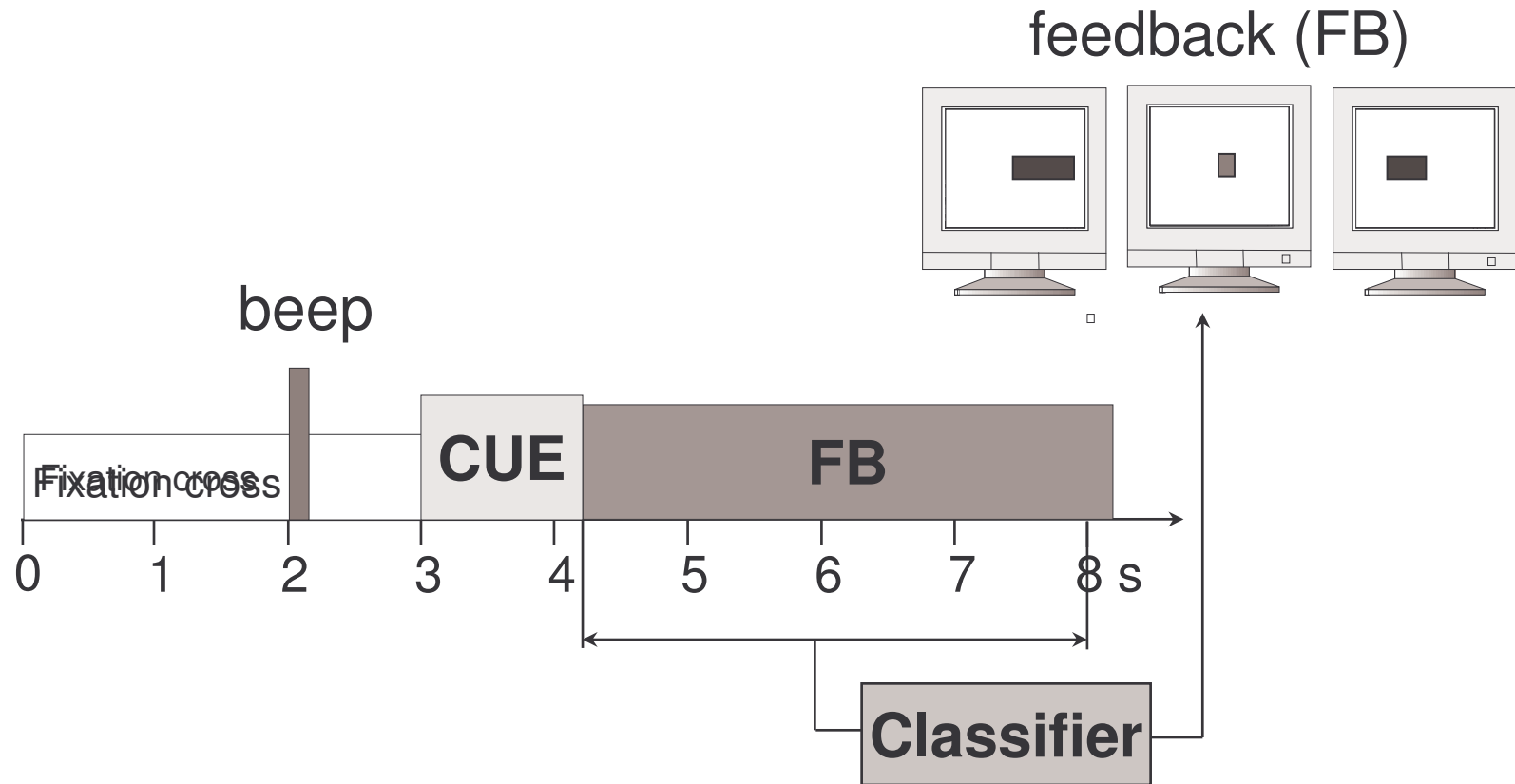
Generate feedback



Static classifier: Schlögl et al. 1997

Adaptive Classifier: Vidaurre et al. 2004

Paradigm with continuous (in time and magnitude) feedback



AAR parameters and LDA

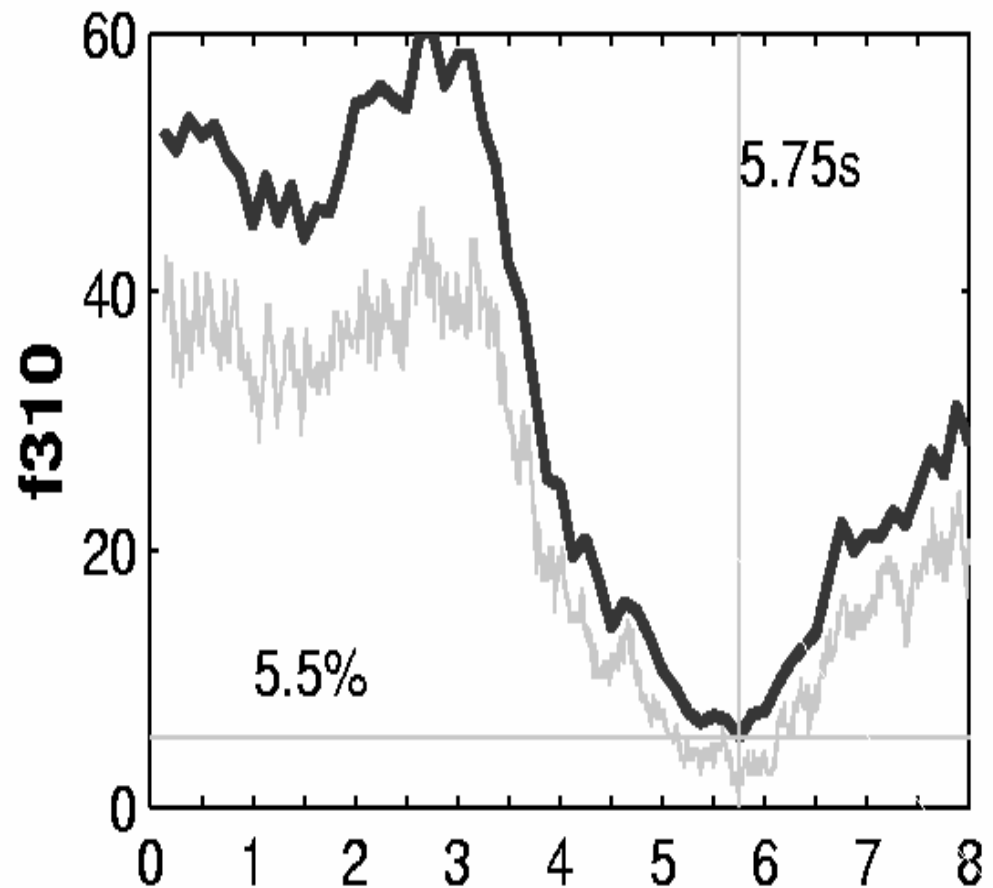
Schlögl et al. 1997, Neuper et al., 1999

Time course of the error rate

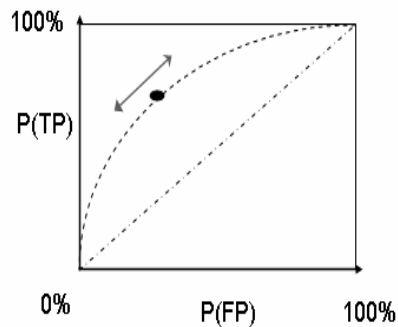
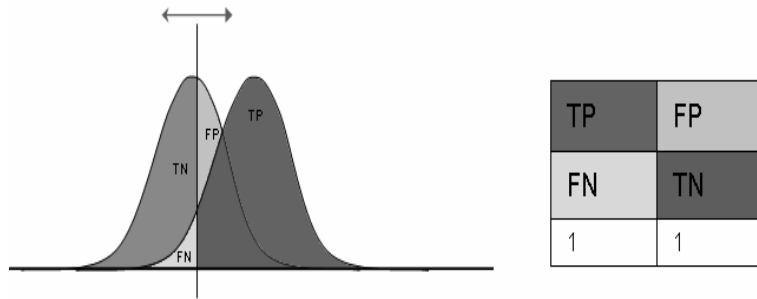
with crossvalidation
(red)

w/o crossvalidation
(cyan)

BCI4c, F3-10, RLS, $p=10$,
UC=0.006,
C3+C4, 80L+80R



SNR, ROC, Confusion matrix



- True positives TP
 - False positives FP
 - True negatives TN
 - False negatives FN
-
- Area under the curve (AUC)

Signal-to-Noise ratio (SNR) and Mutual Information

$$\text{SNR} = \text{signal/noise} = s/n = (s+n)/n - 1$$

$$\text{SNR} = (\mu_1 - \mu_2)^2 / 2(\sigma_2^2 + \sigma_1^2)$$

$$\text{MI [bit]} = \frac{1}{2} * \log_2(\text{SNR} + 1)$$

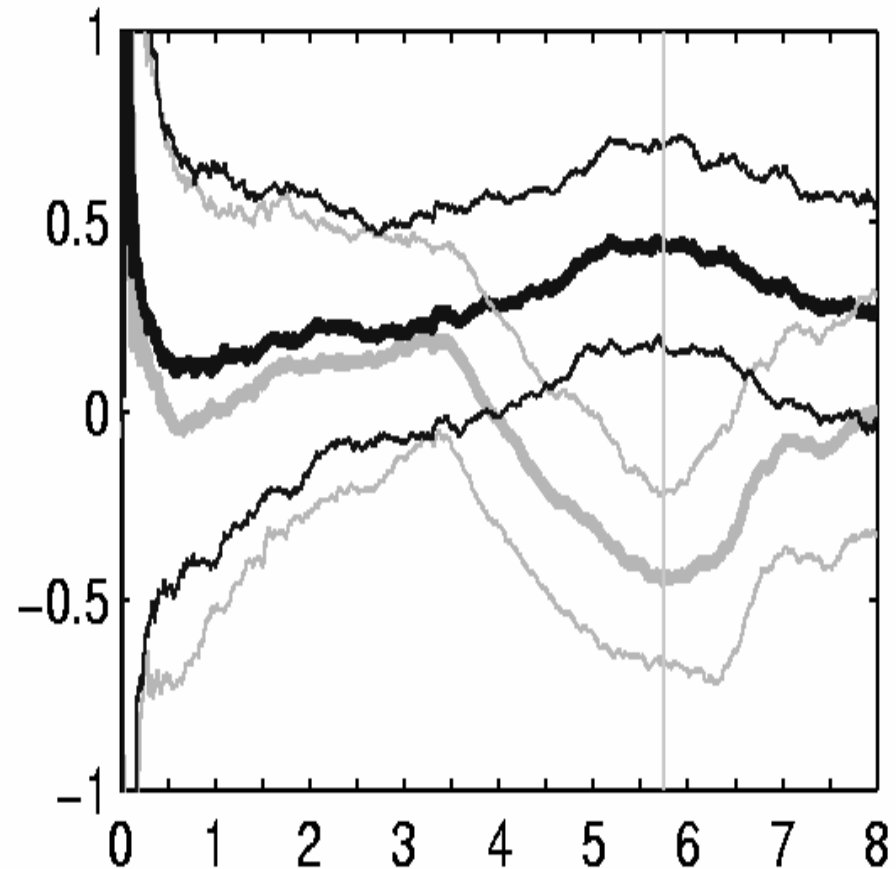
Inter-trial variance and between class differences

Distance d

$$d(t) = \sum_i a_i(t) * w_i + w_0$$

is continuous in time and
in magnitude

BCI4c, F3-10, RLS,
 $p=10$, $UC=0.006$,
C3+C4, 80L+80R



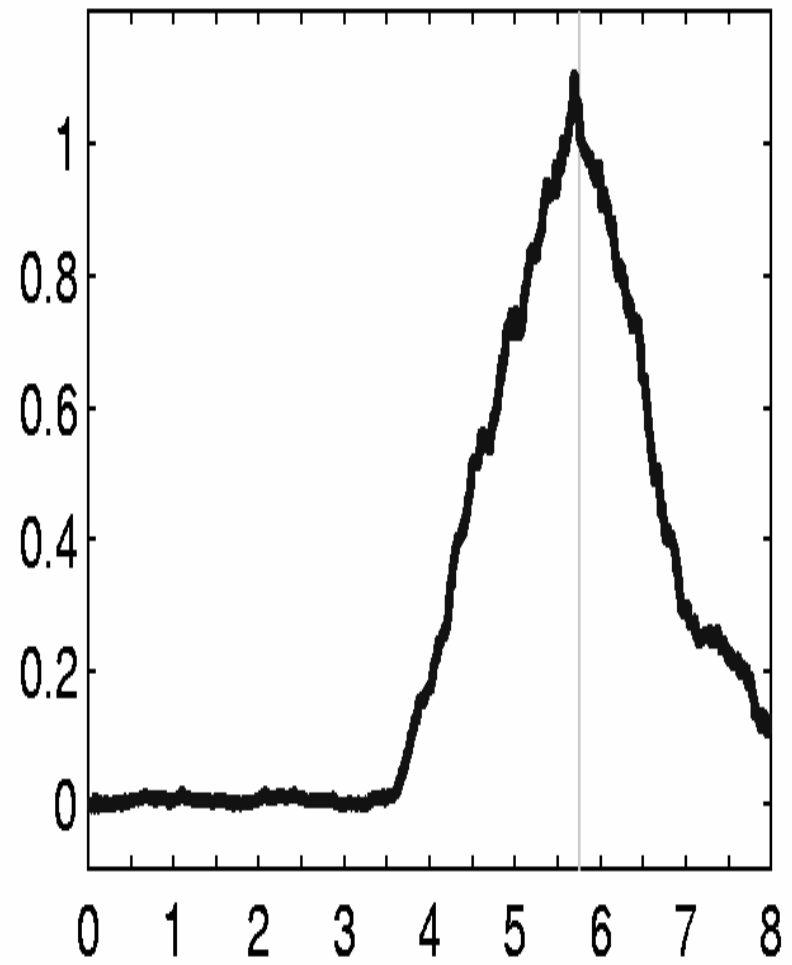
Mutual Information = Transfer information

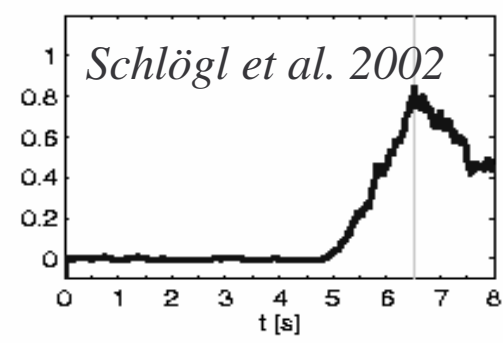
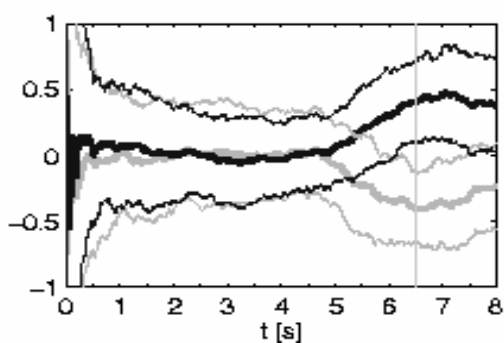
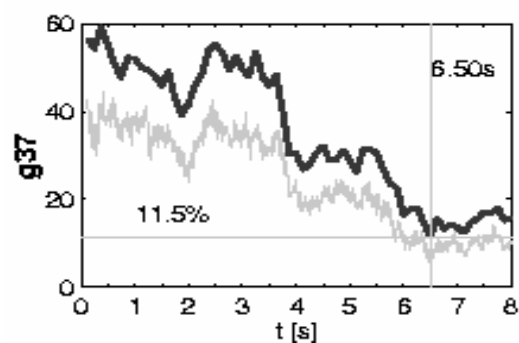
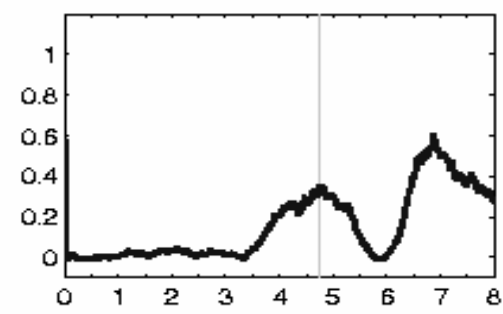
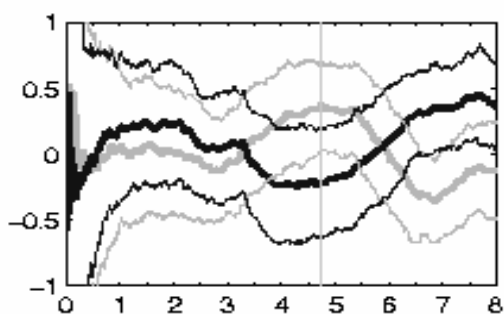
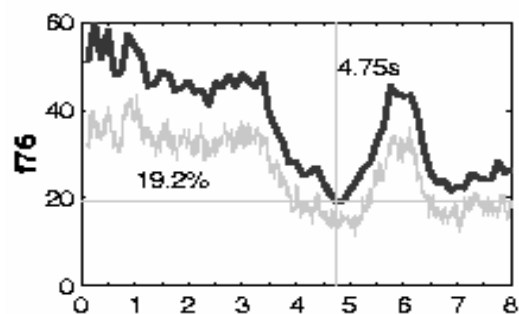
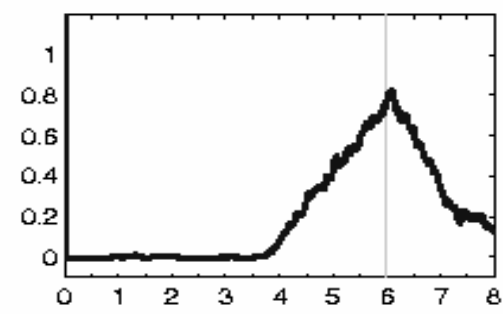
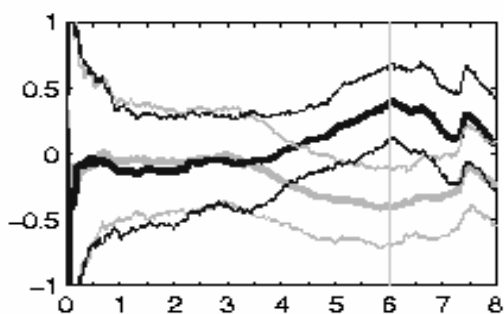
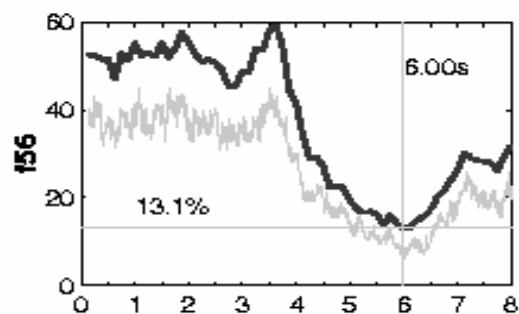
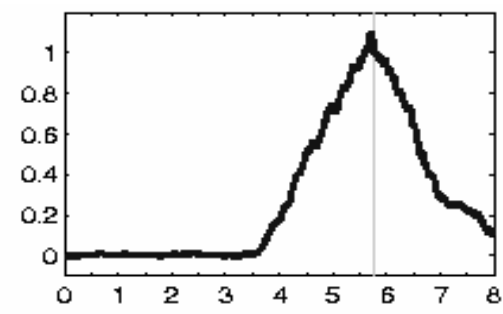
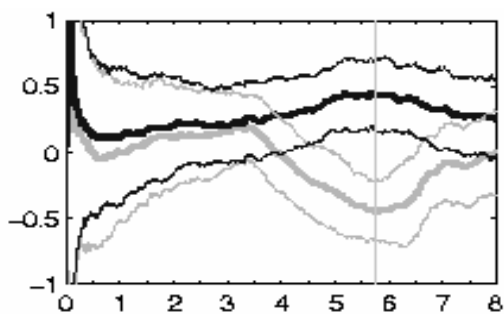
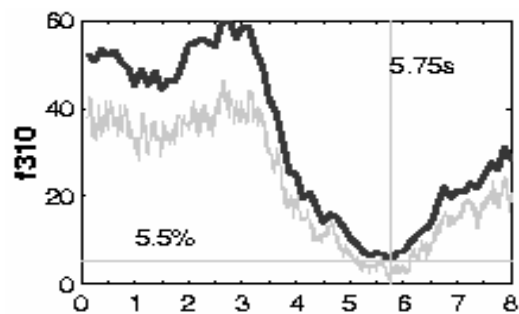
SNR =

Entropy [bit]

$$I_t = \frac{1}{2} \log_2 (S_t/N_t + 1)$$

BCI4c, F3-10, RLS, p=10,
UC=0.006,
C3+C4, 80L+80R



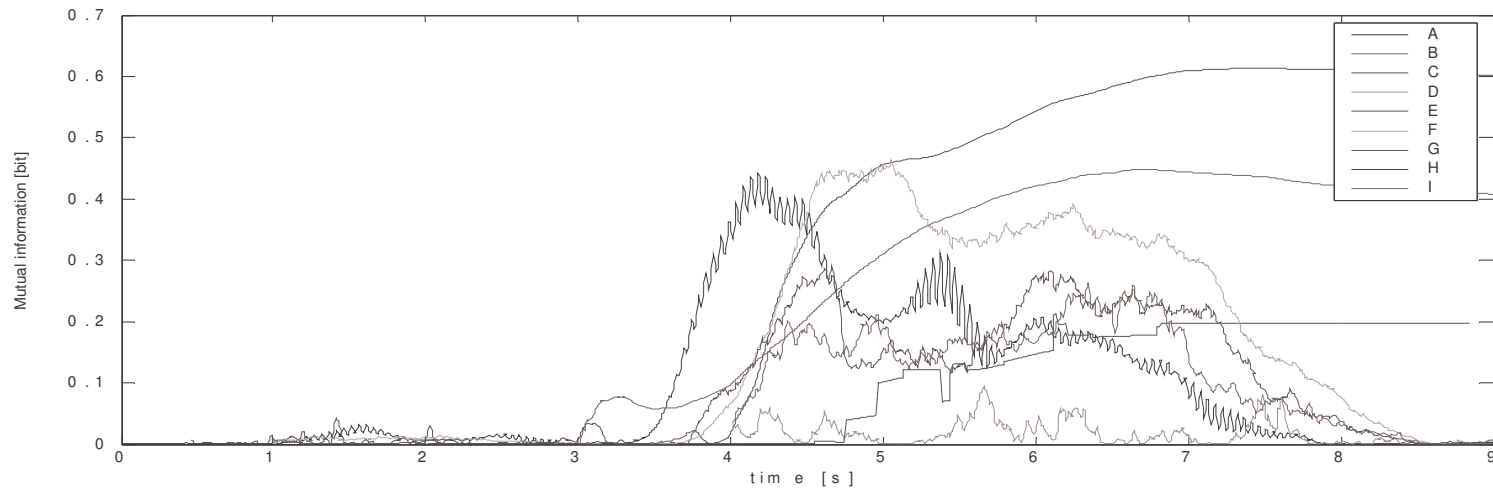
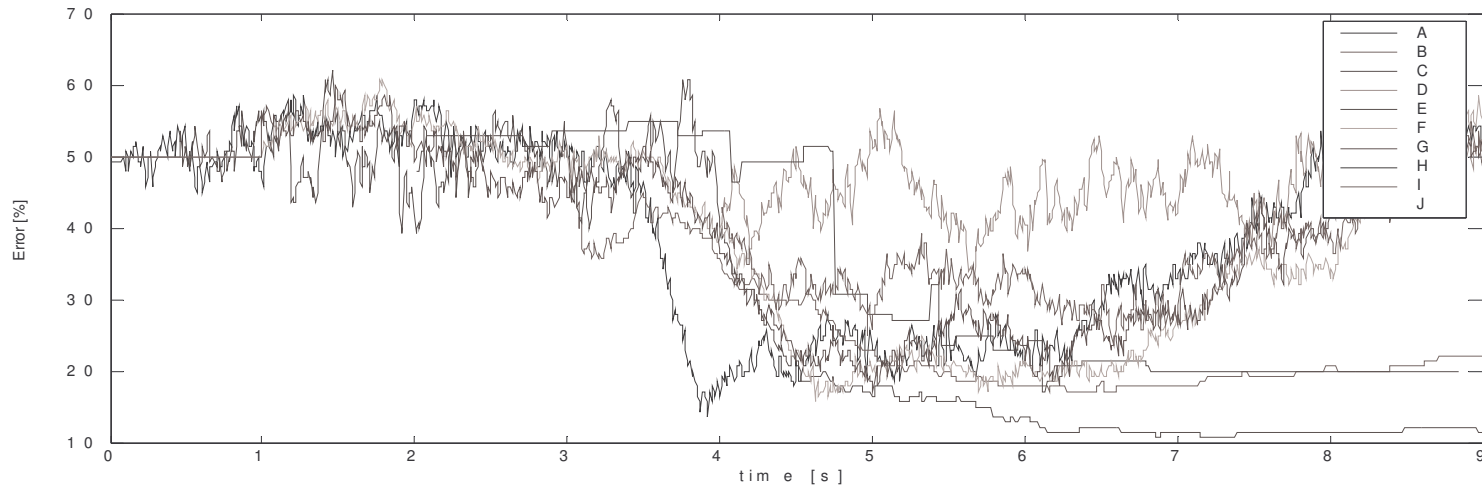


Schlögl et al. 2002

Question

- Which method should be implemented online?
- Which criteria to use? ERR[%] or MI[bit] or something else?

Error rate & Mut. Information



Steepness of mutual information

- „speed“ (i.e. time delay) combined with classification accuracy
- Evaluation of signal processing (feature extraction and classification)

ERR[%] vs. MI [bit]

- ERR% uses only the SIGN and ignores magnitude
- Error rate is limited by 0%,
- Err% close to 0% resolution is limited by N
- MI takes into account the information of the magnitude
- MI no upper limit,
- can indicate large SNR

\Rightarrow MI

Cohen's Kappa coefficient

confusion Matrix H of size MxM for M classes

N samples

n_{oi} , n_{io} : sum of rows and columns

p_e chance expected agreement

se(kappa) standard error, confidence interval

Useful for more than 2 classes

BCI competition 2005, data set IIIa

Schlögl et al. (accepted)

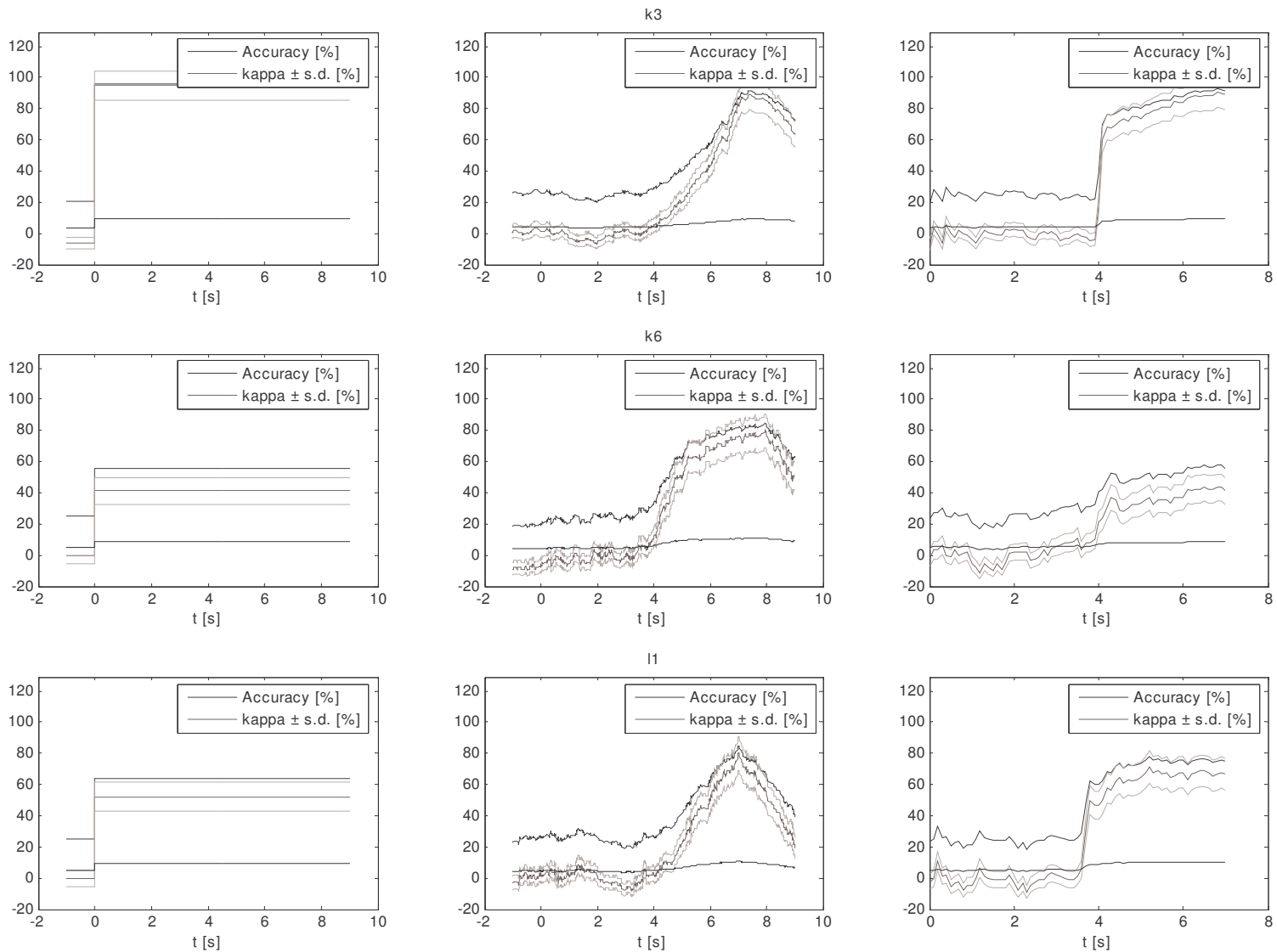
$$ACC = p_0 = \frac{\sum_i H_{ii}}{N}$$

$$p_e = \frac{\sum_i n_{oi} \times n_{io}}{N \times N}$$

$$\kappa = \frac{p_0 - p_e}{1 - p_e}$$

$$se(\kappa) = \frac{\sqrt{p_0 + p_e^2 - \sum_i [n_{oi} \times n_{io} \times (n_{oi} + n_{io})] / N^3}}{(1 - p_e)\sqrt{N}}$$

ACC [%], KAPPA \pm SE



BCI competition 2005, data set IIIa

Summary of
Signal processing & classification
Workshop B
Evaluation criteria

BCI 2005 meeting
Rensellarville, NY, USA

Evaluation criteria

- Classification accuracy (or Error rate)
- Receiver-operator-characteristics curve (ROC)
- Signal-to-noise ratio (SNR)
- Mutual information and Information transfer
- Steepness of the mutual information
- Cohen's Kappa coefficient
- Correlation coefficient R^2
- Letter's per minute
- HF-difference

Questions

- Which criteria to use for evaluating feature extraction and classifiers ?
- Which criteria should be used in Online BCI experiments?

- ERR%/Accuracy, SNR, MI, Steepness of MI, Kappa (Alois)
- R², ITR, Letter per Minute, Duration until successful (Dennis McFarland)
- HF-Difference (Jane Huggins)
- ROC, AUC (Jeremy Hill)

Summary (1)

Justification for criteria of the BCI competition

ERR rate (Accuracy)

Kappa

„Steepness of the mutual information“

Summary (2)

- Agreement that Shannon's Communication theory should be applied
- Different models of the communication channel: „Is the brain digital or analog?“

Summary (3)

- Recommended evaluation of asynchronous BCI
- + Confusion Matrix (TP,FP,TN,FN) and „rebounding time“
- Other proposed criteria are not recommended

Summary (4)

- Asking for software repository for the different evaluation criteria
- BIOSIG provides a starting point
<http://biosig.sf.net>