Donders Institute
for Brain, Cognition and Behaviour

New methods and application domains for stimulus driven BCI's

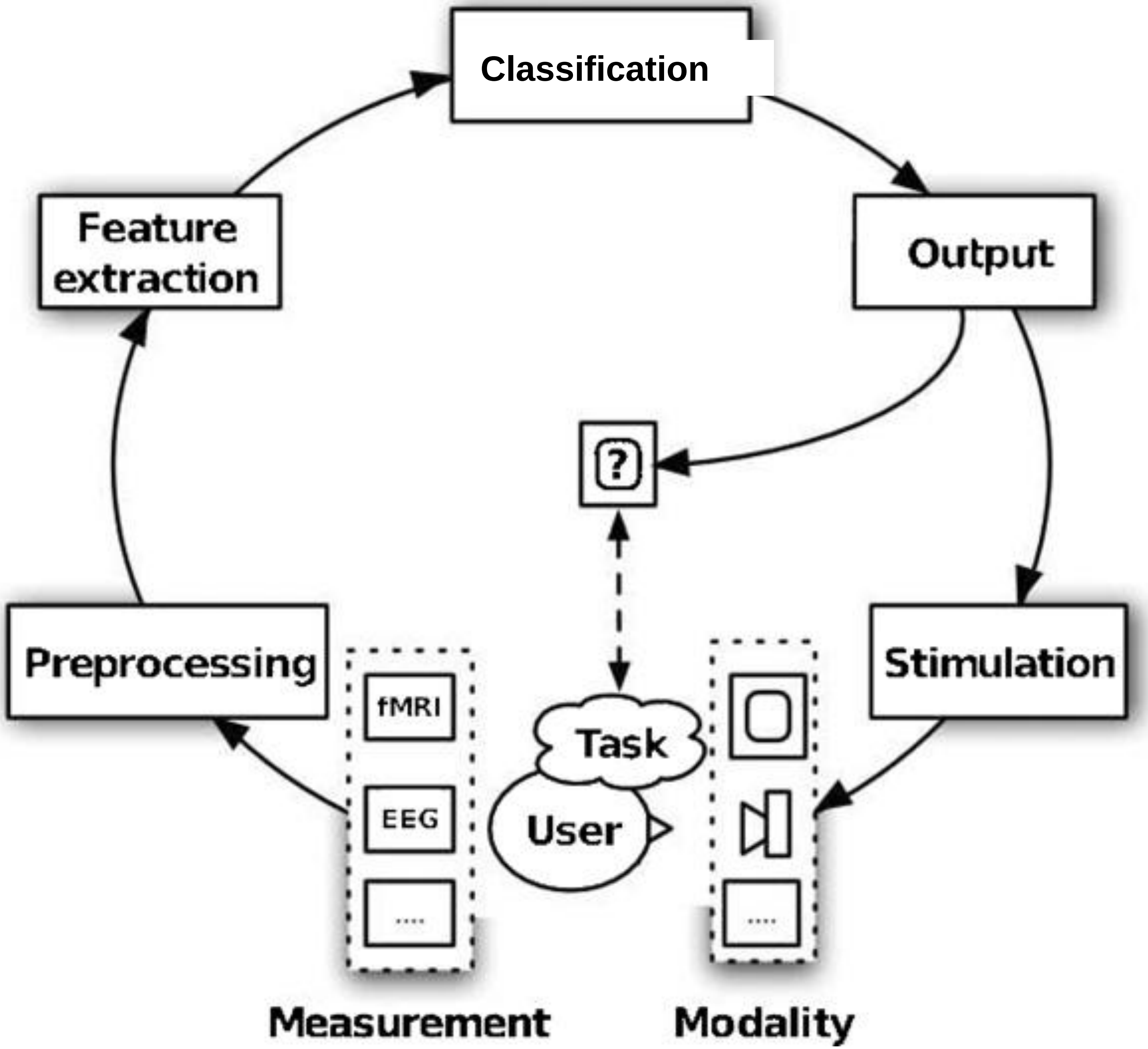
Galin Bajlekov
Alex Brandmeier
Jason Farquhar
Marcel van Gerven
Jeroen Geuze
Christian Hoffman
James McQueen
Makiko Sadakata
Loukianos Spyrou

Peter Desain

Radboud University Nijmegen

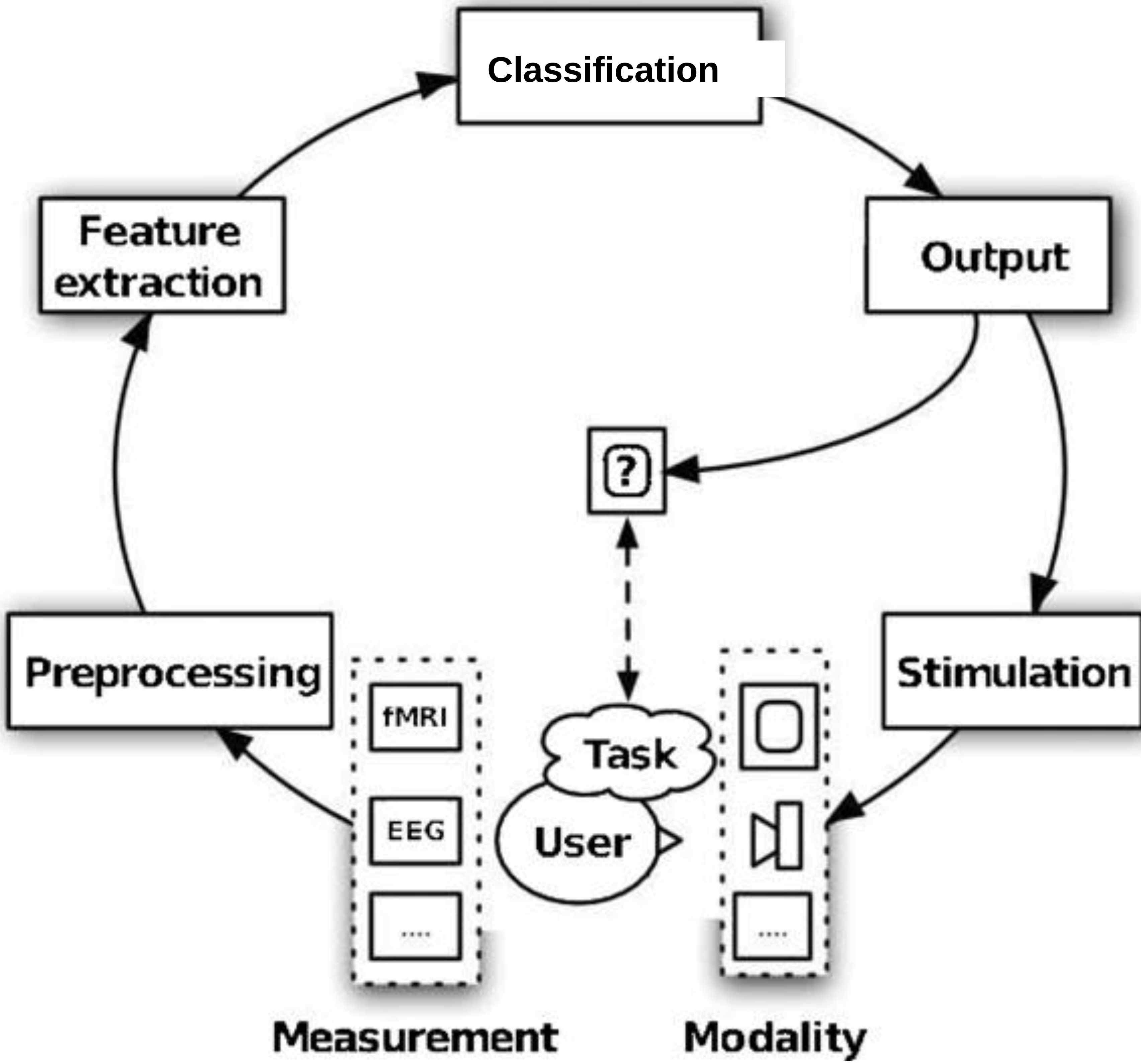
Contents

BCI Cycle


Classification
Feature extraction
Output
?
Preprocessing
Stimulation
fMRI
EEG
....
Task
User
....
Measurement
Modality

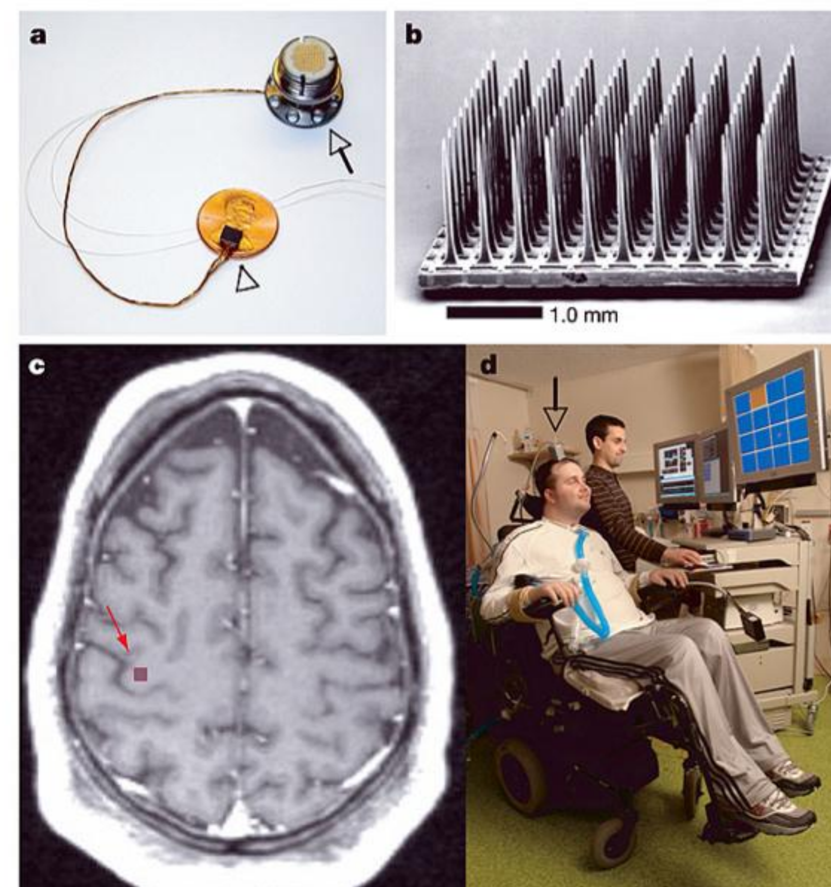

A B C D E F
G H I J K L
M N O P Q R
S T U V W X
Y Z 1 2 3 4
5 6 7 8 9 _

BCI Cycle

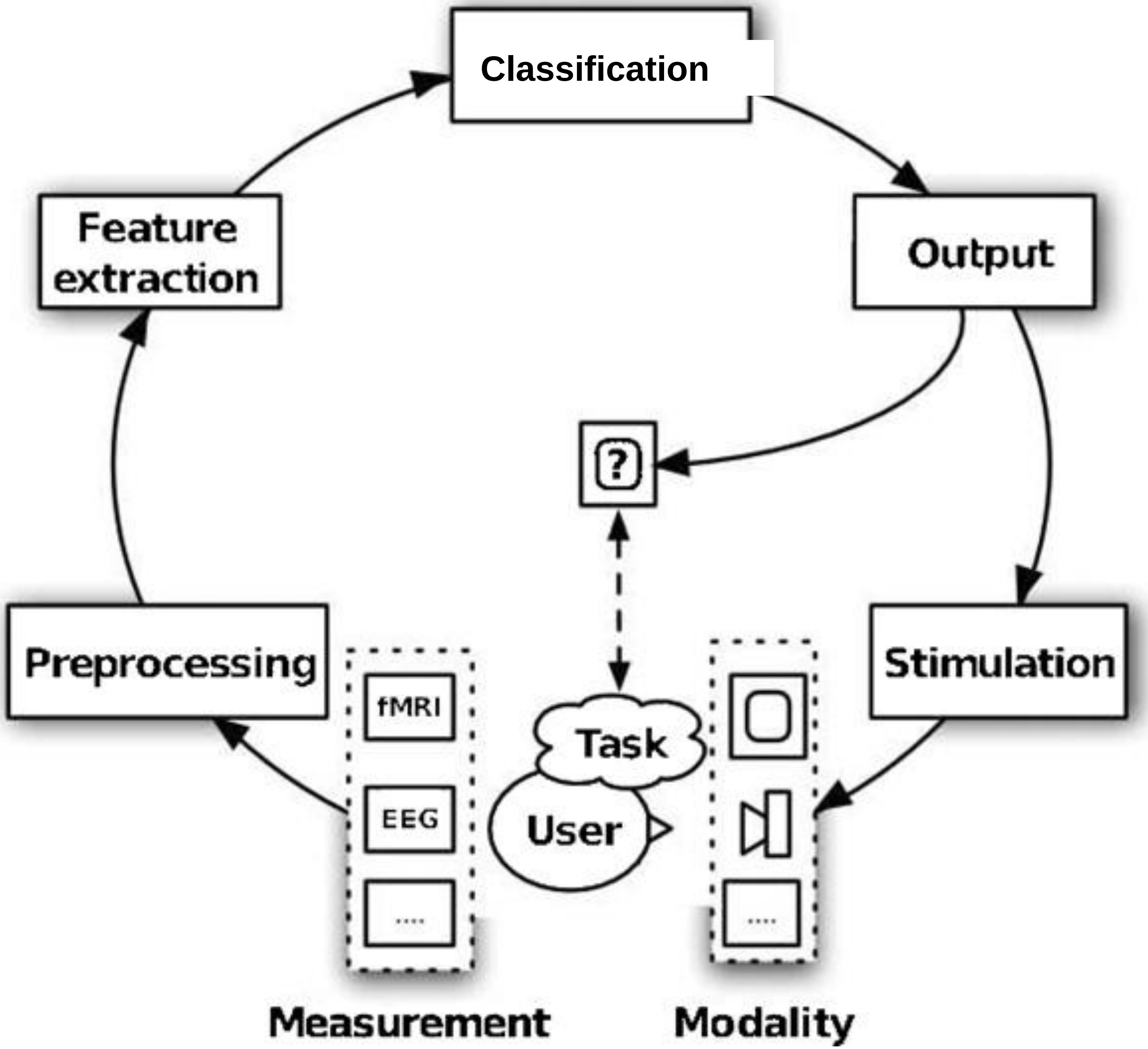

Classification
Feature extraction
Output
?
Preprocessing
Stimulation
fMRI
EEG
....
Task
User
....
Measurement
Modality

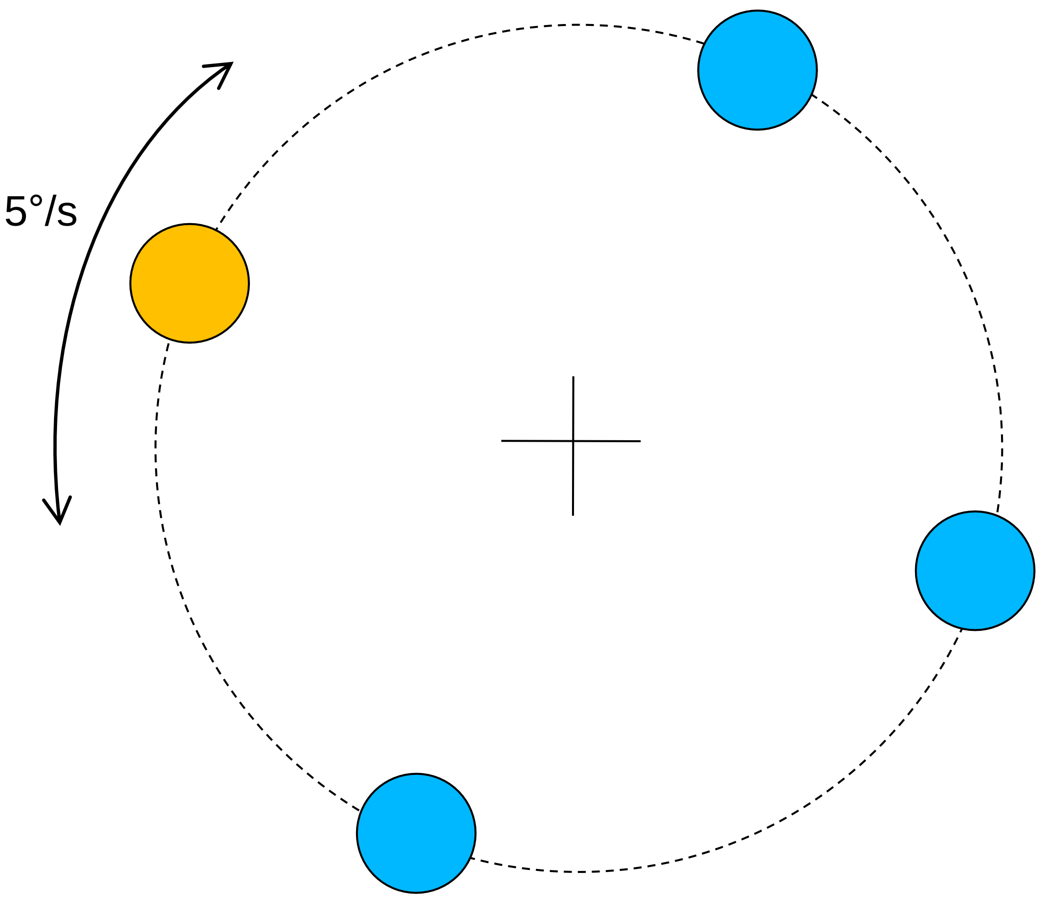
Magnetoencephalography (MEG)

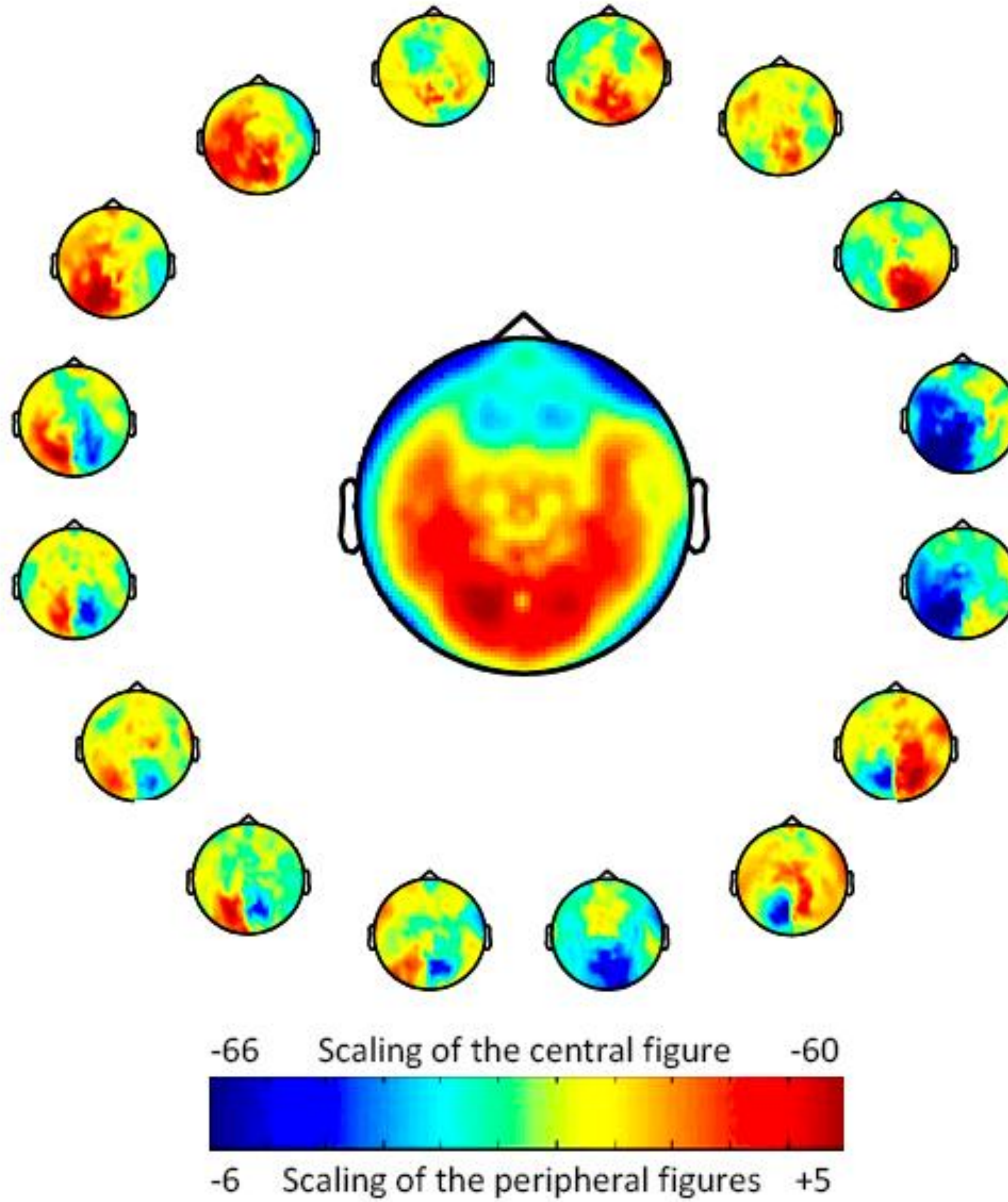
Functional Magnetic Resonance Imaging (fMRI)

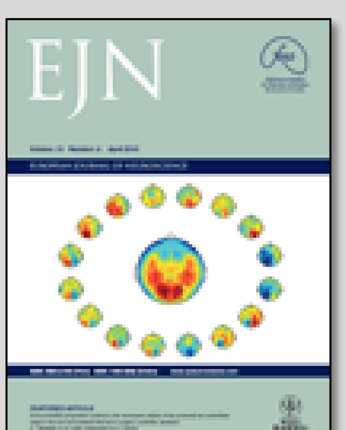


BCI Cycle


Classification
Output
Stimulation
Modality
Measurement
Preprocessing
Feature extraction
fMRI
EEG
....
Task
User
?
....

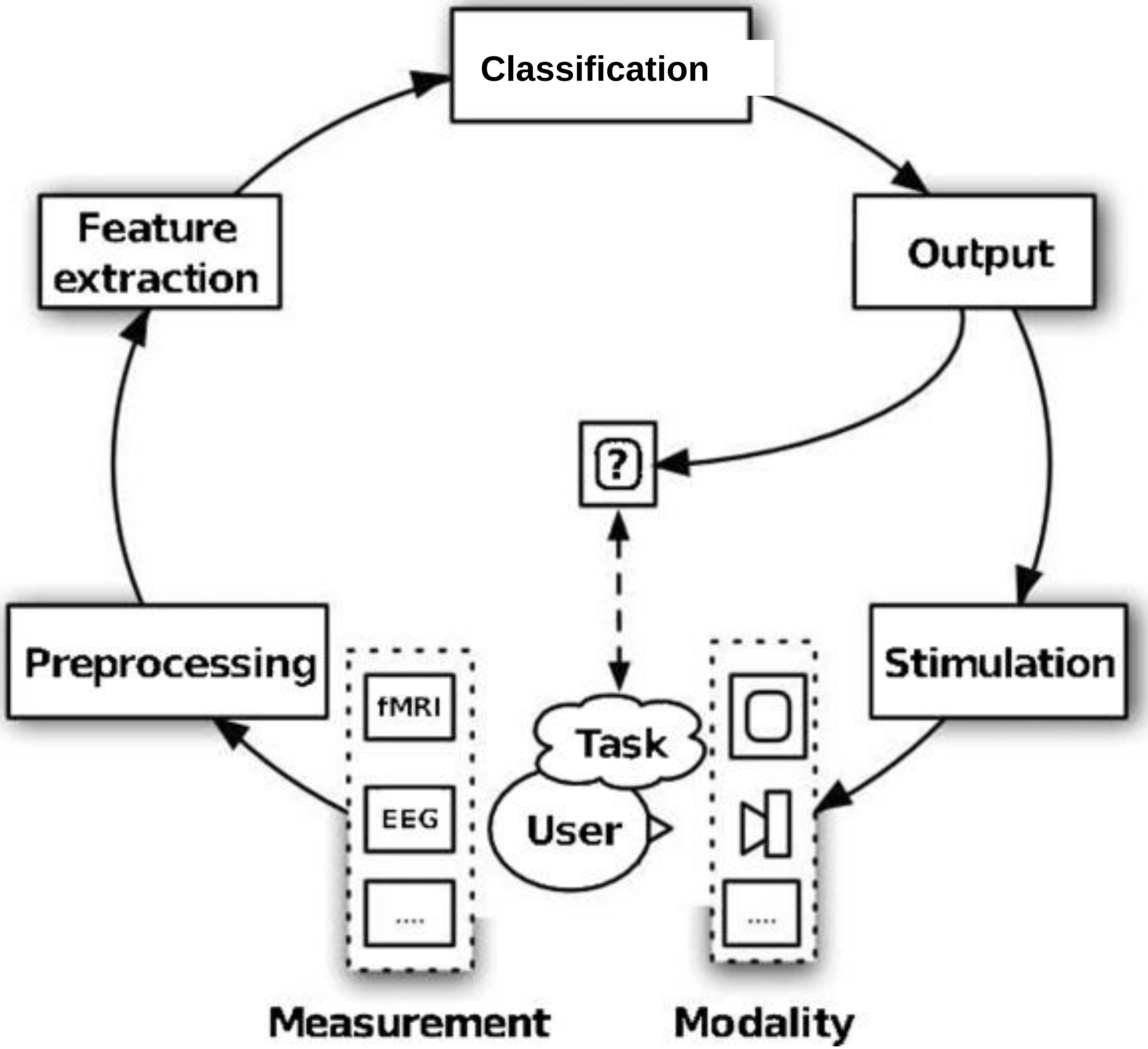


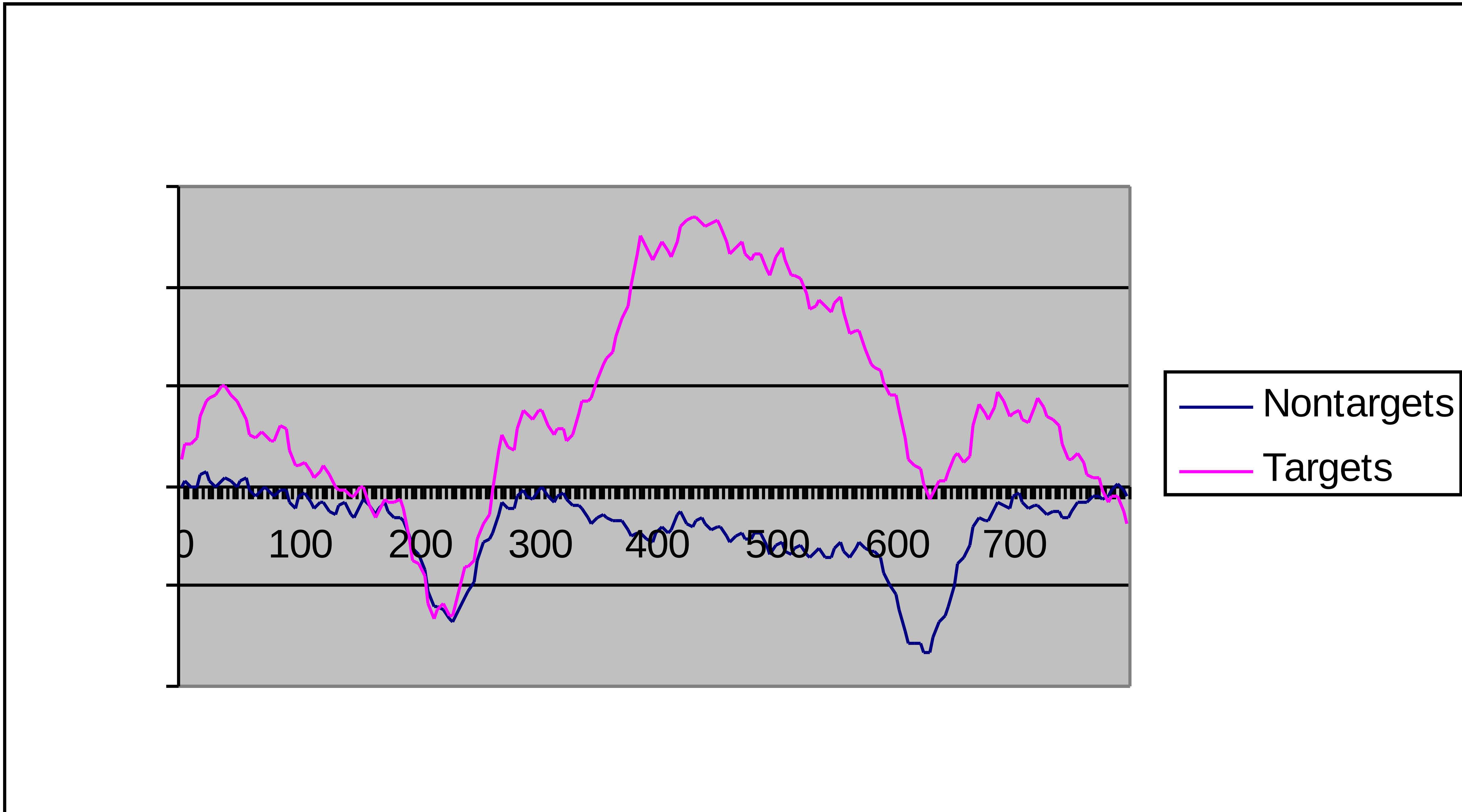




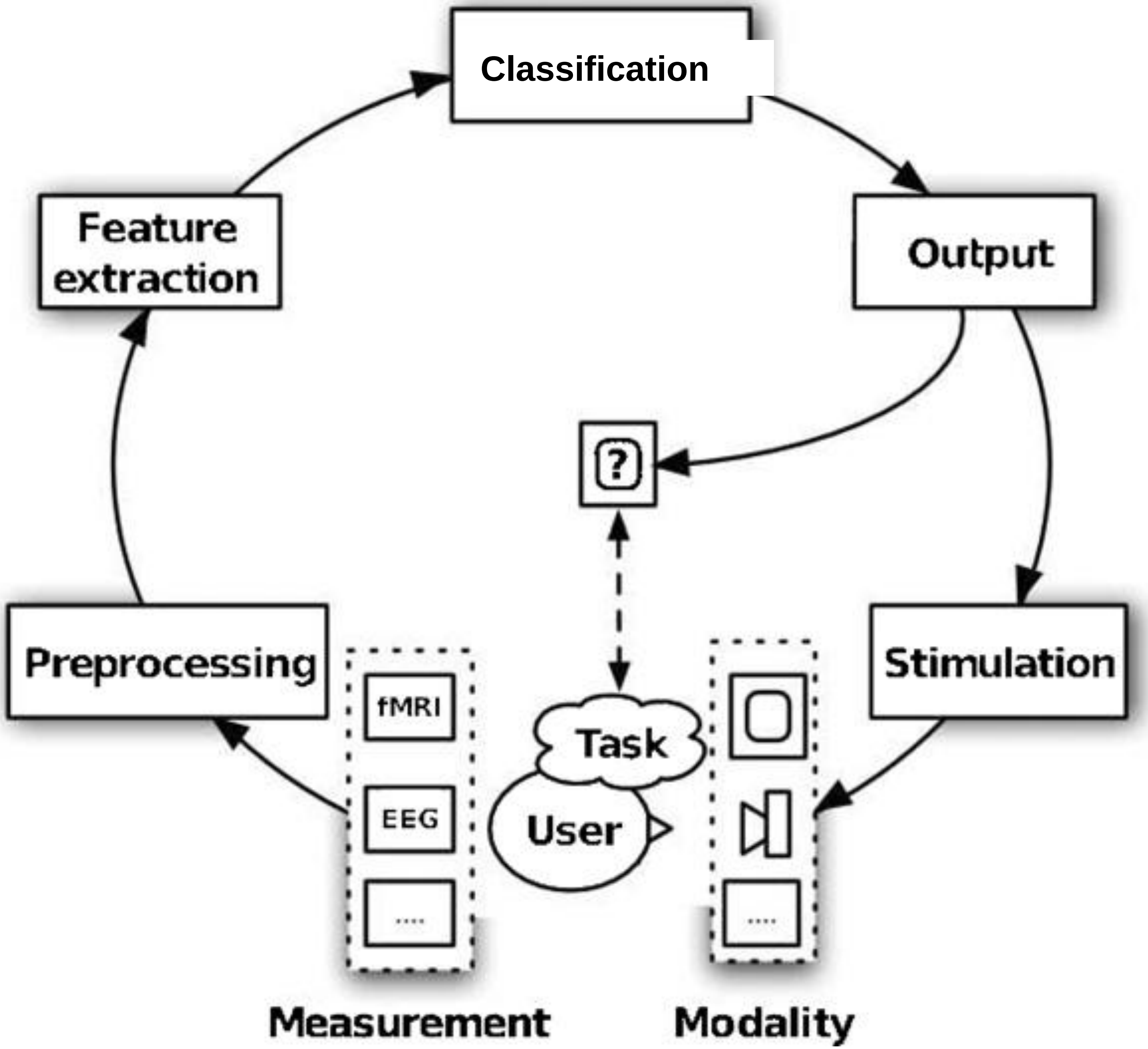
Bahramisharif et al,
Eur J. Neurosci, 2010

BCI Cycle


Classification
Feature extraction
Output
?
Preprocessing
Stimulation
fMRI
EEG
....
Task
User
....
Measurement
Modality

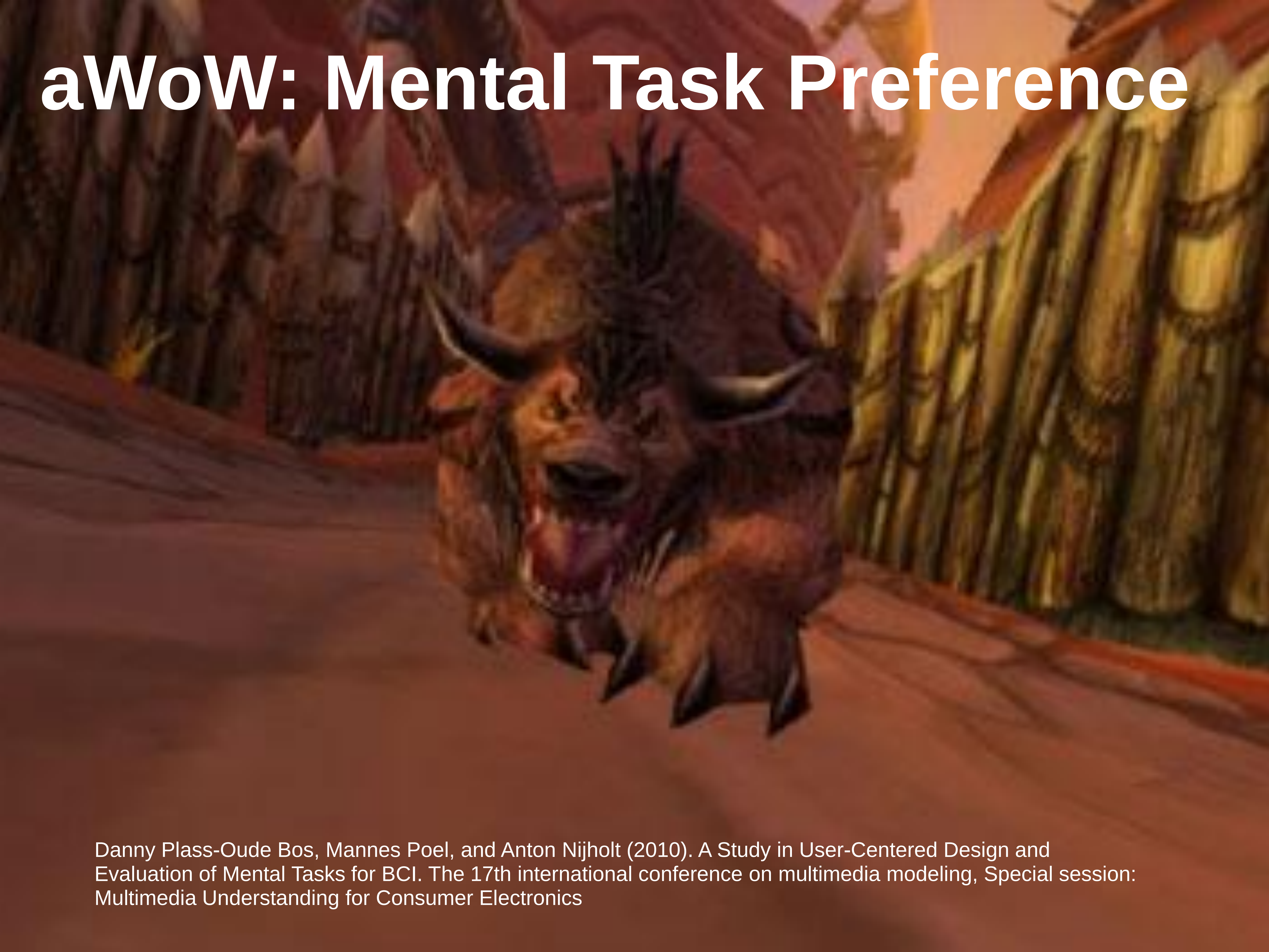

0
100
200
300
400
500
600
700
Nontargets
Targets

BCI Cycle



Classification
Feature extraction
Output
Preprocessing
Stimulation
fMRI
EEG
....
Task
User
....
Measurement
Modality

aWoW: Mental Task Preference



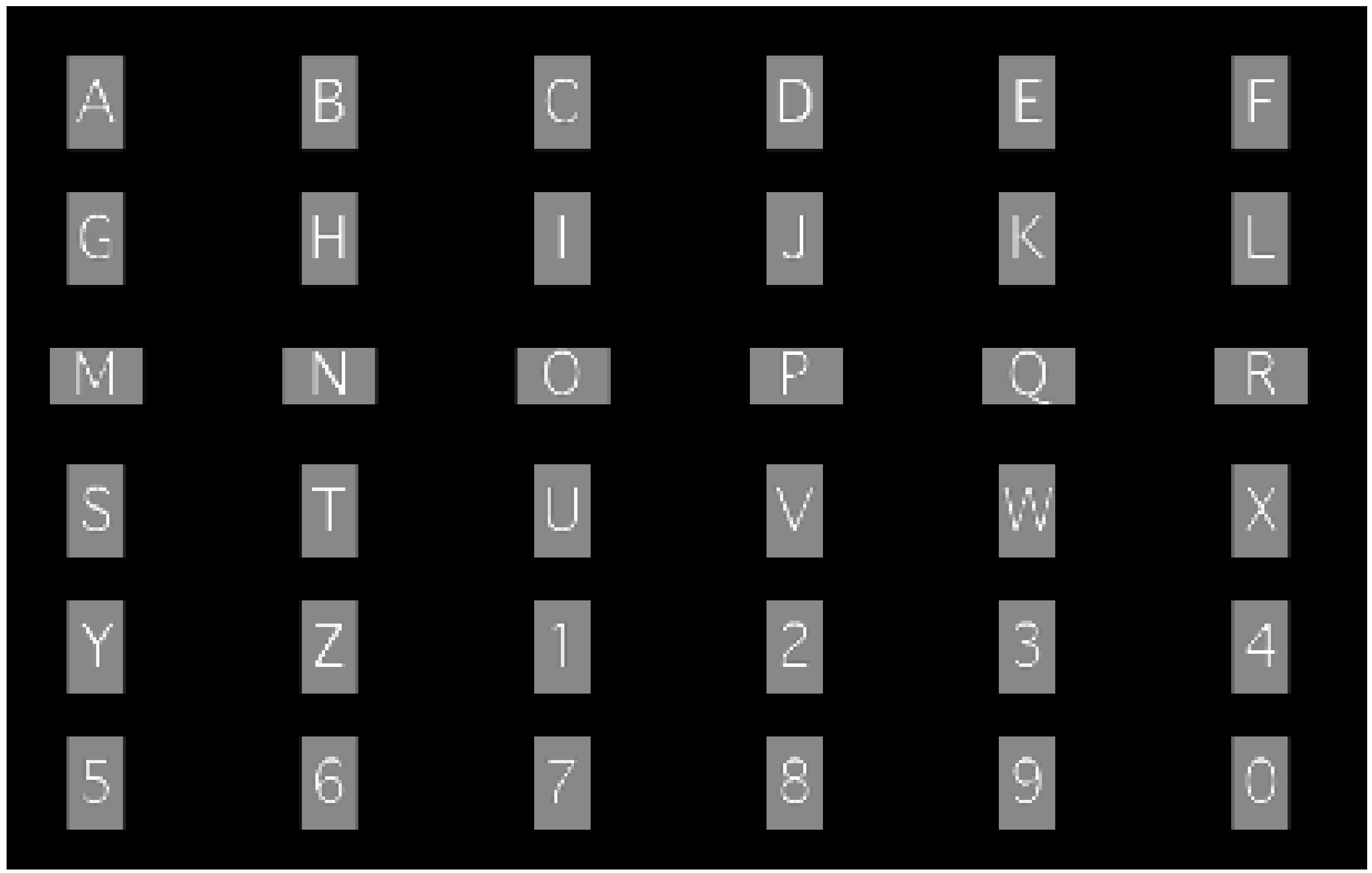
Danny Plass-Oude Bos, Mannes Poel, and Anton Nijholt (2010). A Study in User-Centered Design and Evaluation of Mental Tasks for BCI. The 17th international conference on multimedia modeling, Special session: Multimedia Understanding for Consumer Electronics

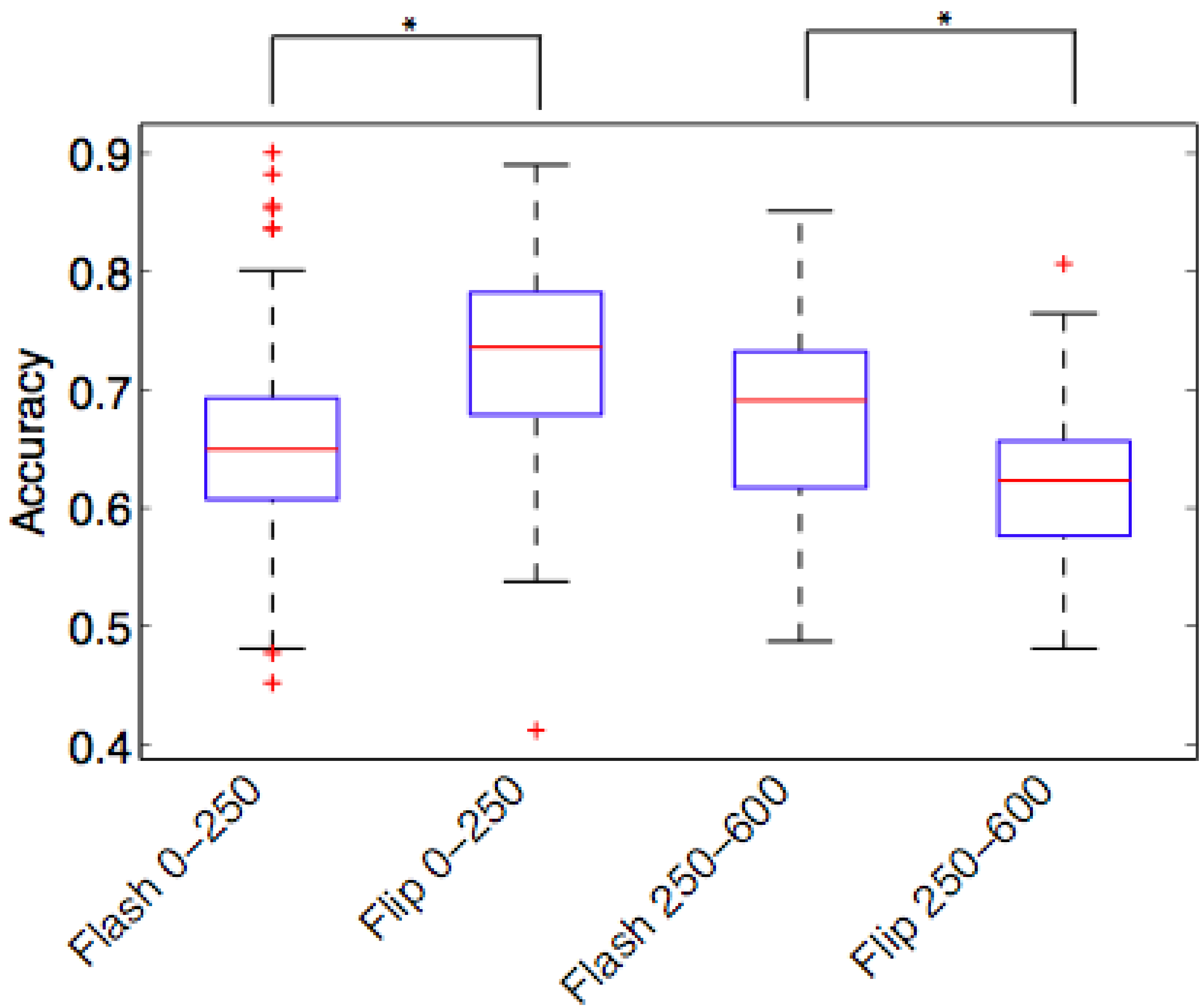
Cell Stimuli: Rate Flash vs Flip, enlarge, even faces

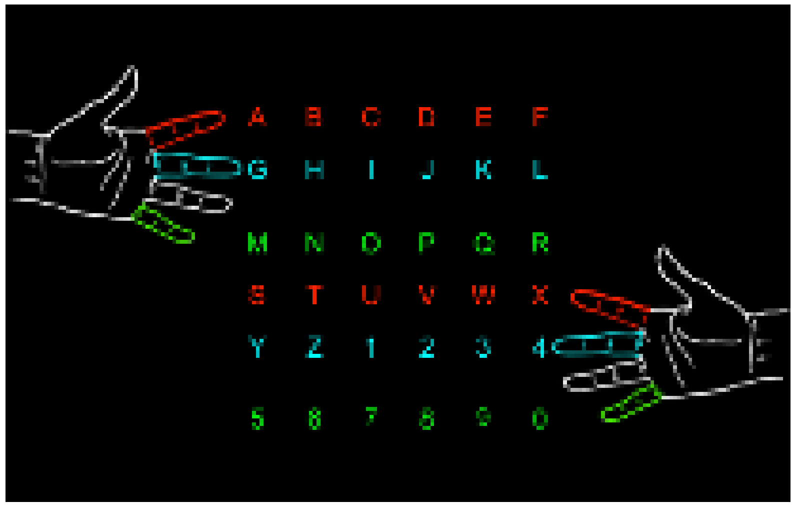
Coding: Row/Col, Scatter

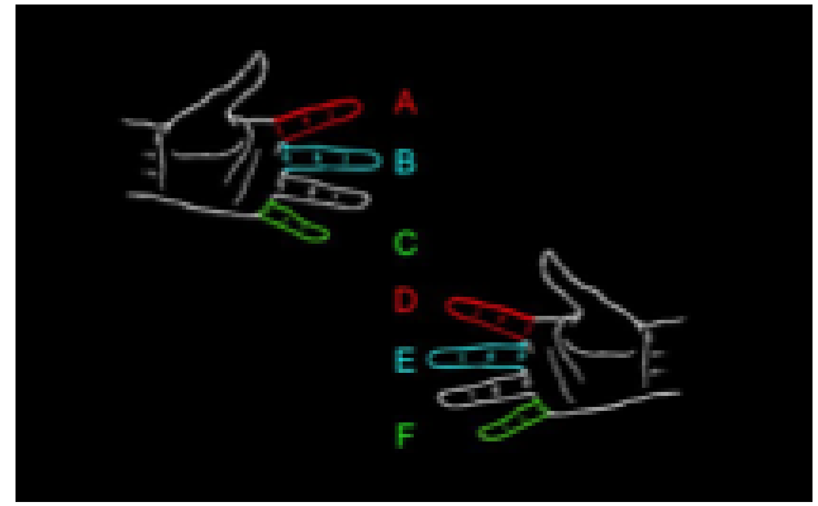
Modality: Auditory, Tactile, multimodal

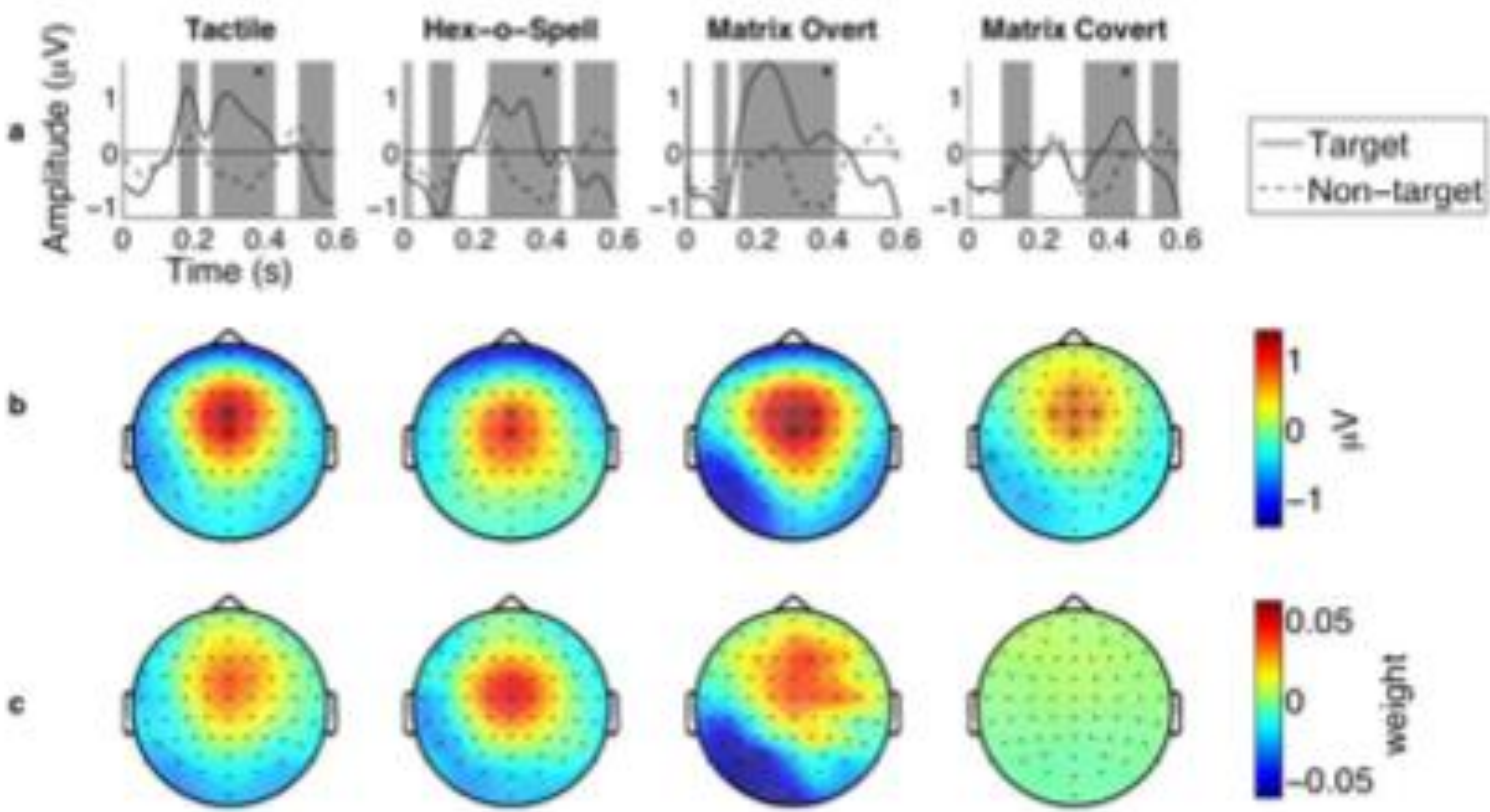
Buttons
  Full Sentence Chat-by-Click
  Locations ...


A B C D E F
G H I J K L
M N O P Q R
S T U V W X
Y Z 1 2 3 4
5 6 7 8 9 0

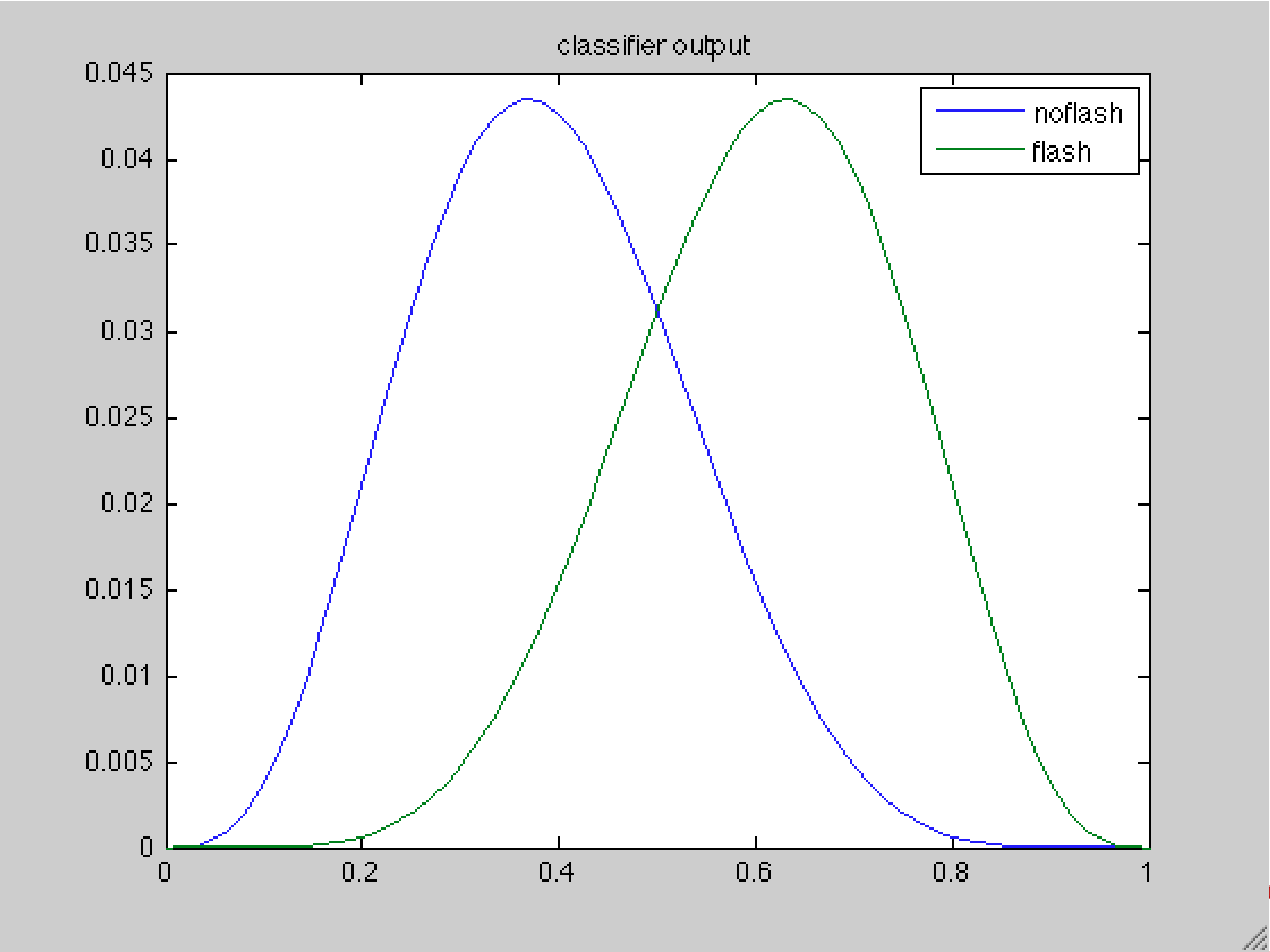

*
*
0.9
0.8
0.7
0.6
0.5
0.4
Accuracy
Flash 0–250
Flip 0–250
Flash 250–600
Flip 250–600


A B C D E F
G H I J K L
M N O P Q R
S T U V W X
Y Z 1 2 3 4
5 6 7 8 9 0


A
B
C
D
E
F

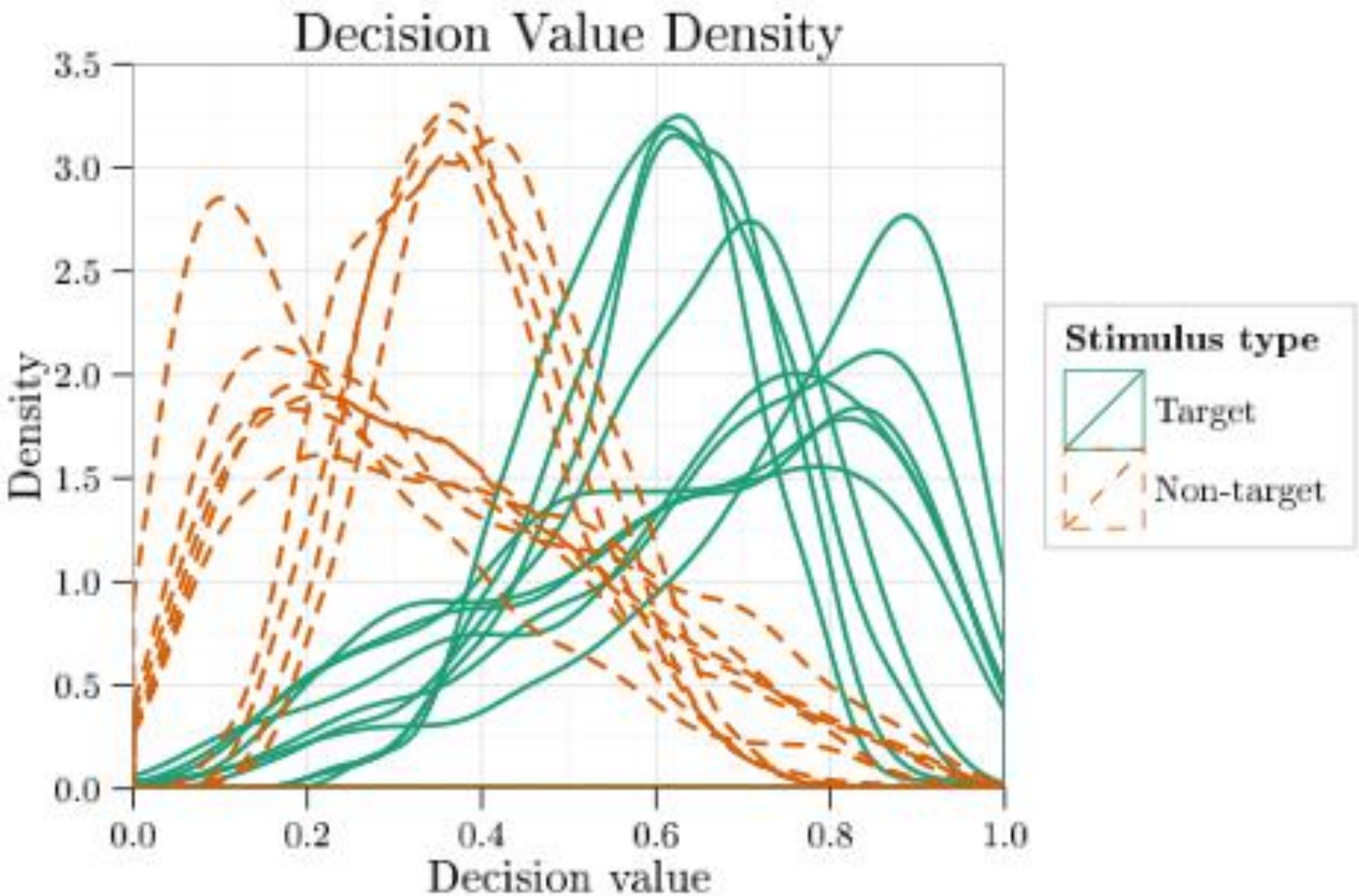

Tactile
Hex-o-Spell
Matrix Overt
Matrix Covert
a
b
c
Amplitude (μV)
Time (s)
Target
Non-target
μV
weight
0.05
-0.05

Classifier output Flash – Non flash

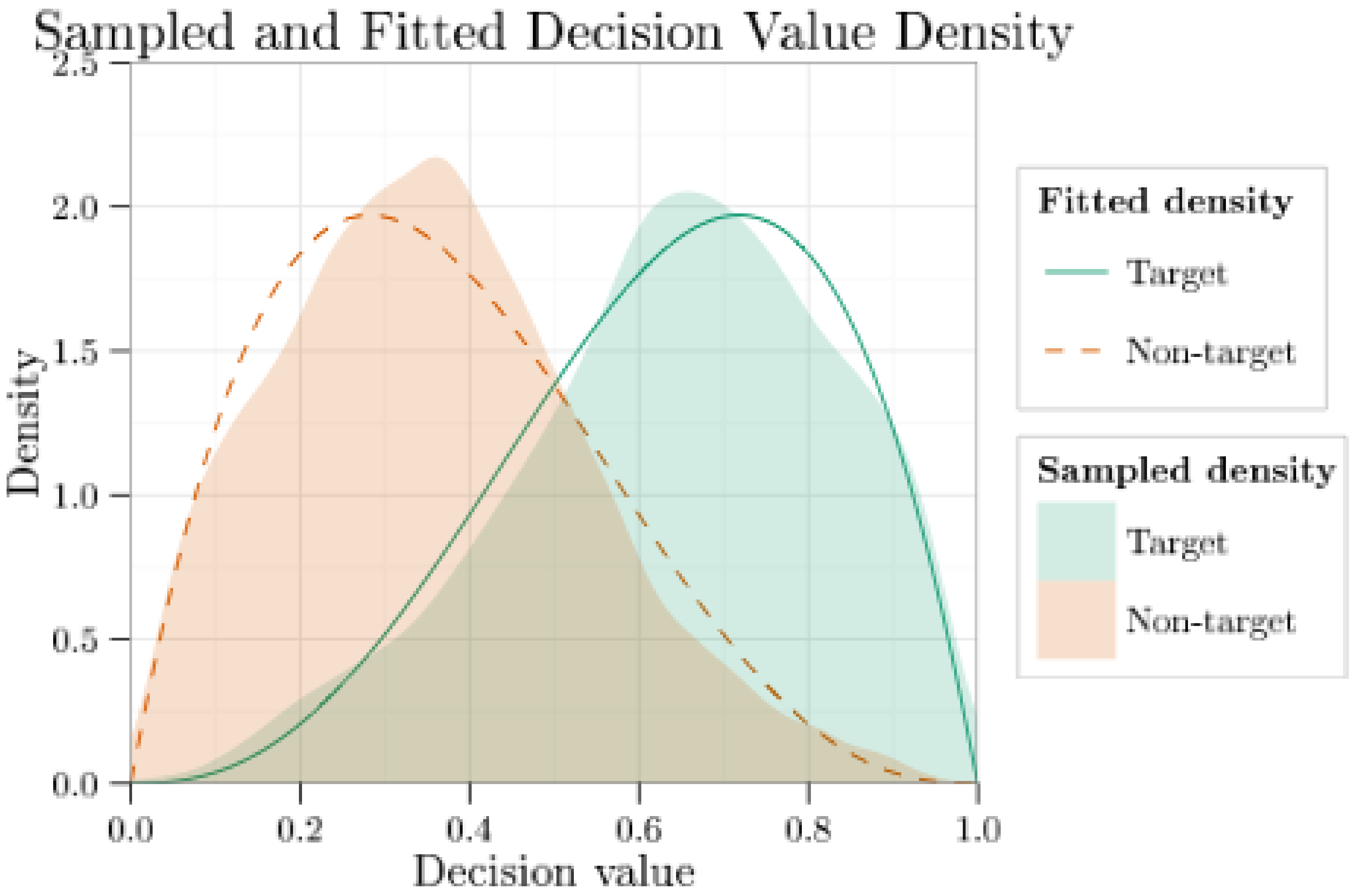


megen

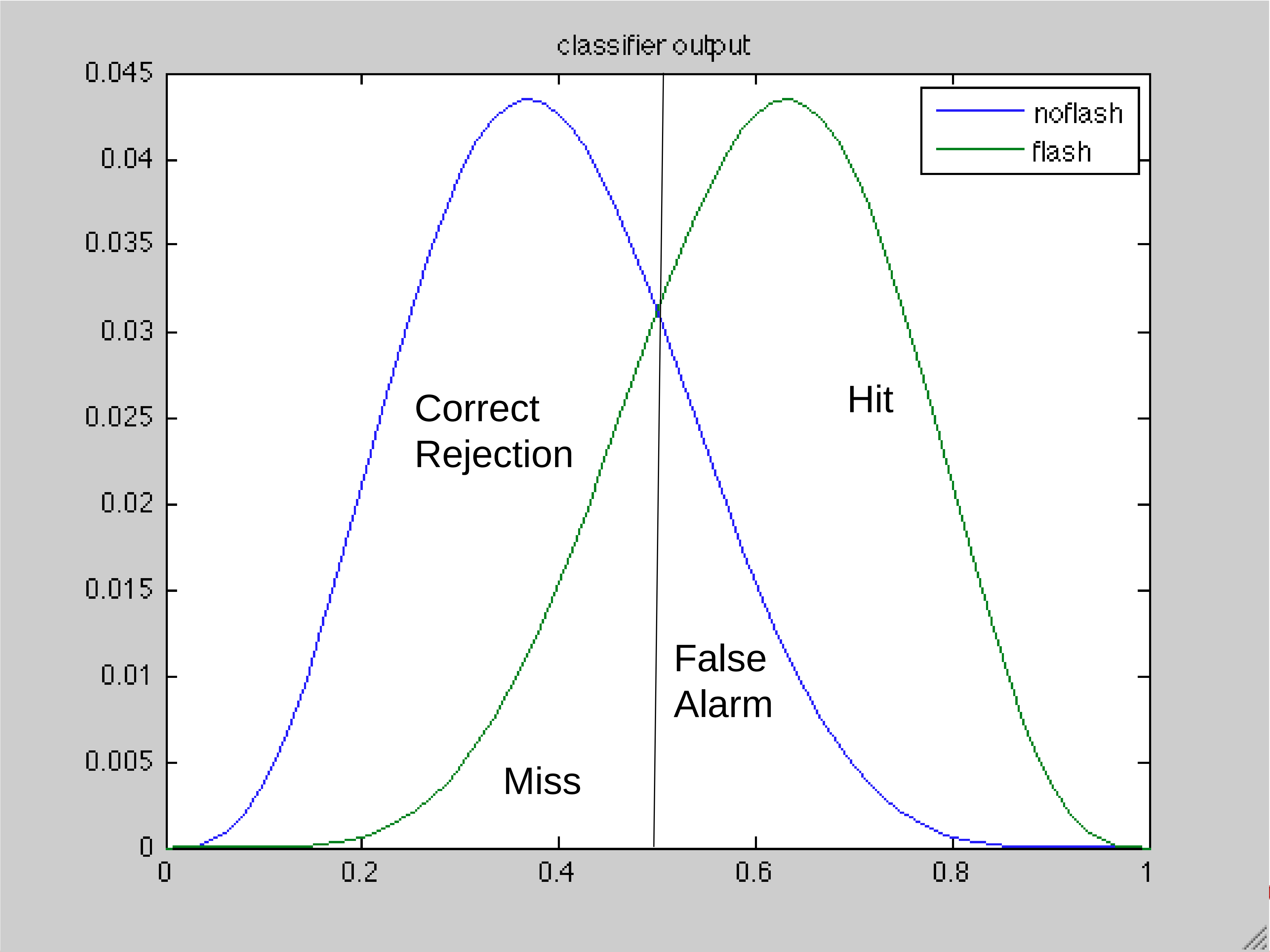
Classifier output Flash – Non flash



Classifier output Flash – Non flash



Classifier output Flash – Non flash



Confusion Matrix

Detected Presented	Flash	No Flash
Flash	**Hit**	**Miss**
No Flash	**False Alarm**	**Correct Rejection**

-> % Correct
-> amount of information transmitted in one flash (in Bits)
-> ROC (AUC)

Flash Sequence

Bitrate per flash (2 class)
maximizes
Bit rate for number of flashes (36 class)

Optimal coding characters (codebook) -> flash sequence
Row/Col coding

Each flash produces evidence (classifier output)
How to combine?

Normalize

Flash Sequence

Classifier output = c, roughly reflects P(target is flashed)

Distribute as evidence for targets, based on flashed row/col

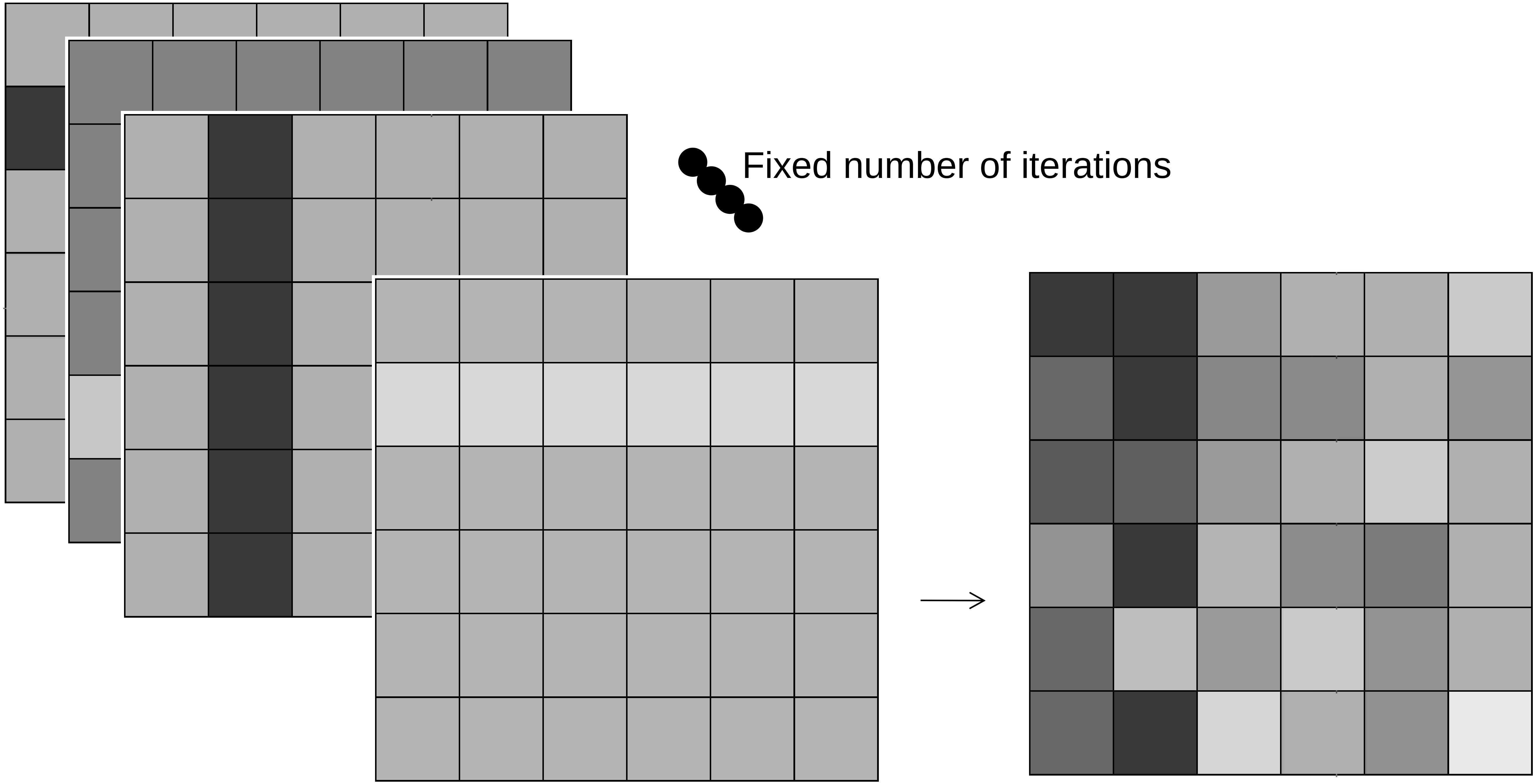
Counter evidence for non-flased row/cols

c $\longrightarrow$

1-c	1-c	1-c	1-c	1-c	1-c
1-c	1-c	1-c	1-c	1-c	1-c
c	c	c	c	c	c
1-c	1-c	1-c	1-c	1-c	1-c
1-c	1-c	1-c	1-c	1-c	1-c
1-c	1-c	1-c	1-c	1-c	1-c

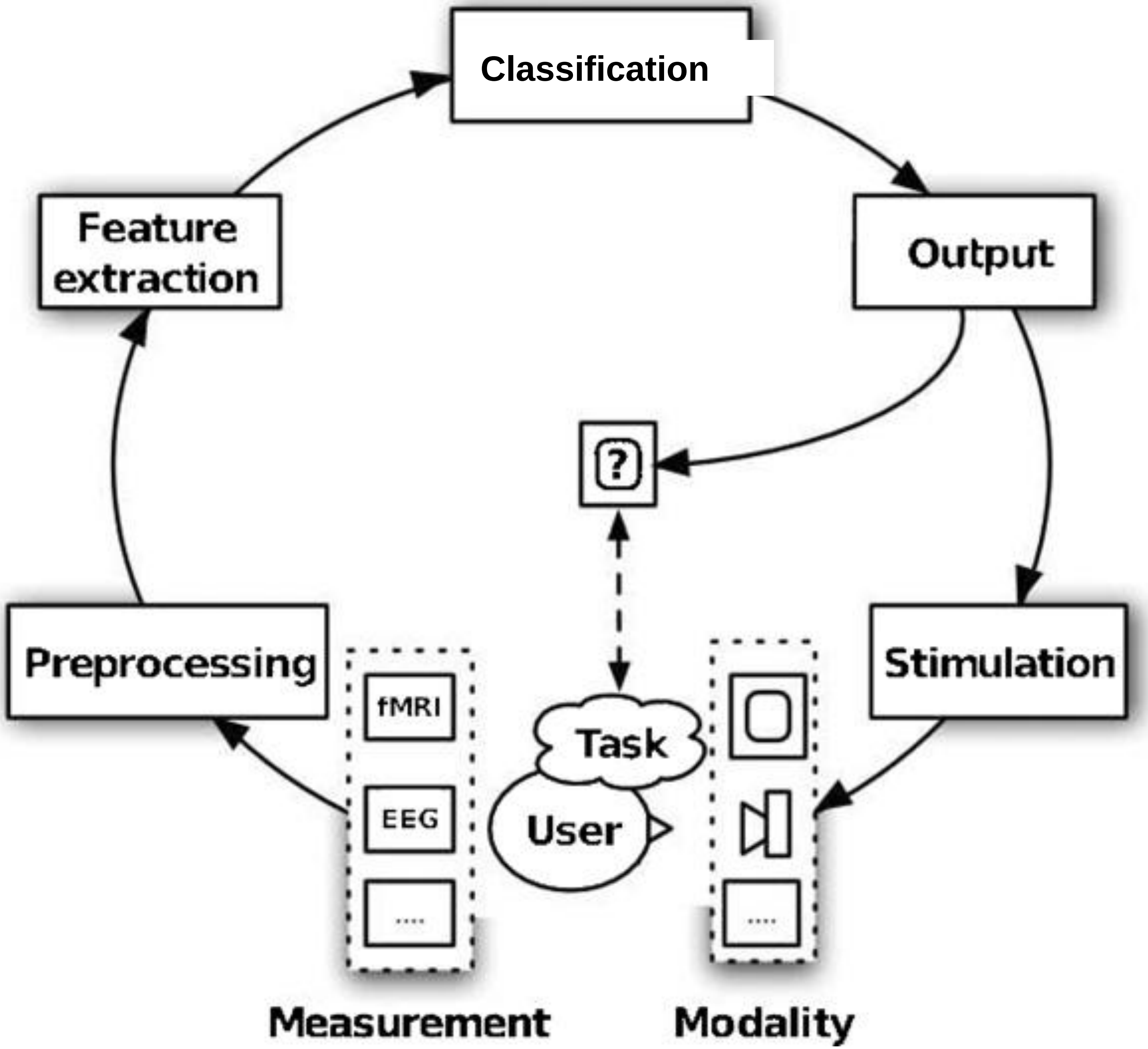
Normalize

Flash Sequence

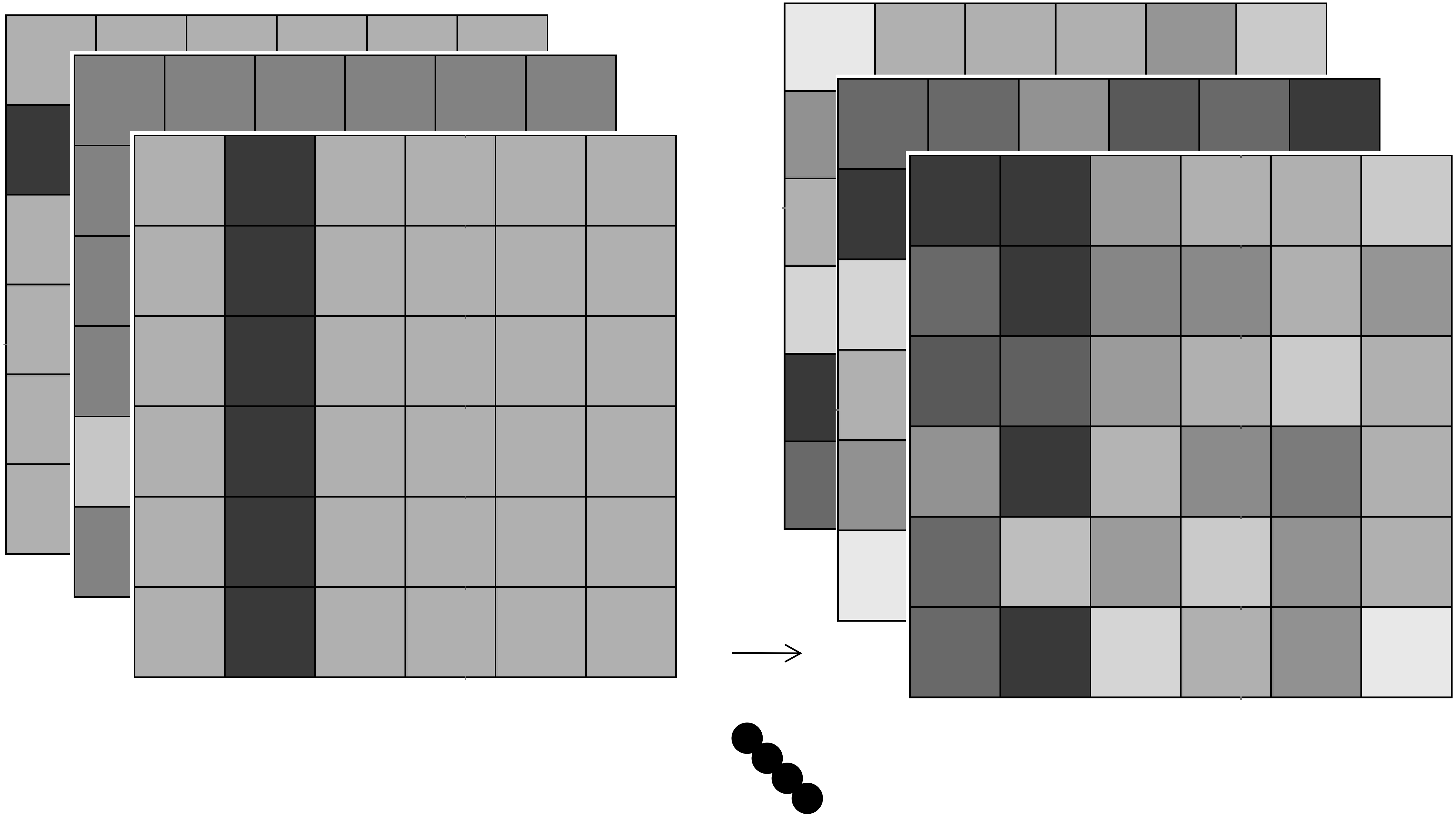


Multiply and re-normalise

Exploit BCI Cycle, make it loop faster



Flash Sequence, maintain current belief state



Flash Sequence, Online Incremental Detection

Early stopping,   careful, not just threshold !

Next flash choice, complication: late responses

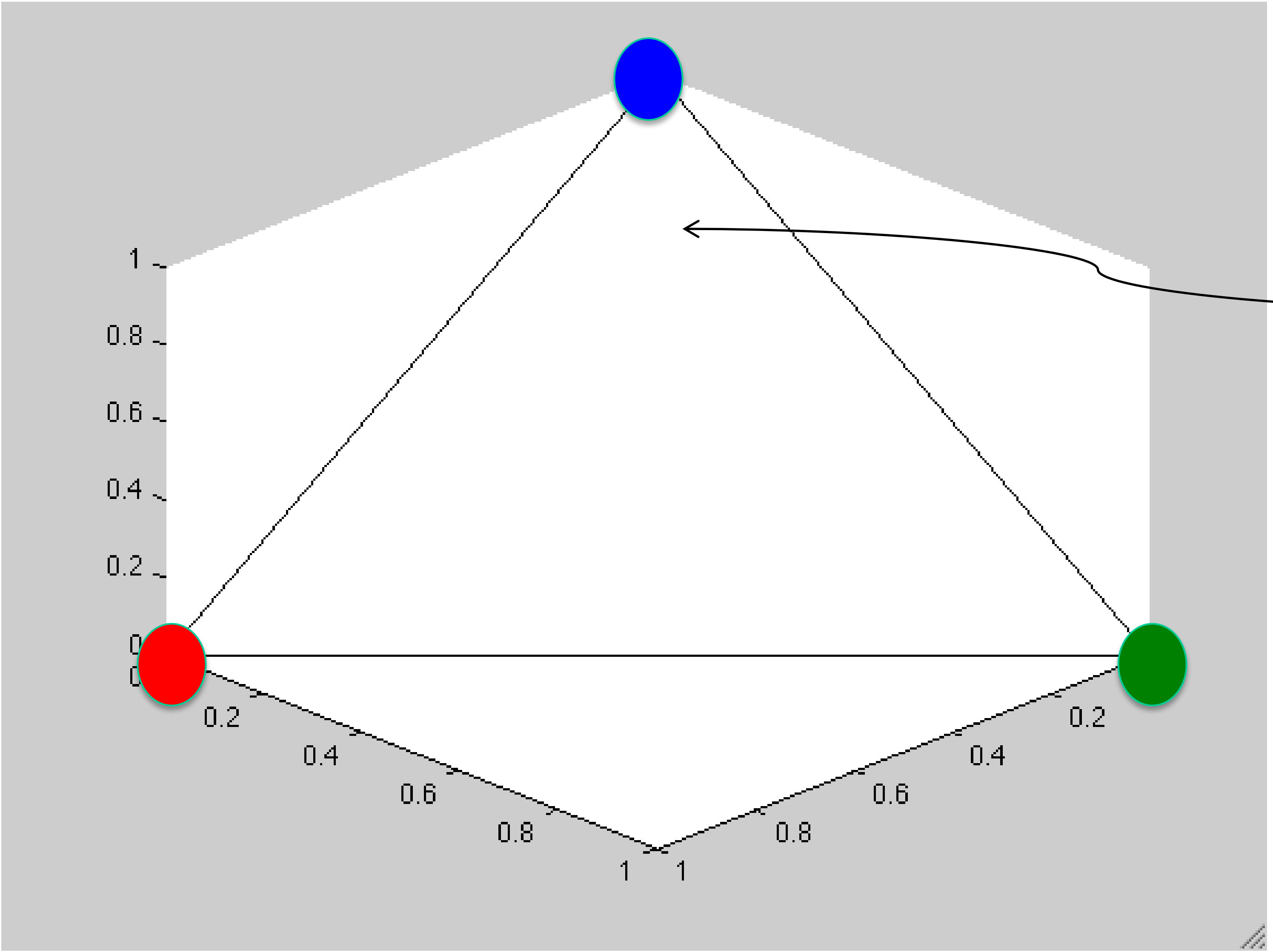
vs Random (permutation)

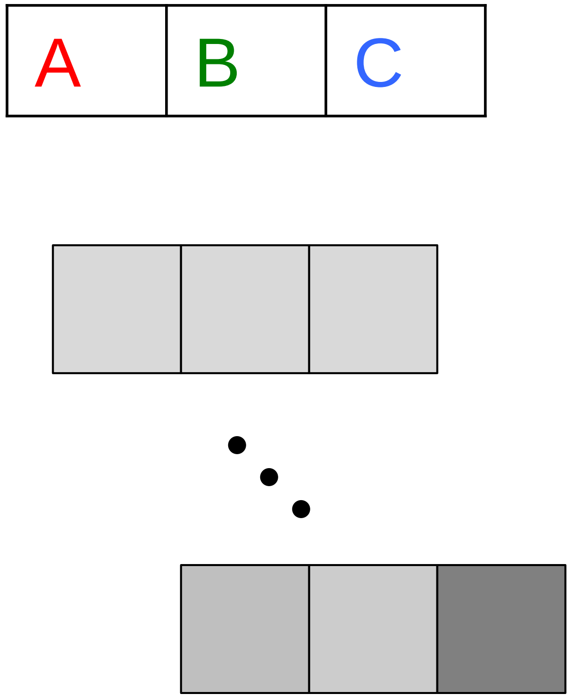
Local decision possible, simple rule?

Normalize

Adaptive Speller Matrix, visualization?

Simplify: 3D
$p_a+p_b+p_c=1$





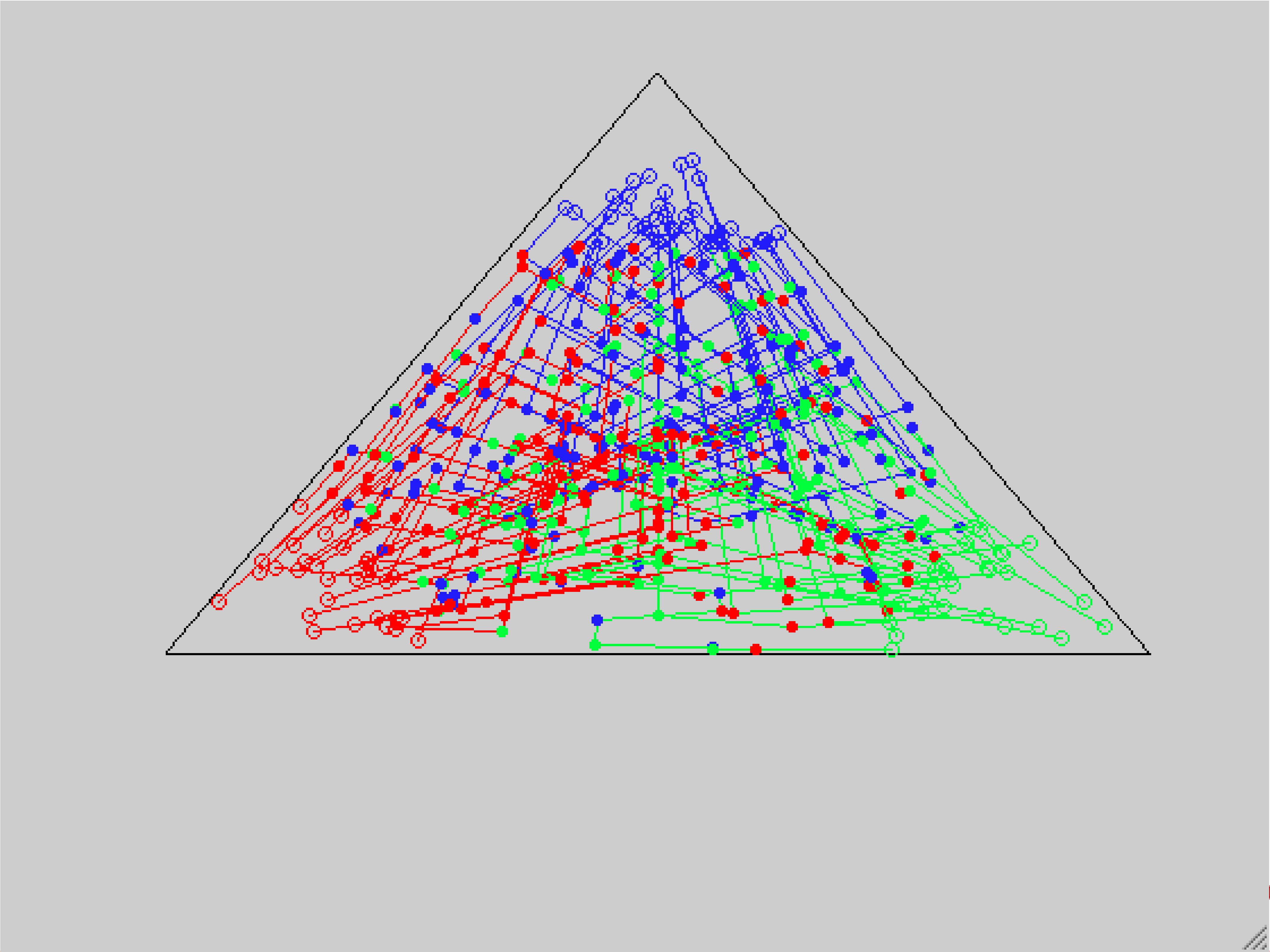
State =
Point in believe space

Flash Sequence

0) Flash Random

Normalize

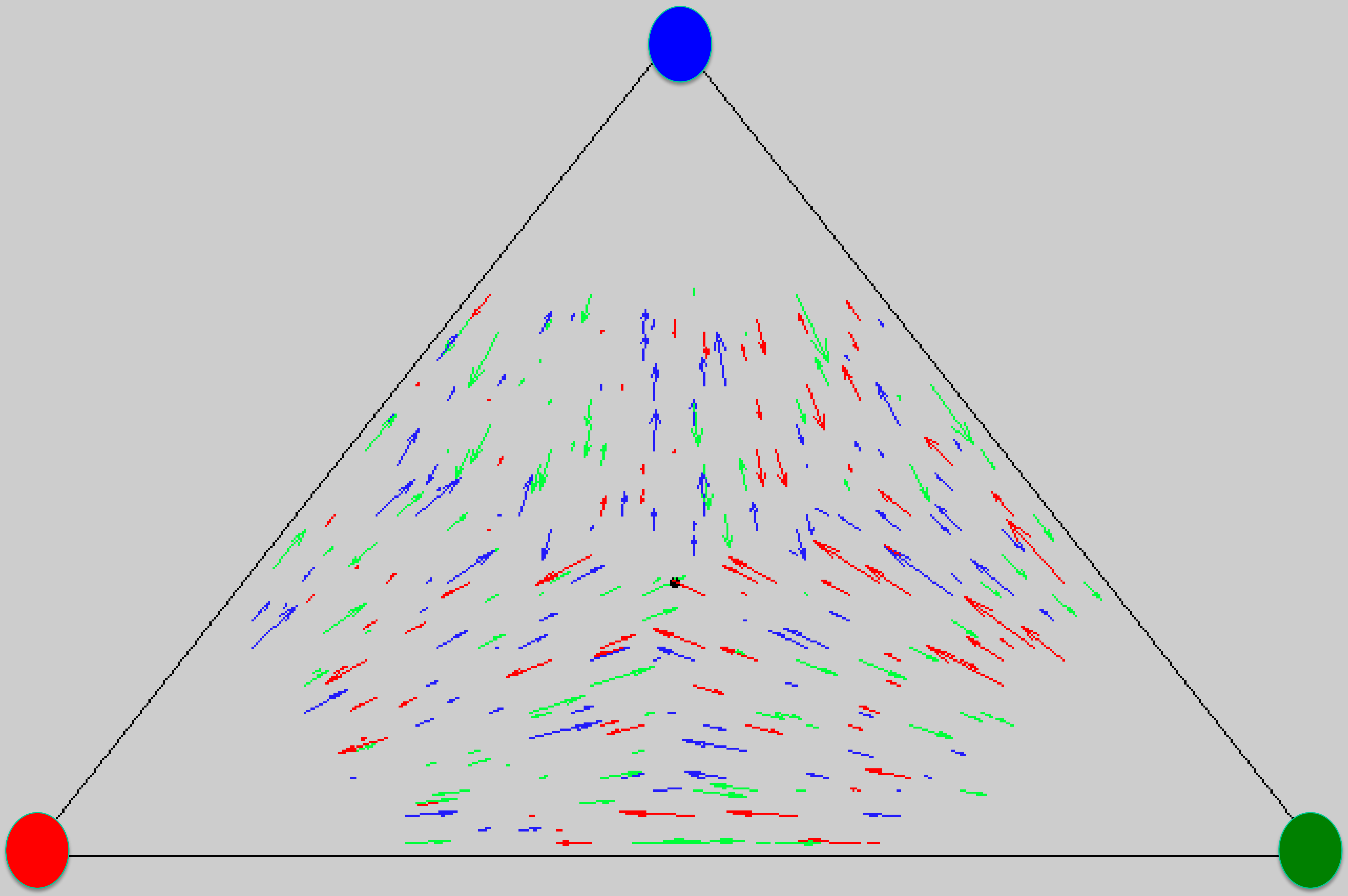
Results (flash random)

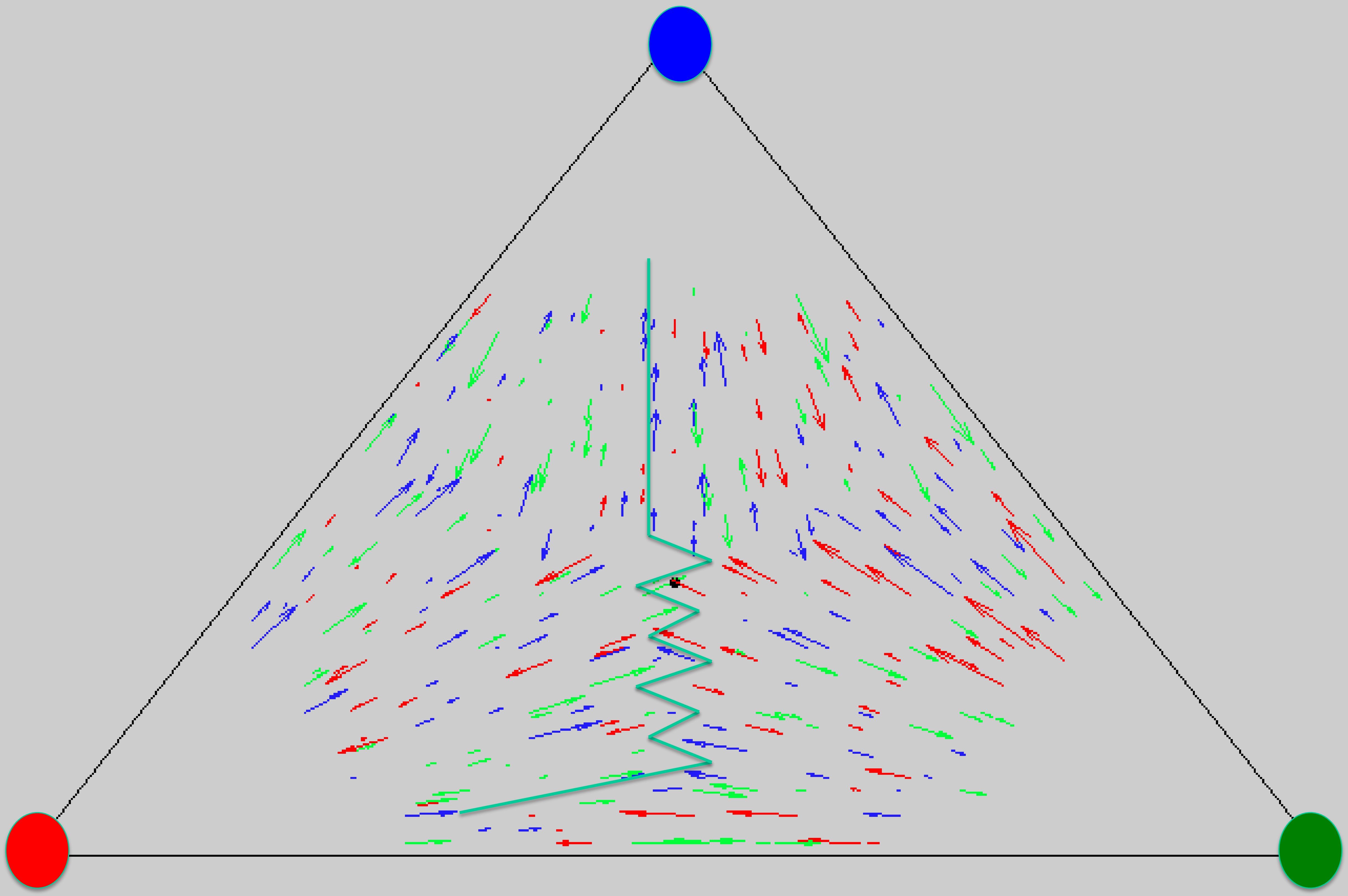


Flash Sequence

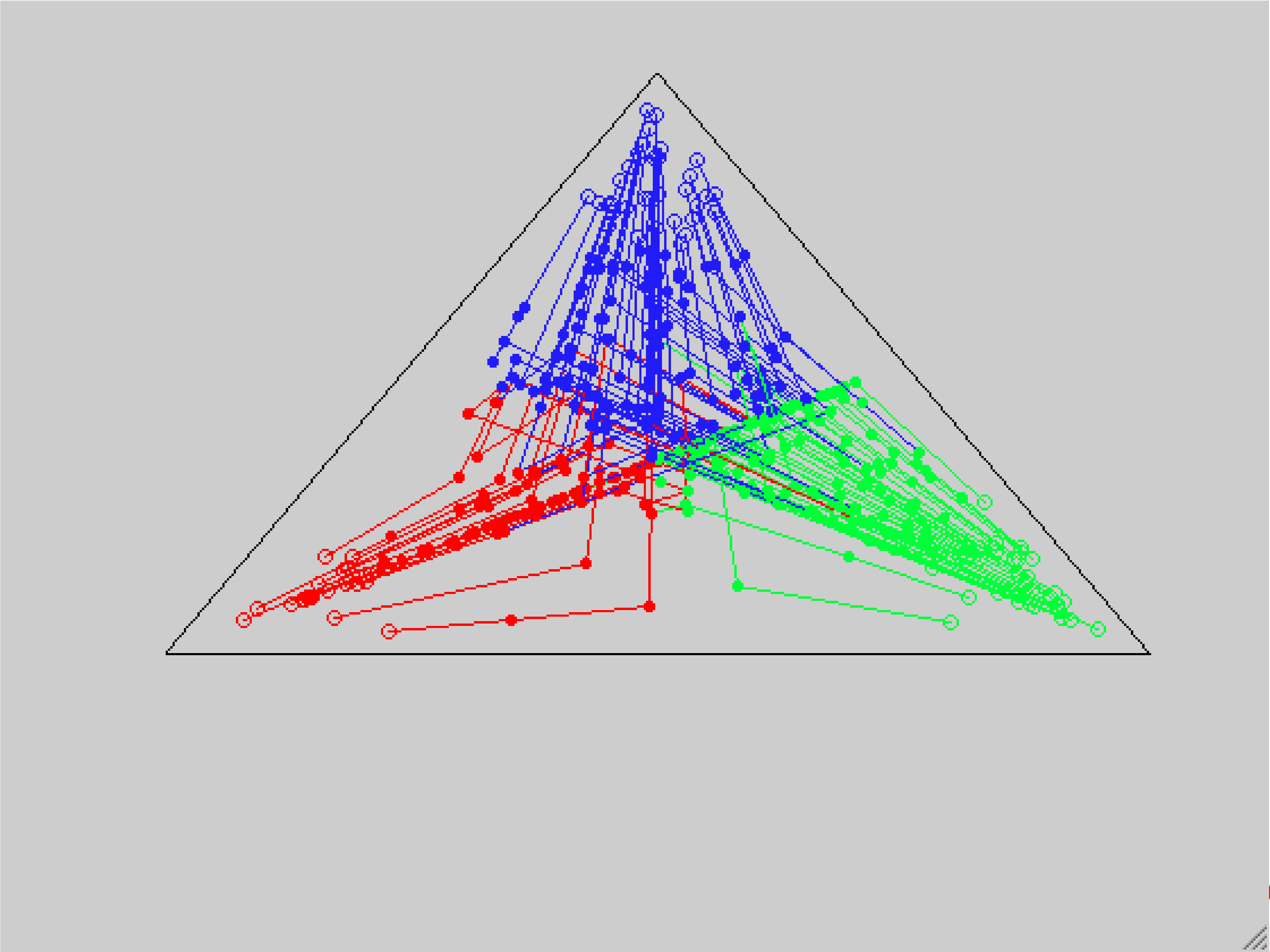
0) Flash Random
1) Flash Most Promising Candidate

Normalize





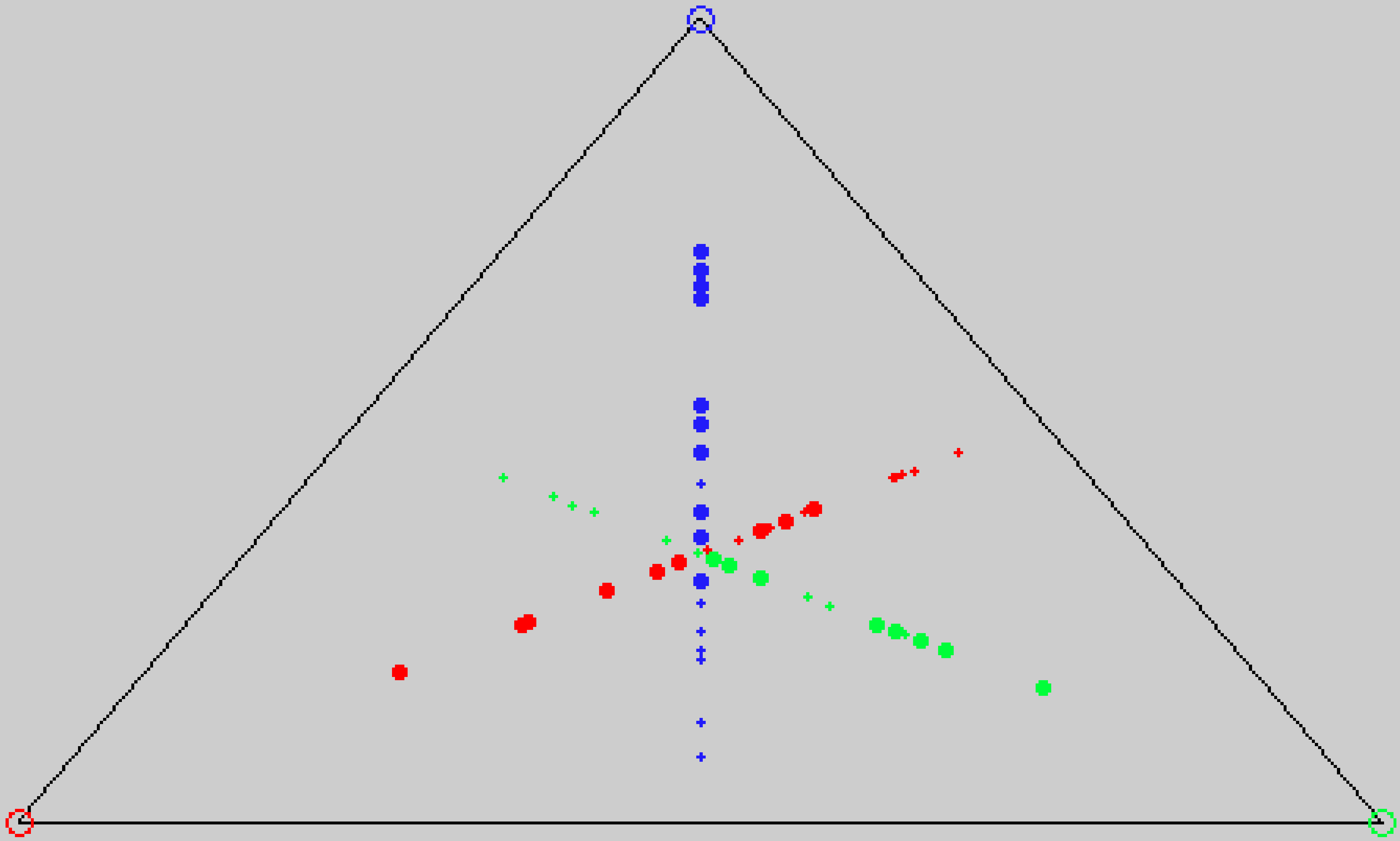
Results (flash max)



Flash Sequence

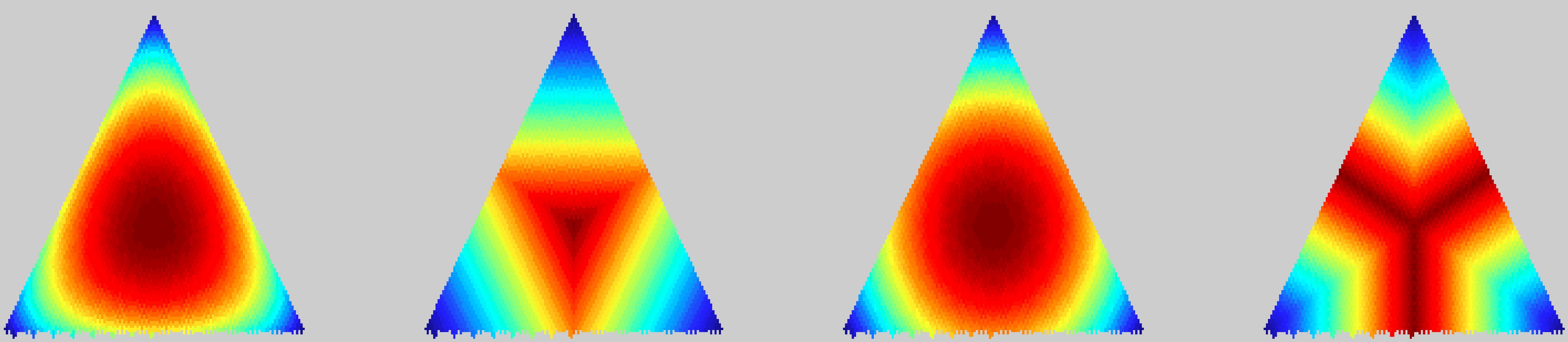
0) Flash Random
1) Flash Most Promising Candidate
2) Flash second best
3) Optimize expected criterion
   - model next belief state distribution assuming target, given flash
   - calculate expected criterion (over targets)
   - pick best flash

Normalize



Objective to be minimized

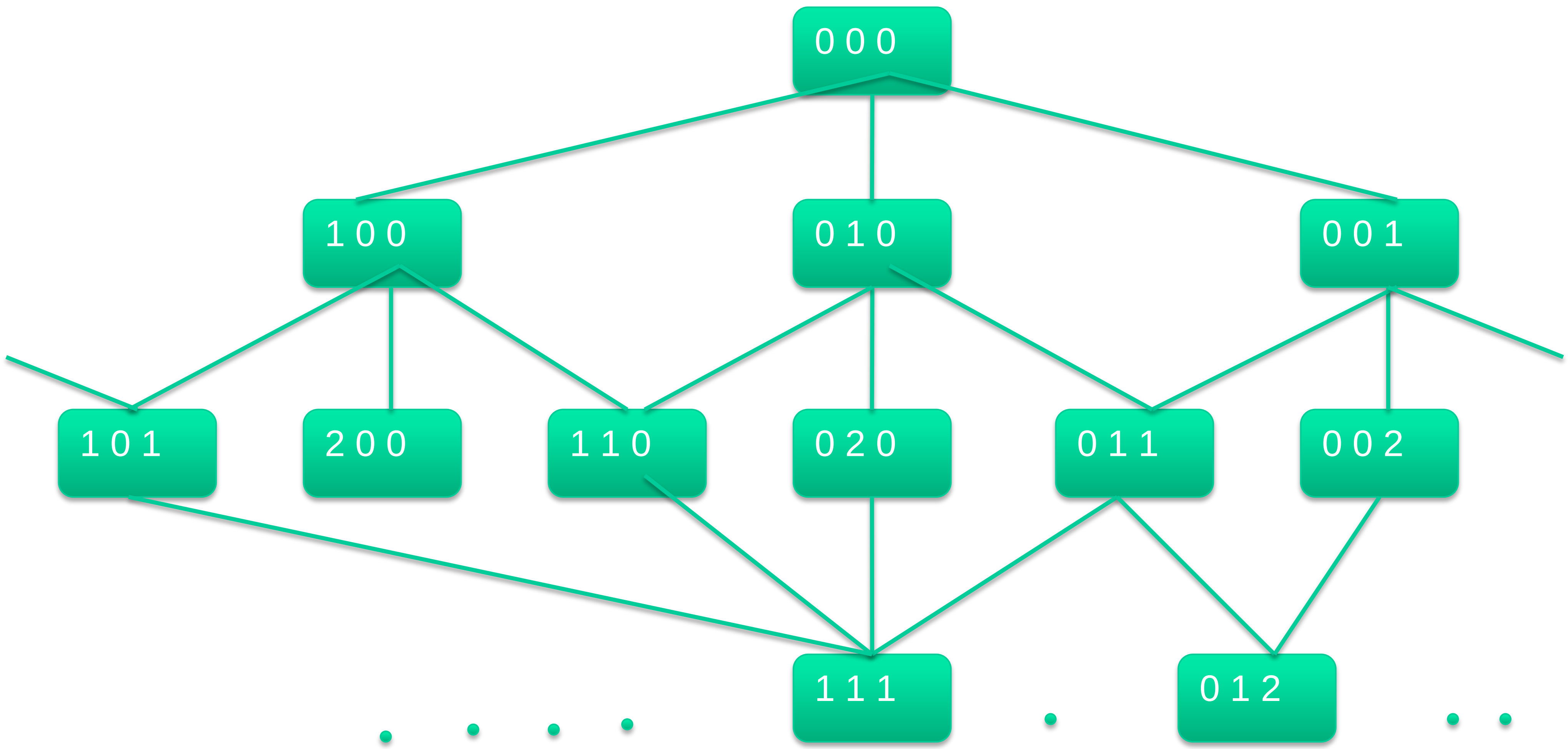
Entropy Error Distance from uniform (Gini) Margin



1) Flash best candidate -
2) Flash second best
3) Optimize expected criterion
4) Fully model belief state distributions, optimize outcome
    - early stopping trivial (expected % correct > ...)
    - doable
        - exploit symmetries, order independencies

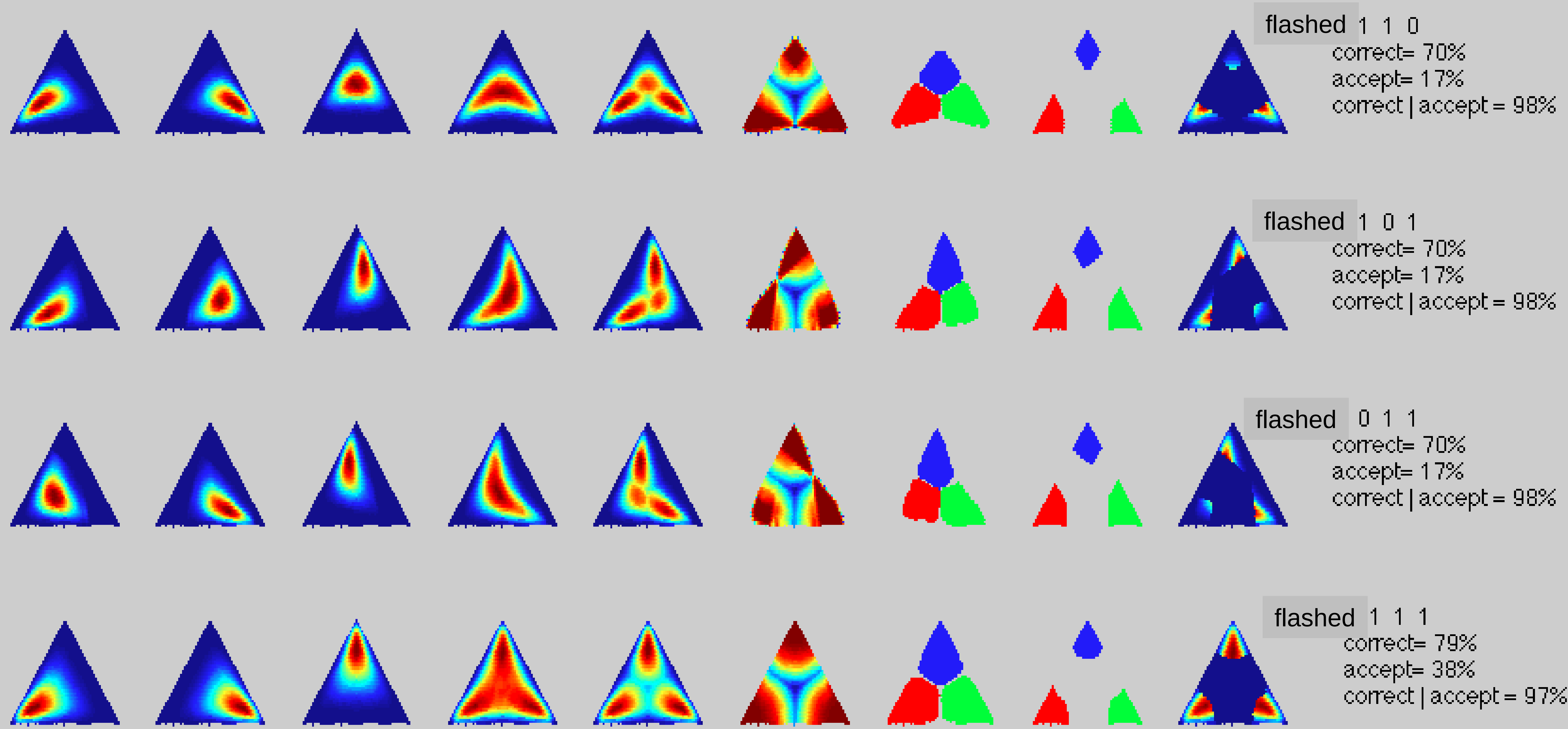
Normalize

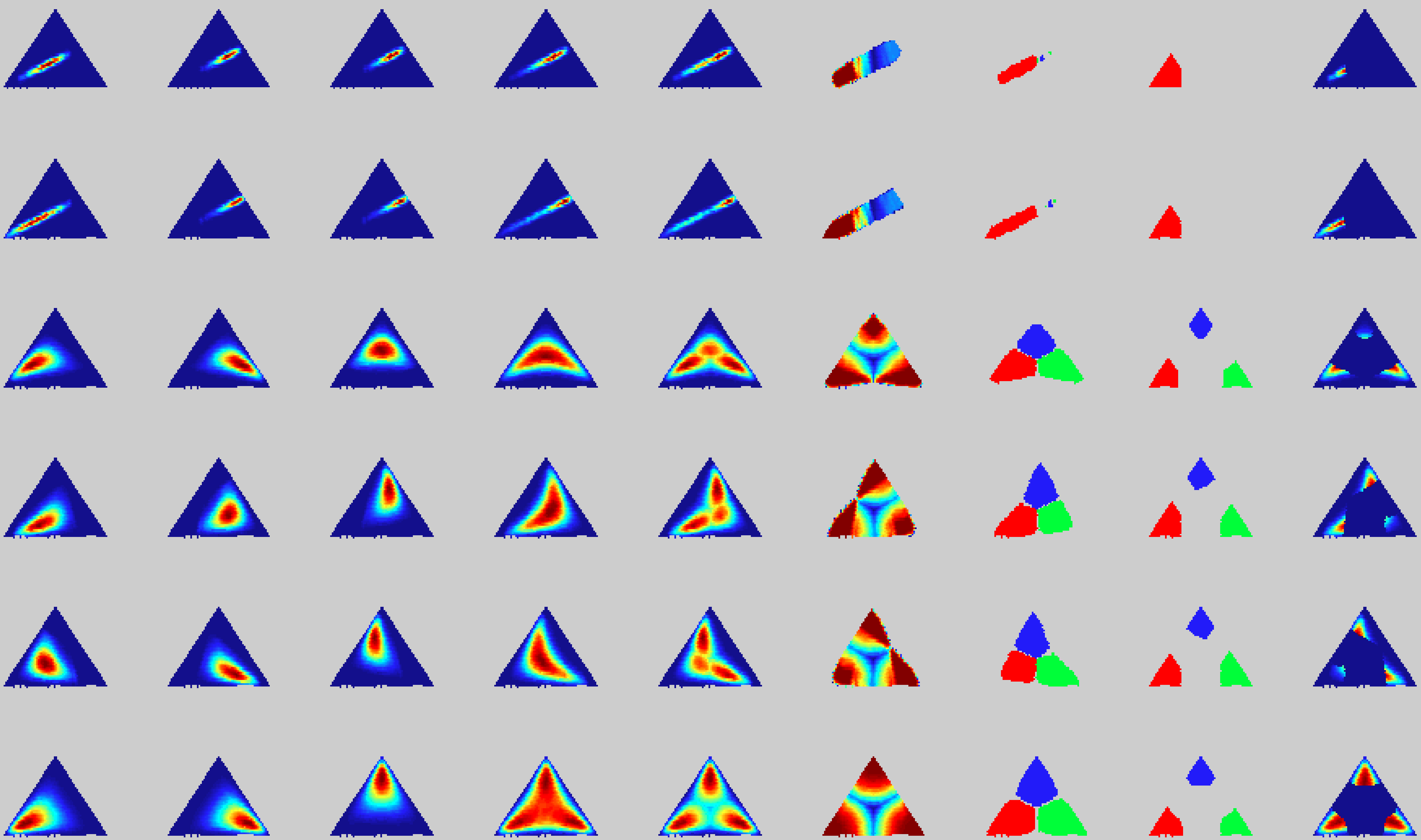
State transitions



Distributions

T=1 T=2 T=3 Av Max Correct best accept correct





0 1 0 0
correct= 52%
accept= 7%
correct | accept = 99%

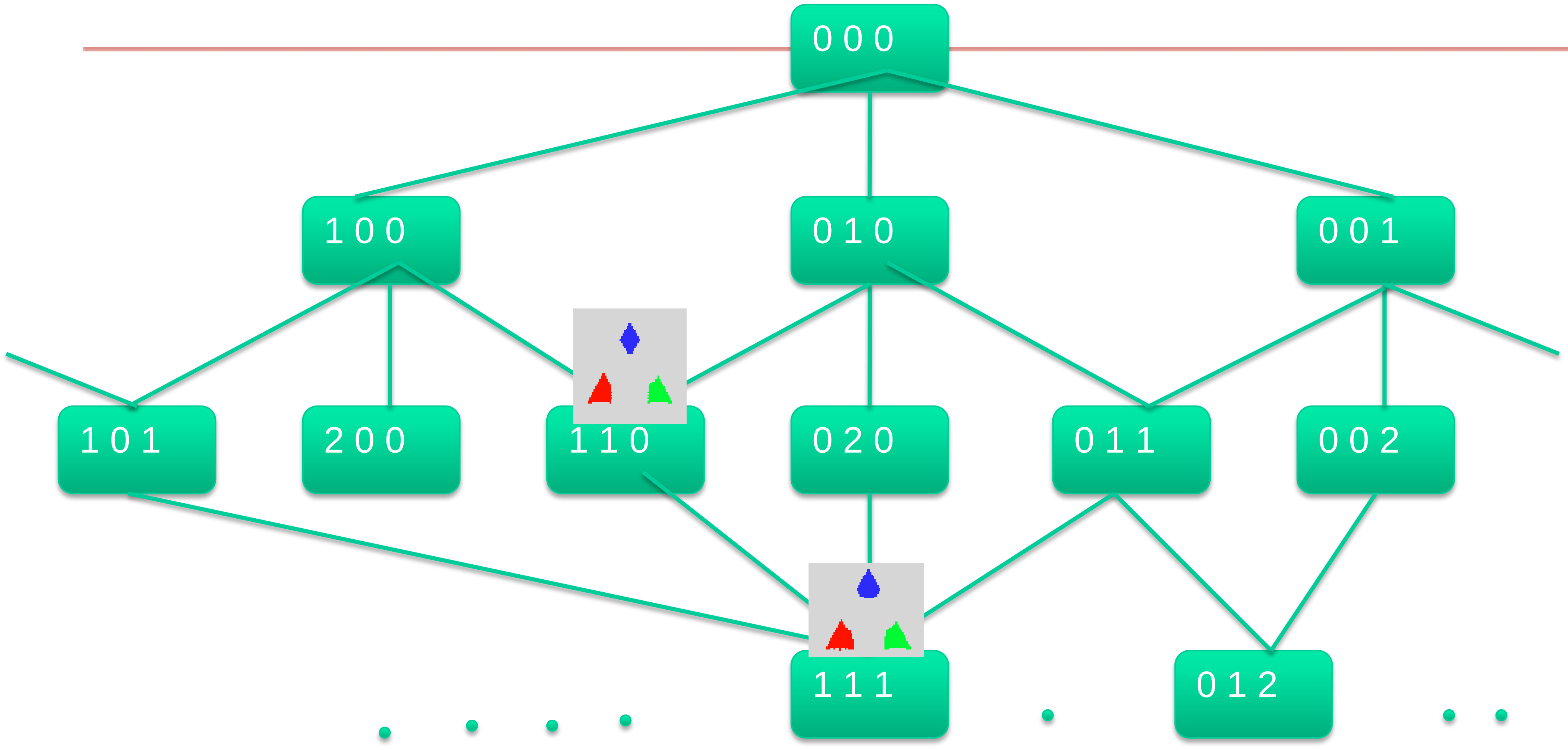
0 2 0 0
correct= 58%
accept= 15%
correct | accept = 100%

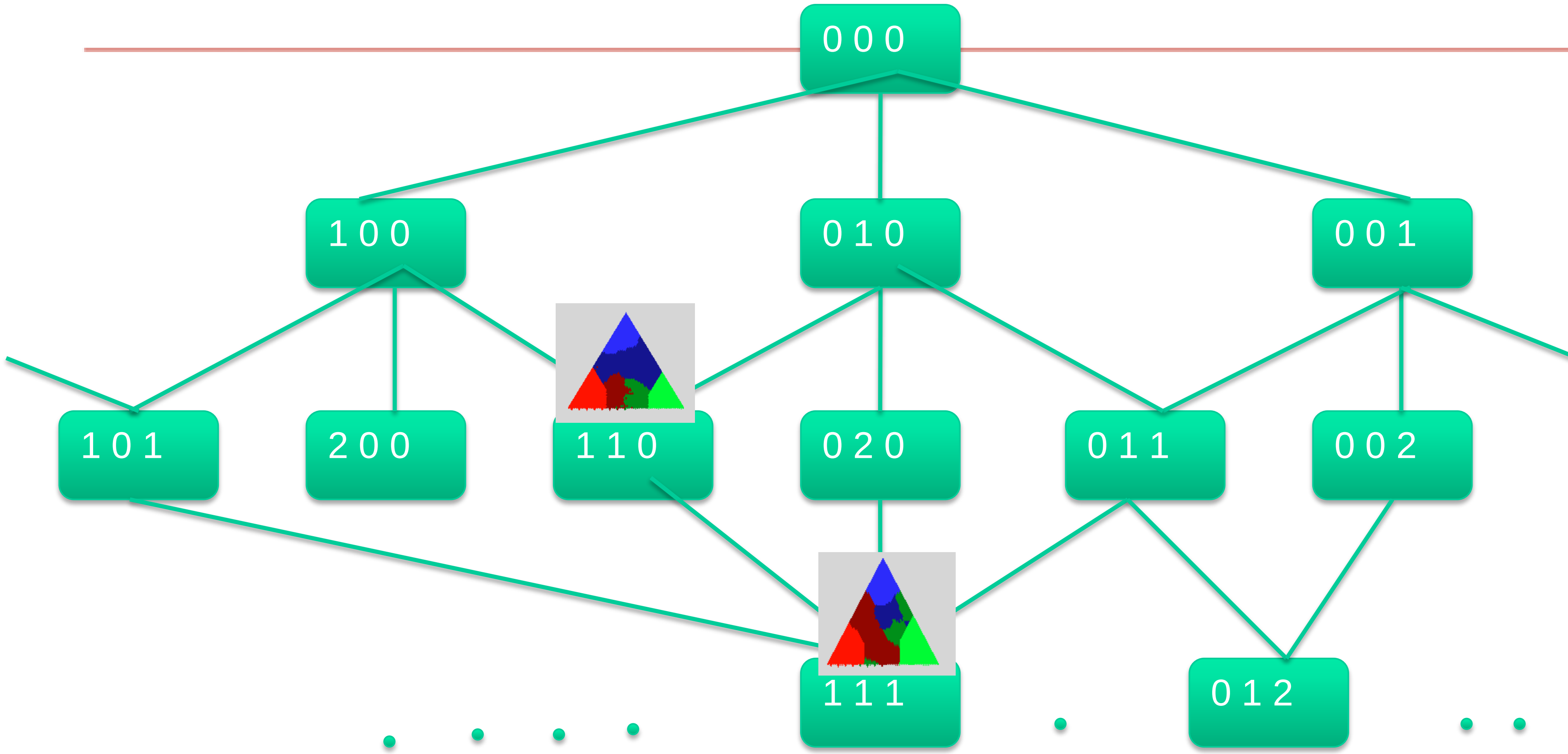
0 1 1 0
correct= 70%
accept= 17%
correct | accept = 98%

0 1 0 1
correct= 70%
accept= 17%
correct | accept = 98%

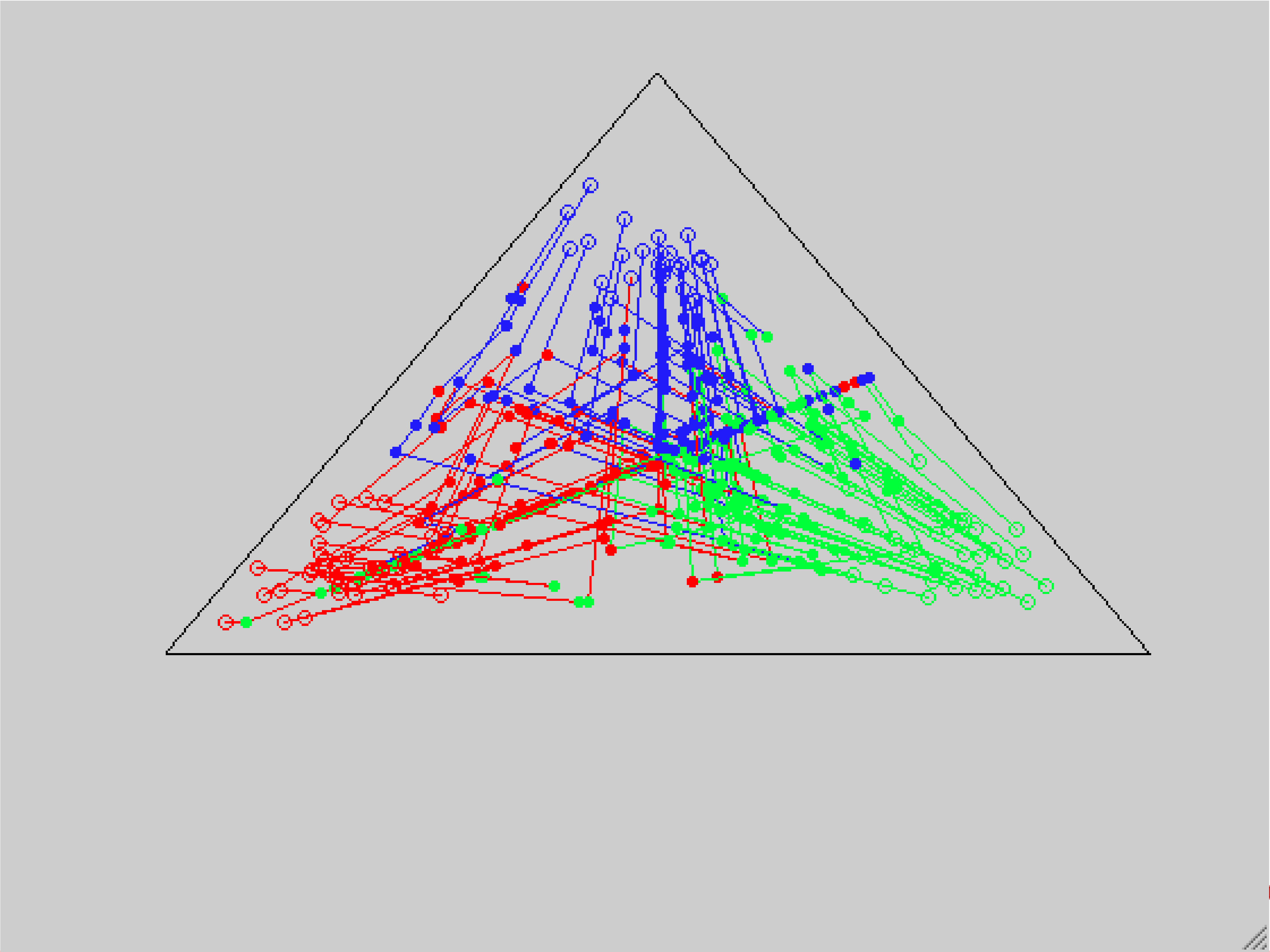
0 0 1 1
correct= 70%
accept= 17%
correct | accept = 98%

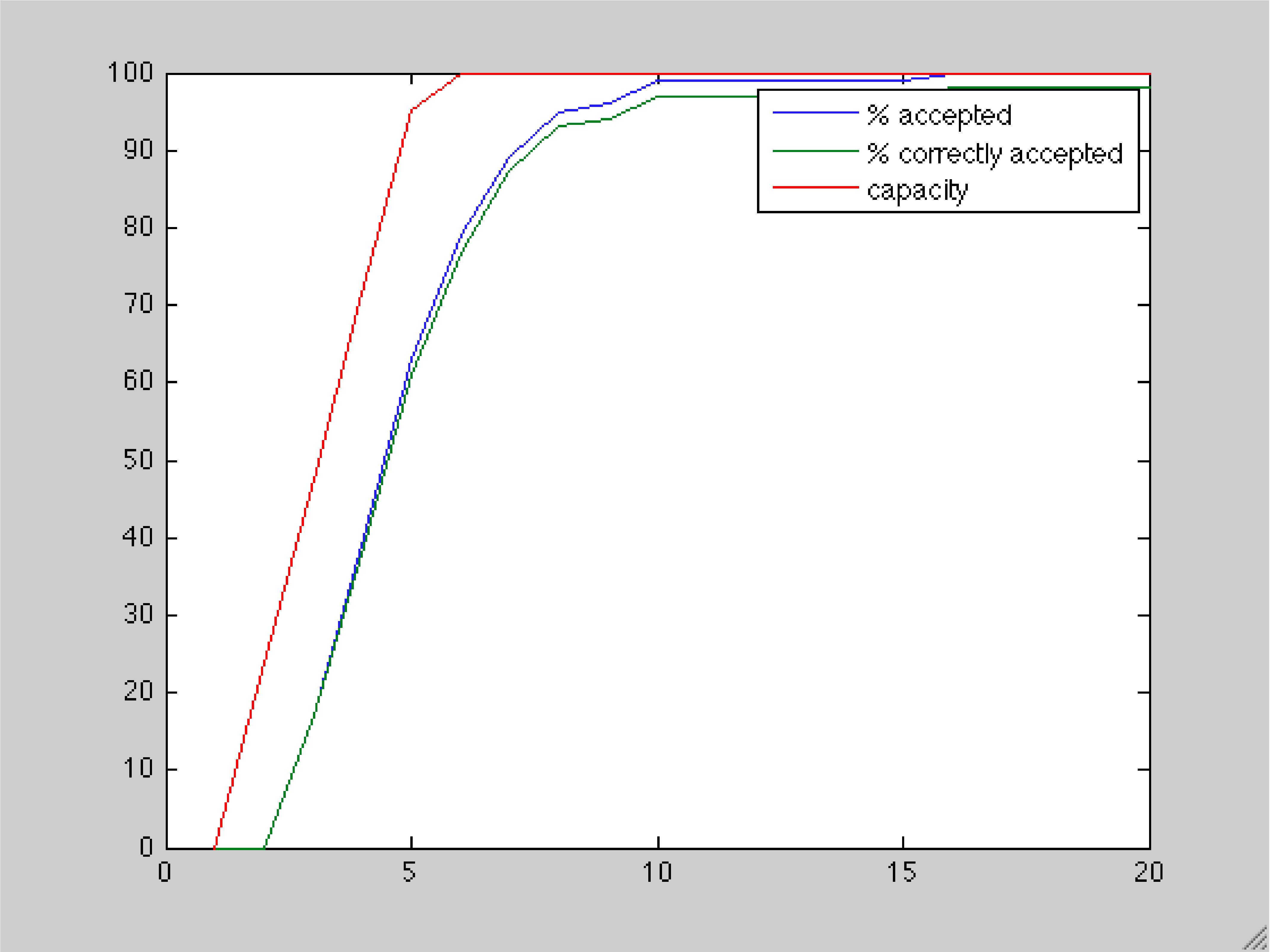
0 1 1 1
correct= 79%
accept= 38%
correct | accept = 97%


000
100
010
001
101
200
110
020
011
002
111
012


0 0 0
1 0 0
0 1 0
0 0 1
1 0 1
2 0 0
1 1 0
0 2 0
0 1 1
0 0 2
1 1 1
0 1 2

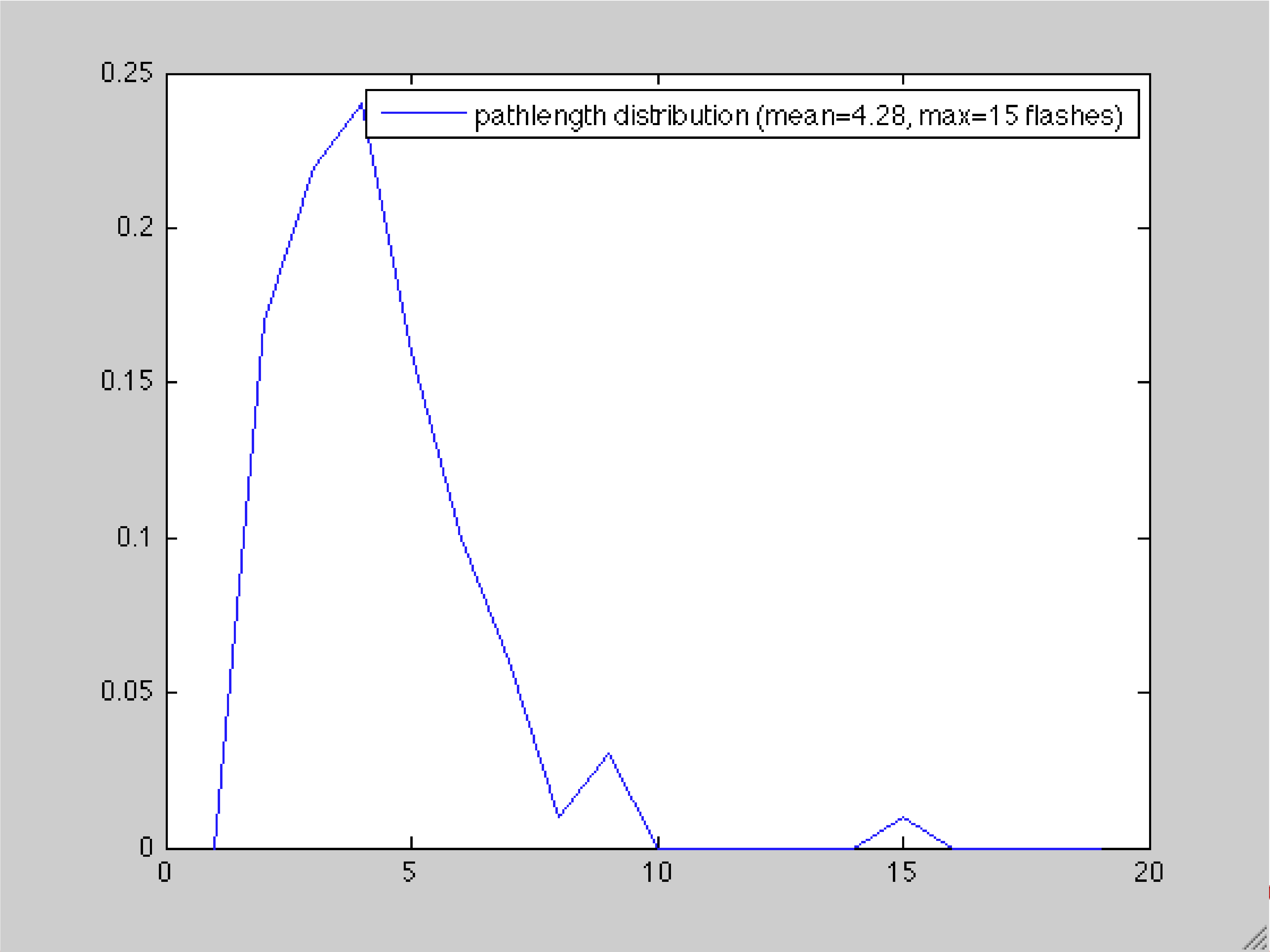
Results (adaptive)





100
90
80
70
60
50
40
30
20
10
0
0
5
10
15
20
% accepted
% correctly accepted
capacity

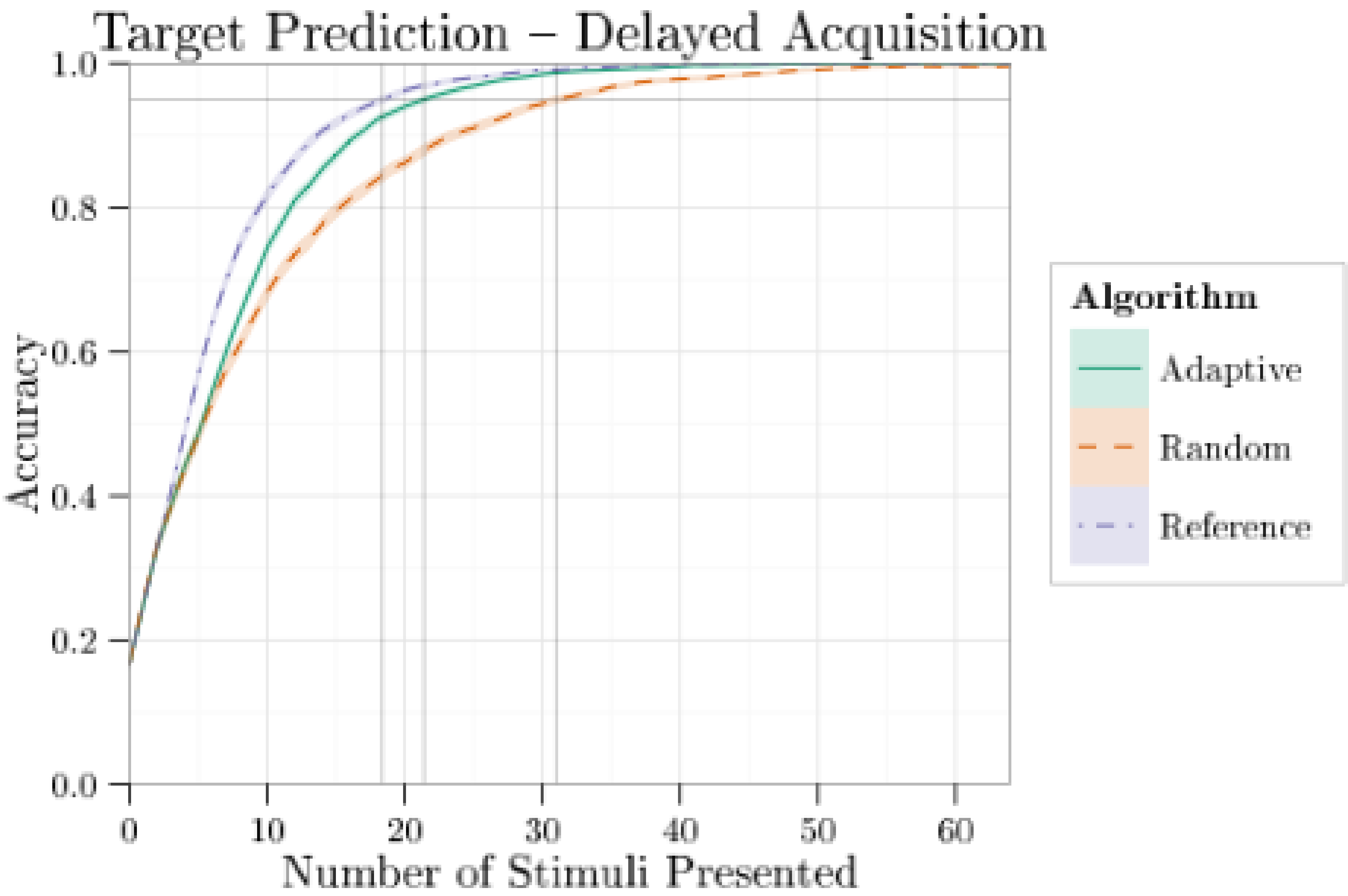
Results



megen

Results: Simulation full Row/Col

Based on empirical classifier output distribution of single flash



Conclusion

Improvement (confirmed in first full online pilot)

Optimality (guaranteed performance)

Optimization Criterion can consider pathlength

However, need to build in/model:

- Refractory period
- Delay
- Periodicity
- ...

Method not yet fully exploited because improvement larger for

- multi-flash
- large number of classes

Domain with large number of classes: words

Hypothesis:

Assume a target word is active (kept in mind)

Presenting a related word gives a detectable response

Like flashing a row gives a detectable response on all targets in that row

There is a very simple relation between row/col and letter in matrix

There is a less systematic (and more sparse) relation between words (associated or not)

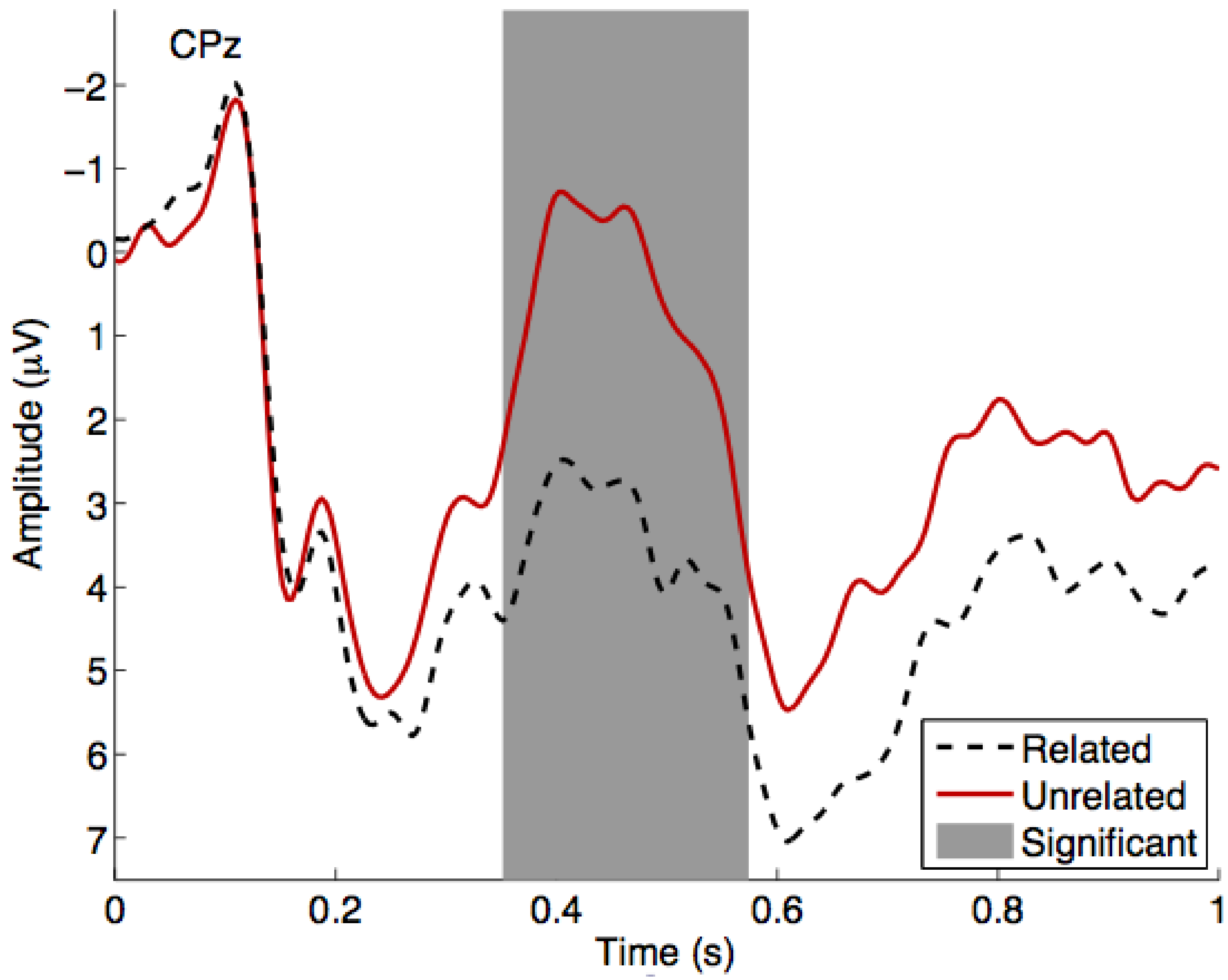
Semantic priming

(non)associated word pairs fom Leuven Database

Prime		Probe
	Unrelated	
tang (pliers)	-	opbrengst (yield)
berg (mountain)	-	drankje (small drink)
eland (moose)	-	eerbied (respect)
rog (ray)	-	maaier (mower)
	Related	
mier (ant)	-	klein (small)
tram (tram)	-	spoor (track)
racket (racket)	-	tennis (tennis)
naald (needle)	-	draad (thread)

Semantic priming

EEG contrast



Semantic priming, study 1

12 subjects, 400 word pairs

Classification rate 60% (+/- 7%)

Careful matching for word frequency, length etc

Accurate Association Database, needed checking

Semantic priming

Can we detect which word subject has in mind using this paradigm?
  a la 20 questions.

Universe
  Target word (belief state dimension)
  Probe word

Present probe, classify: update belief, (non)associated targets up(down)

Probe selection

  Random

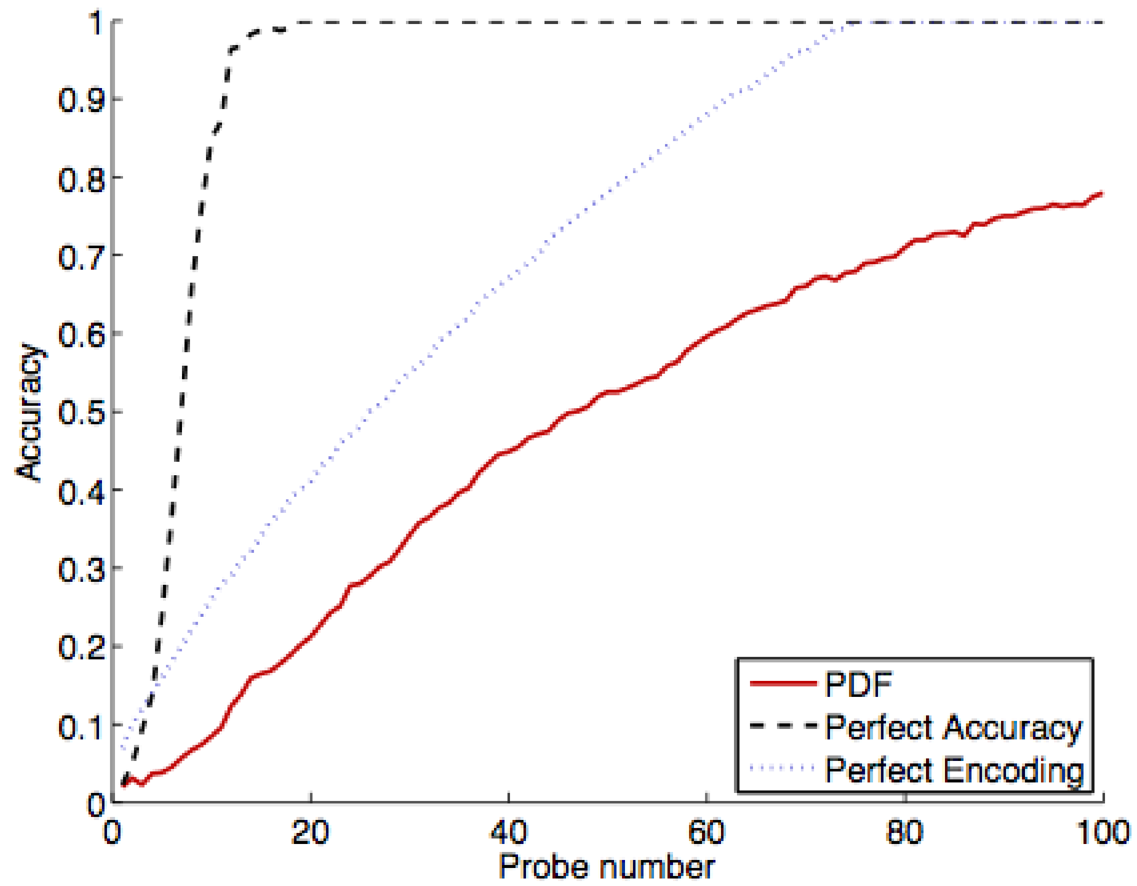
  Ordered (most informative first)

  Adaptive dynamic selection (rule)

Simulate

Semantic priming, simulation, 100 word universes

Words with most associations
Ordered probes



Semantic priming, simulation, word universe size

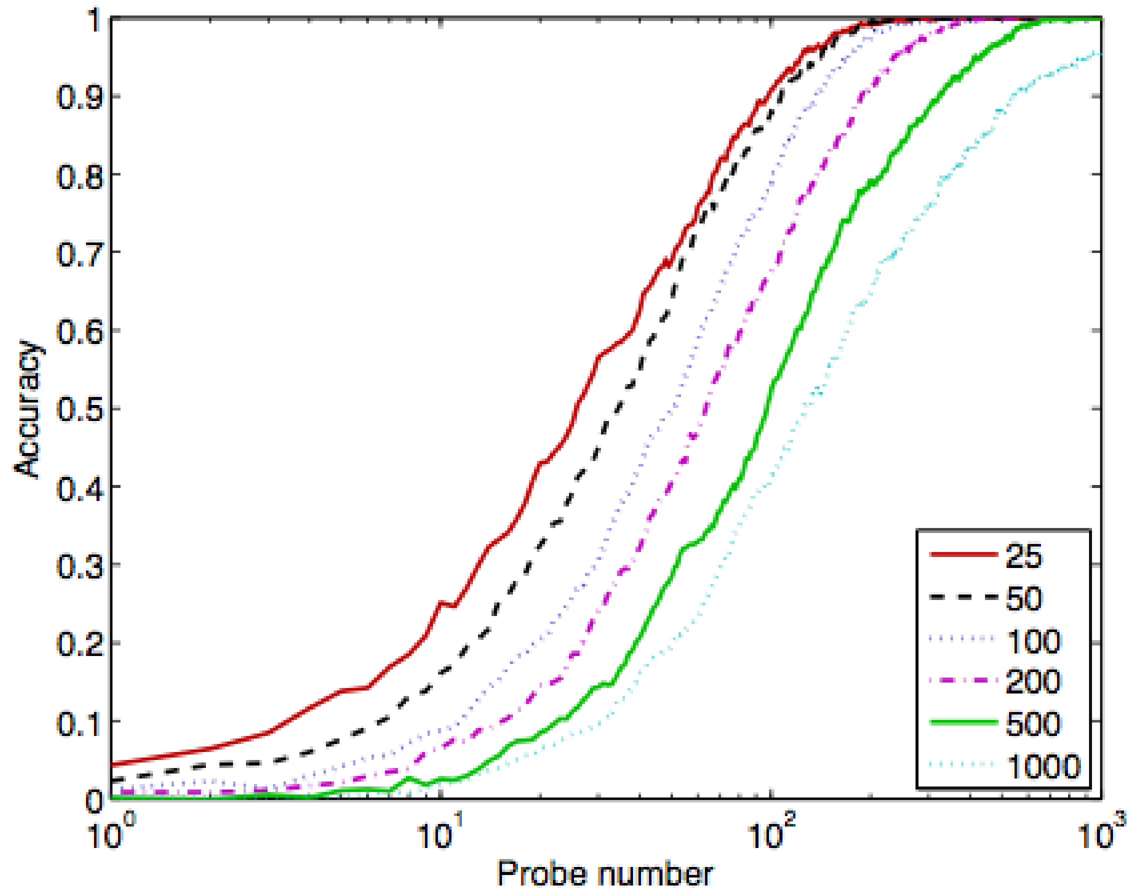


Figure 6. Simulation results for different size universes.

Adaptive semantic word probing

Adaptive model or local rule
  big improvement (large spaces feasible, scales well)

Good databases needed

Adaptation of system can still bring us a lot

Adaptation of user

Desired non-stationarity (learning)

Traditional neuro-feedback (abstract marker, no task instruction)

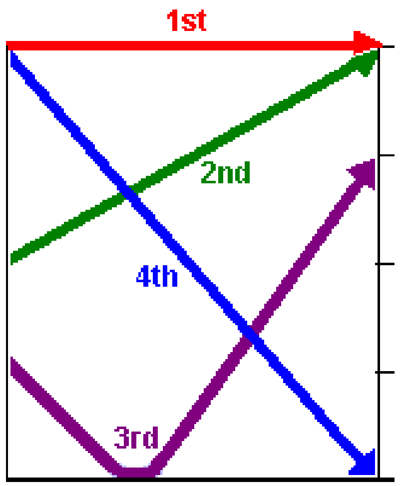
BCI learning, adaptation to output (given task)

BCI supported training of perceptual categories (given marker)

Categorical perception

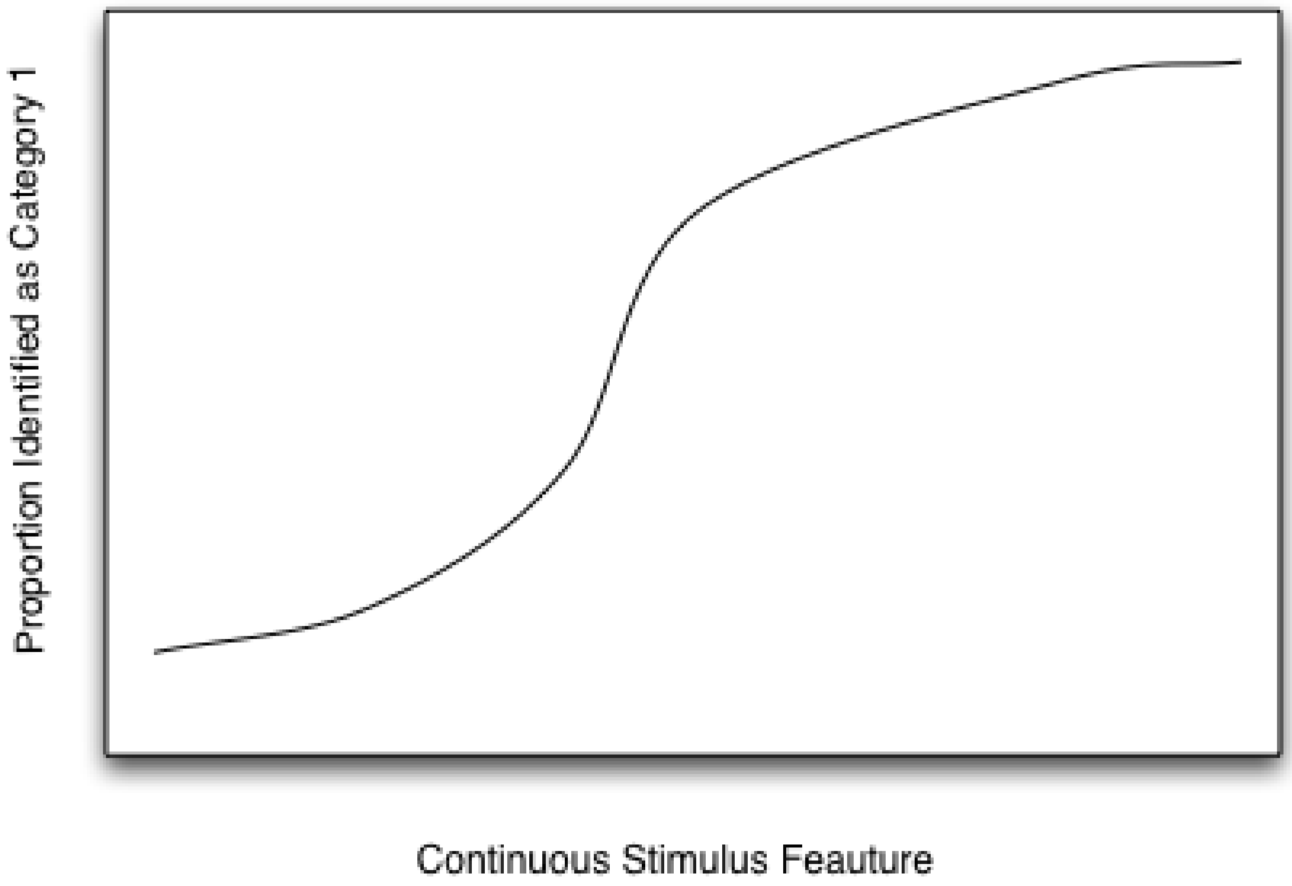
- Continuous variation in stimulus
- Example: speech: d vs t
- Discrete mental representation
- Information lost: efficiency of coding and representation
- Very basic, pervasive process in perception/cognition
- In speech

  pa – ba, assu – asu, r – l, mandarin tones, vowels, ...



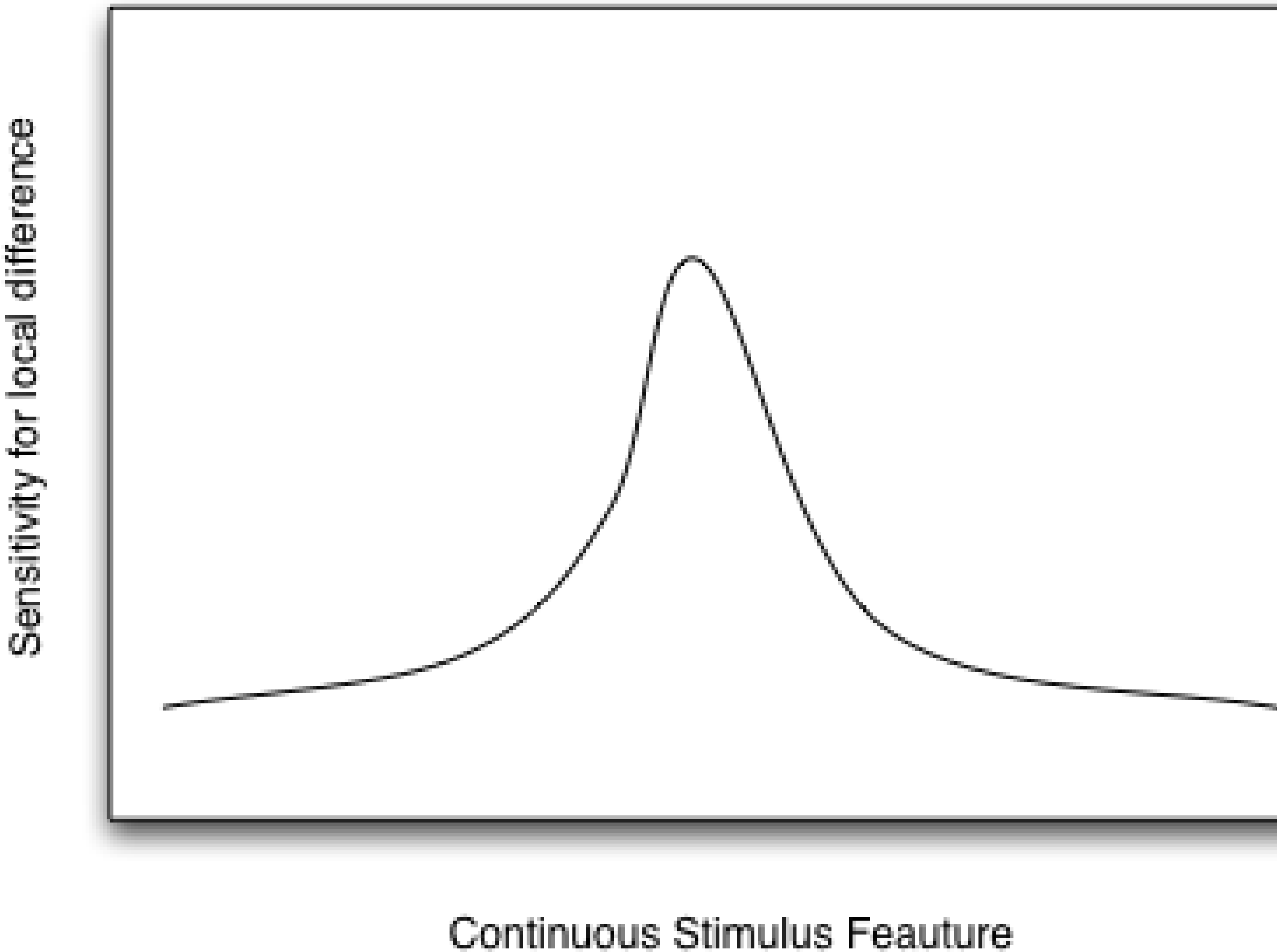
Identification

- Recognize and 'label' stimulus

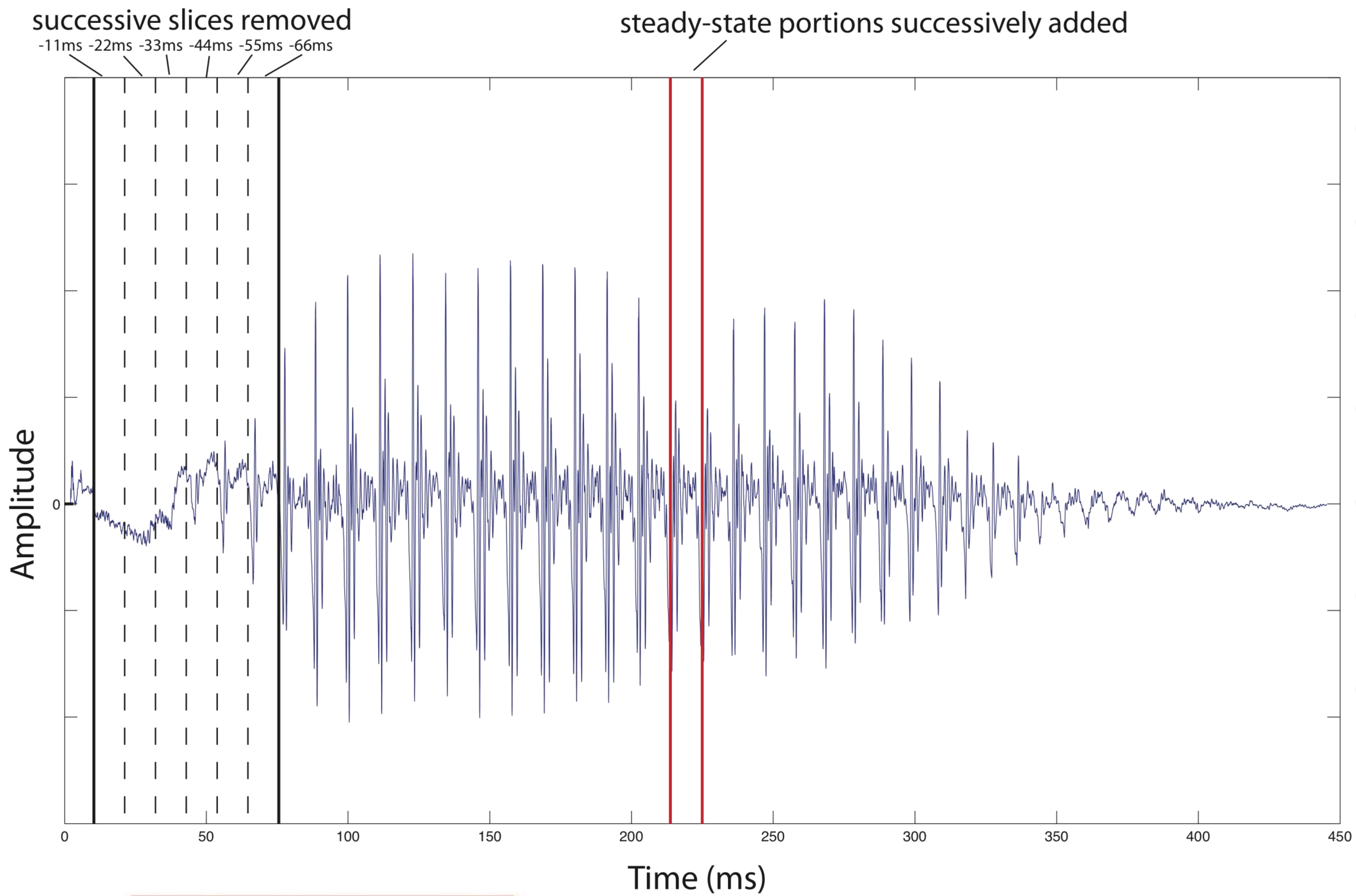


Discrimination

- Sensitivity for differences



Experiment Ba vs Pa; Stimuli



/pa/ - 85 ms VOT

74 ms VOT

63 ms VOT

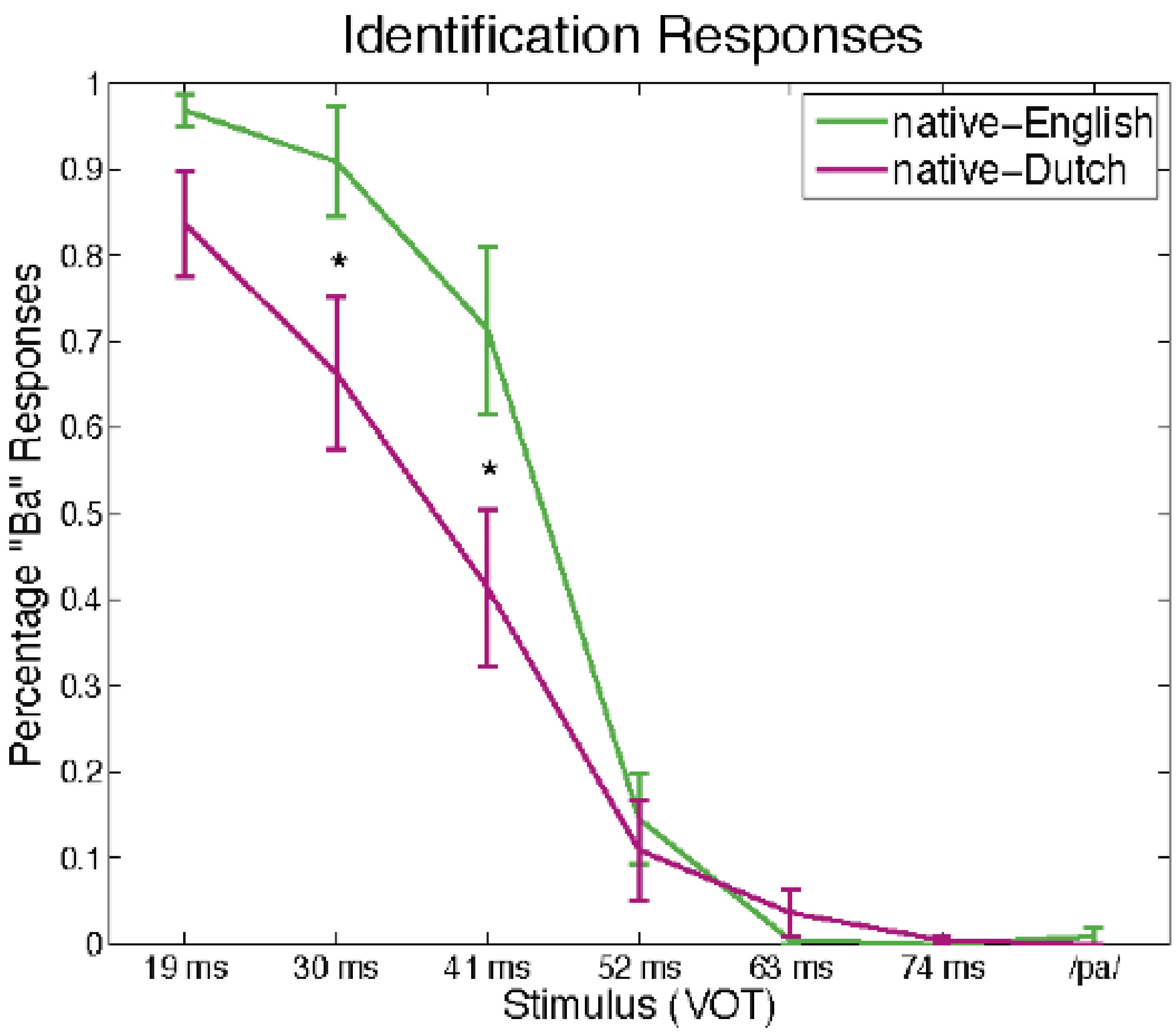
52 ms VOT

41 ms VOT

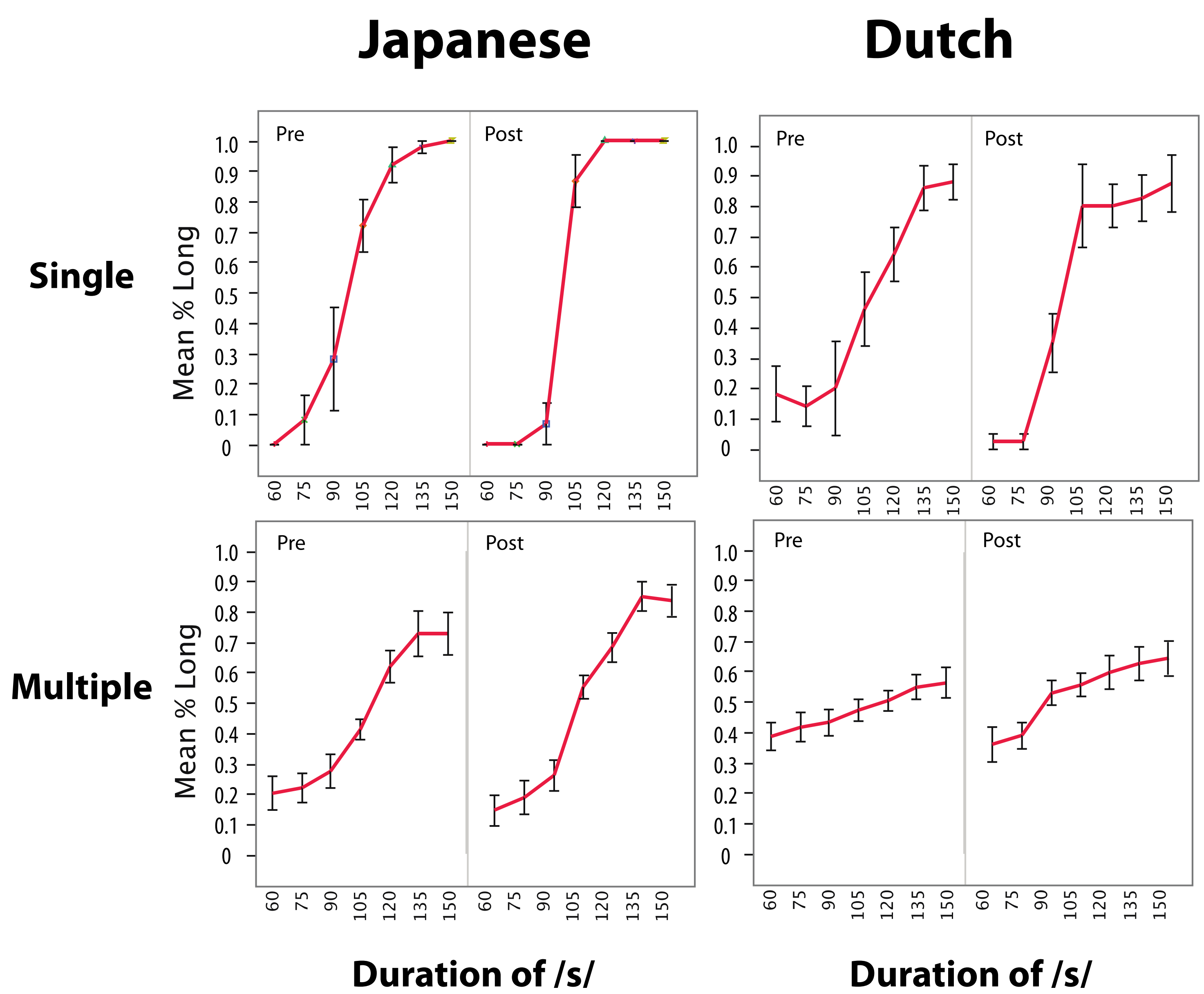
30 ms VOT

19 ms VOT

Behavioral Identification ba vs pa

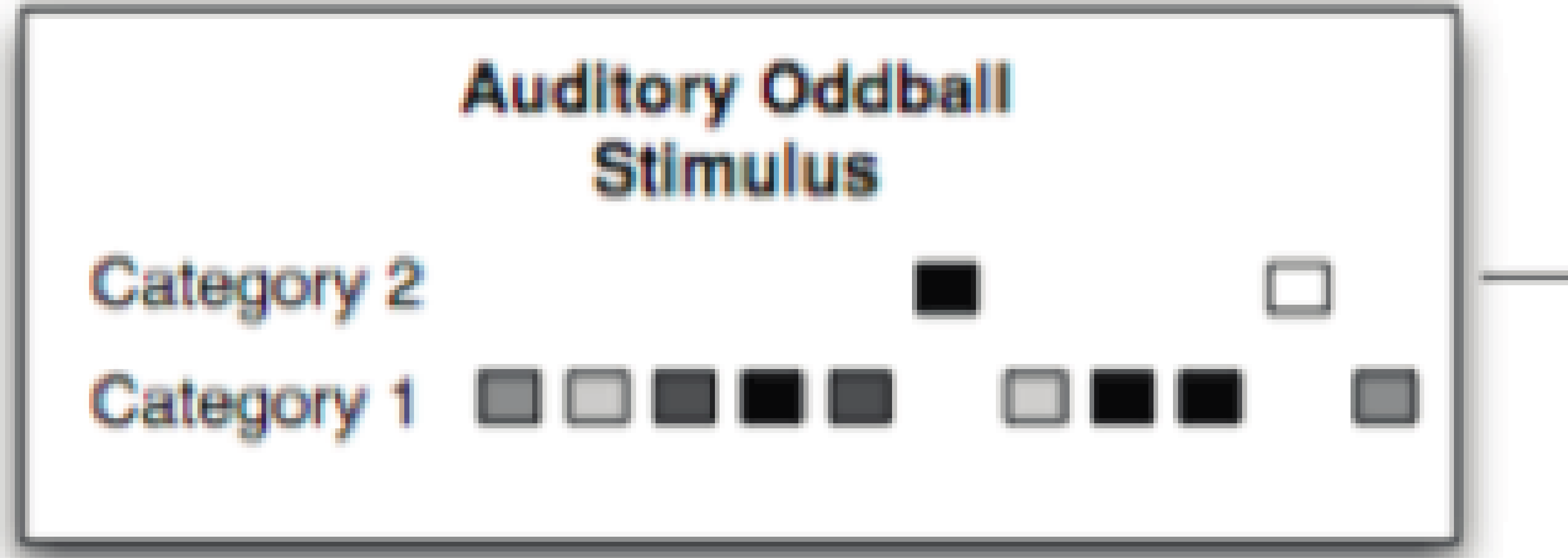


Behavioural Identification s vs ss (before / after exposure during experiment)

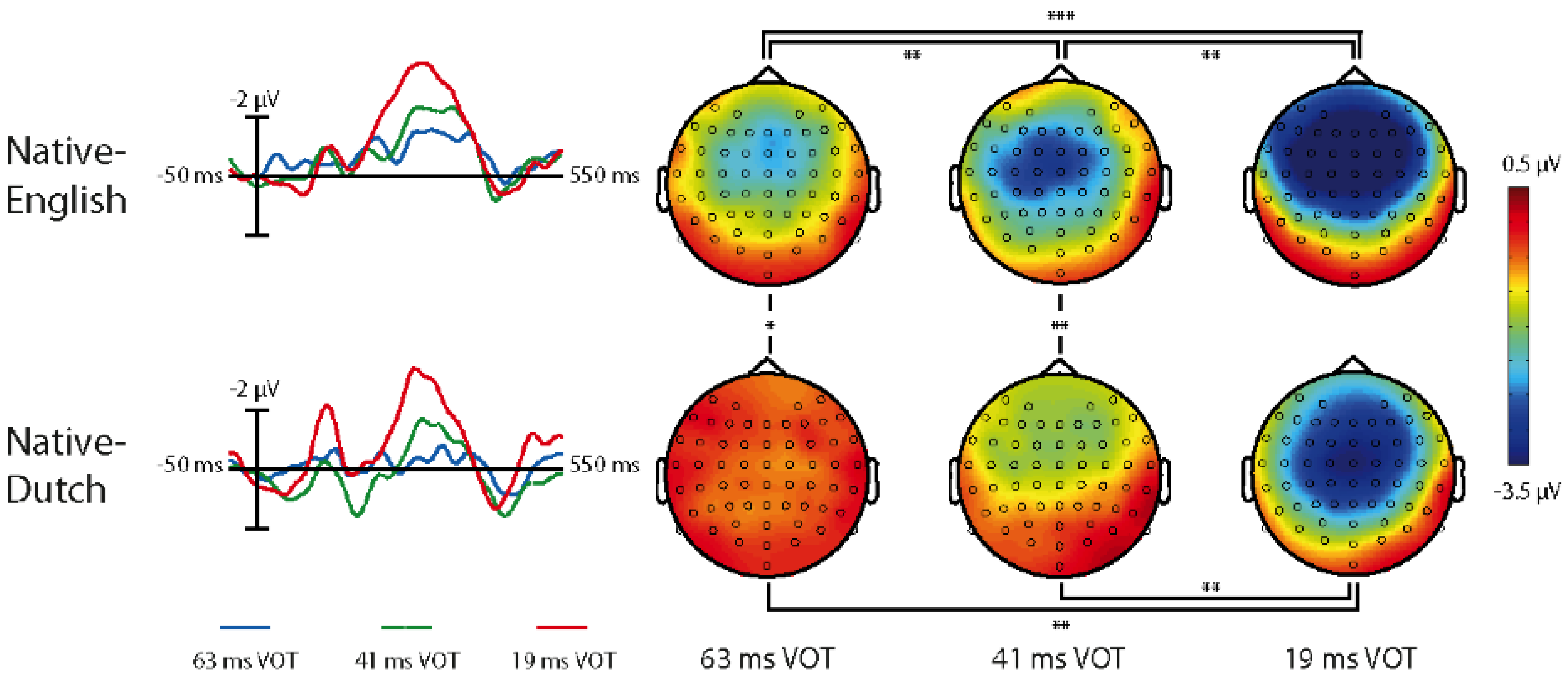


Discrimination test with EEG, Oddball sequence

- Difference perceived? -> Mismatch Negativity Response
- Pre-attentive
- Even present before behavioral response



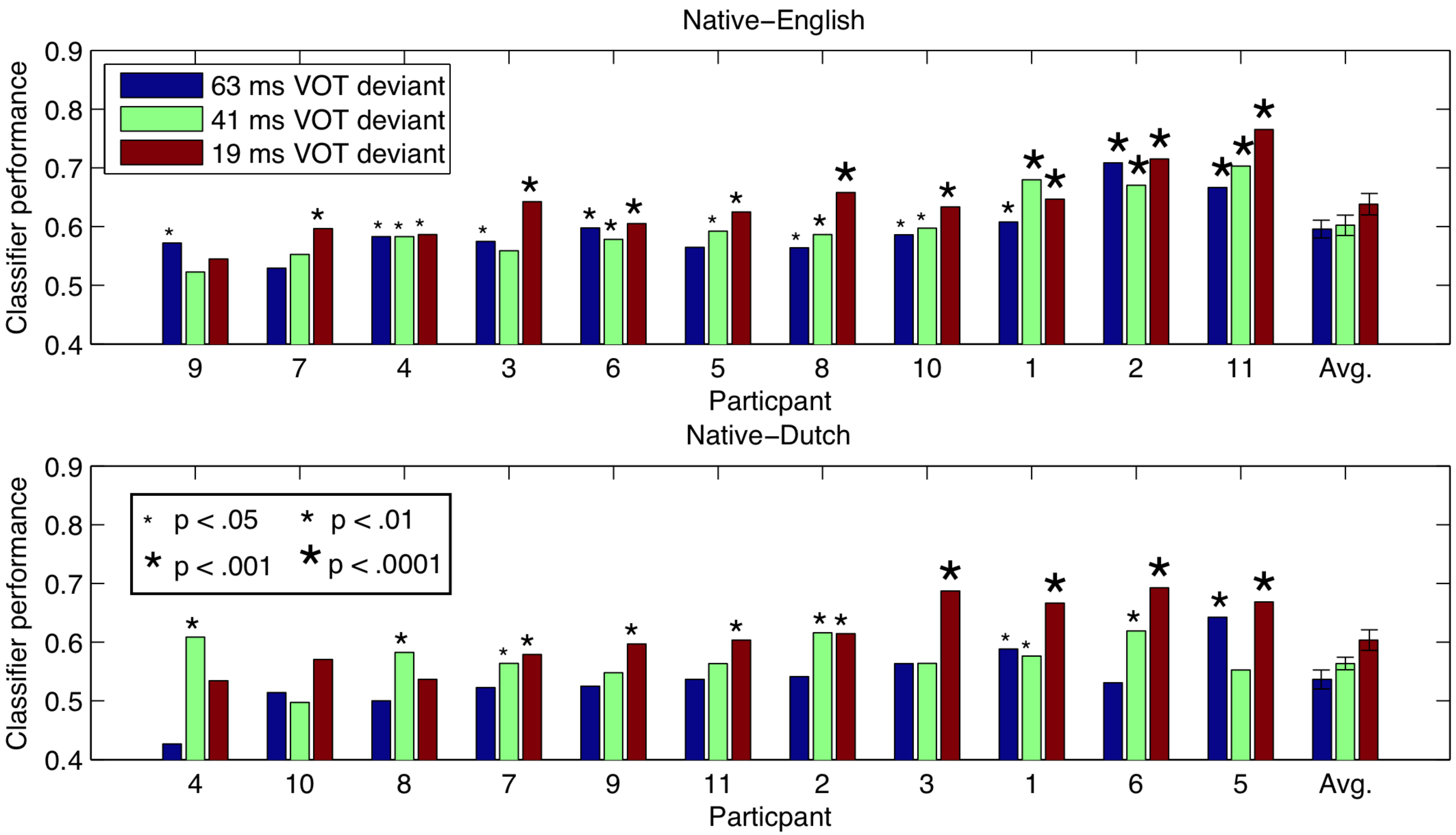
Results: Mismatch negativity responses



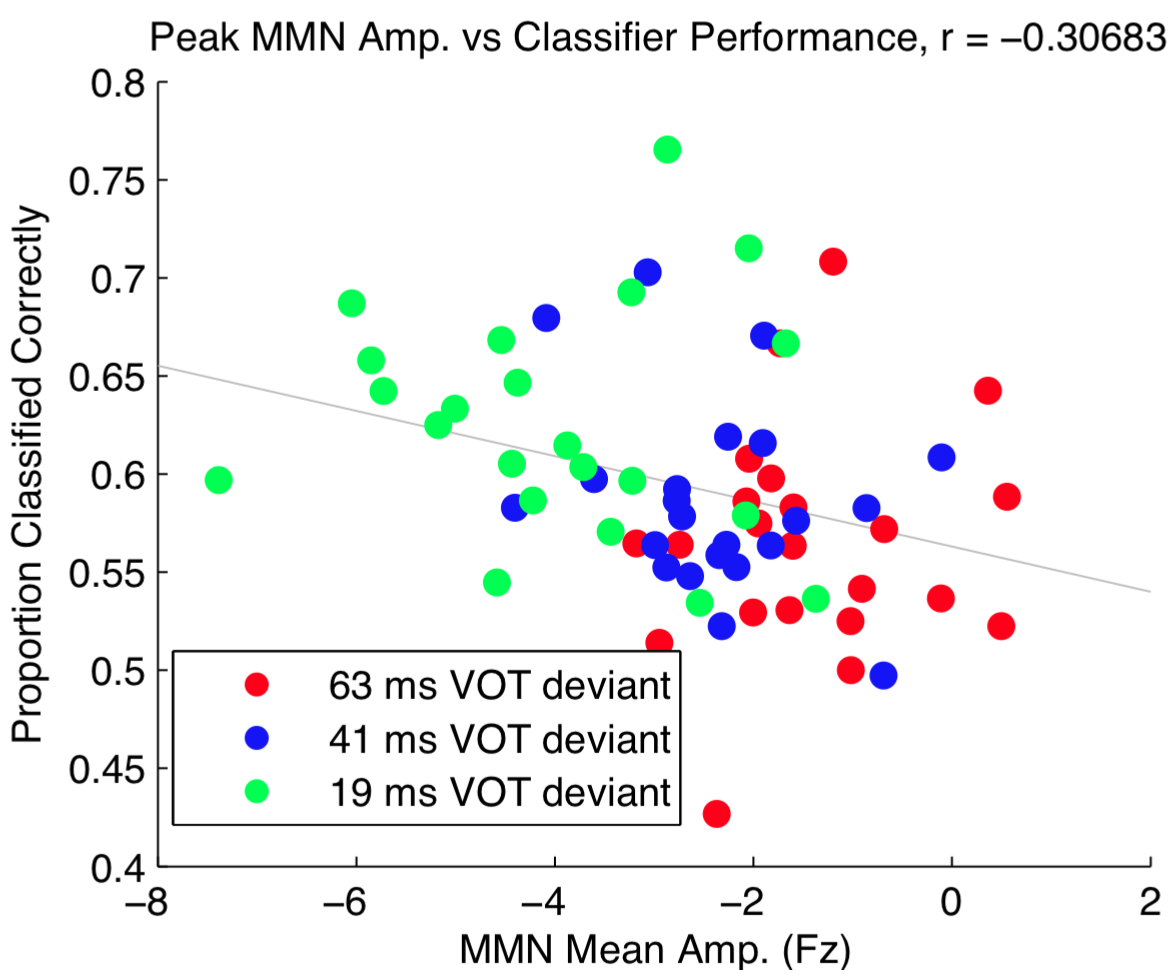
Single trial detection?

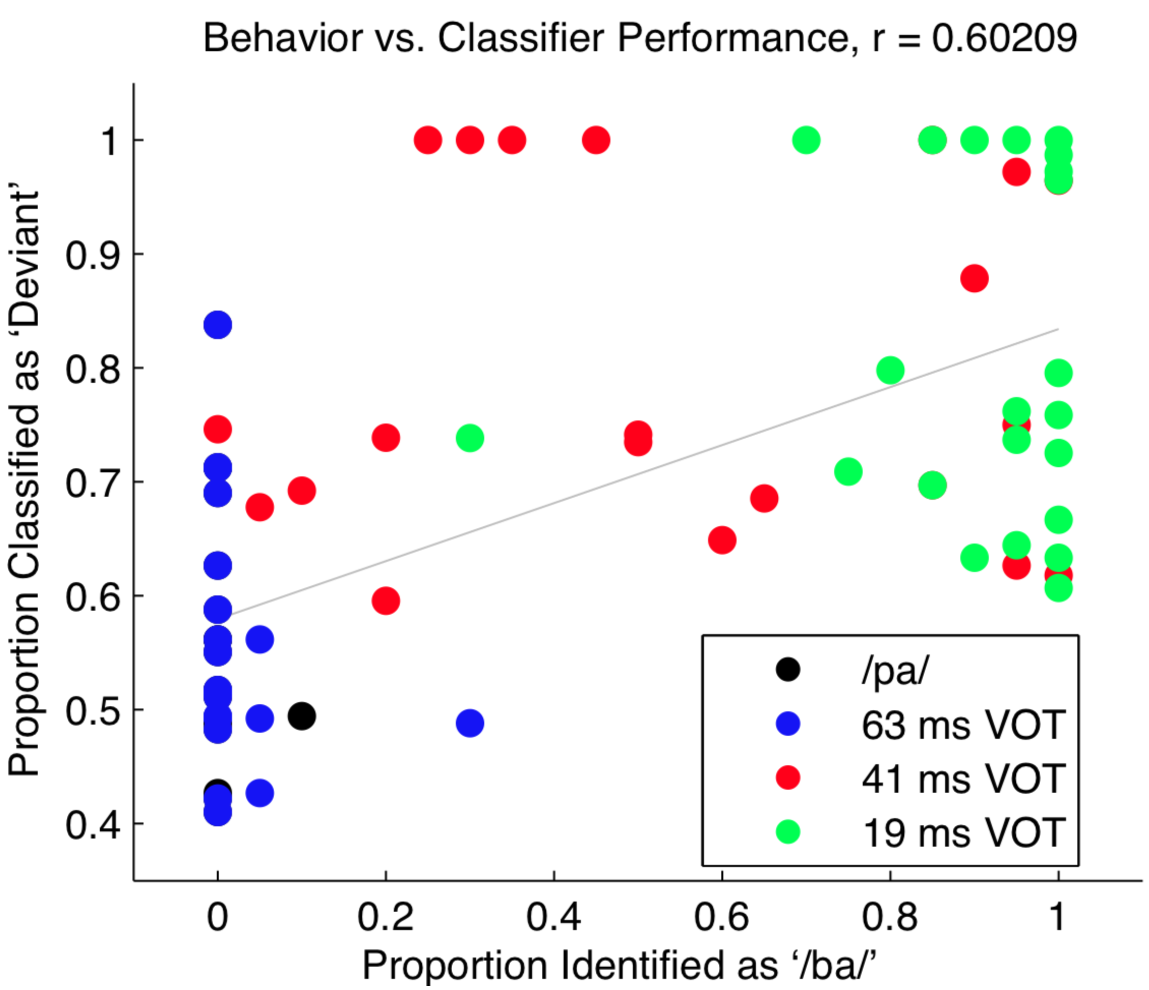
- Discrimination
  - pre-processing
  - classifier
  - cross validation
  - ... all standard BCI practice

Within-participant classification analysis ba vs pa

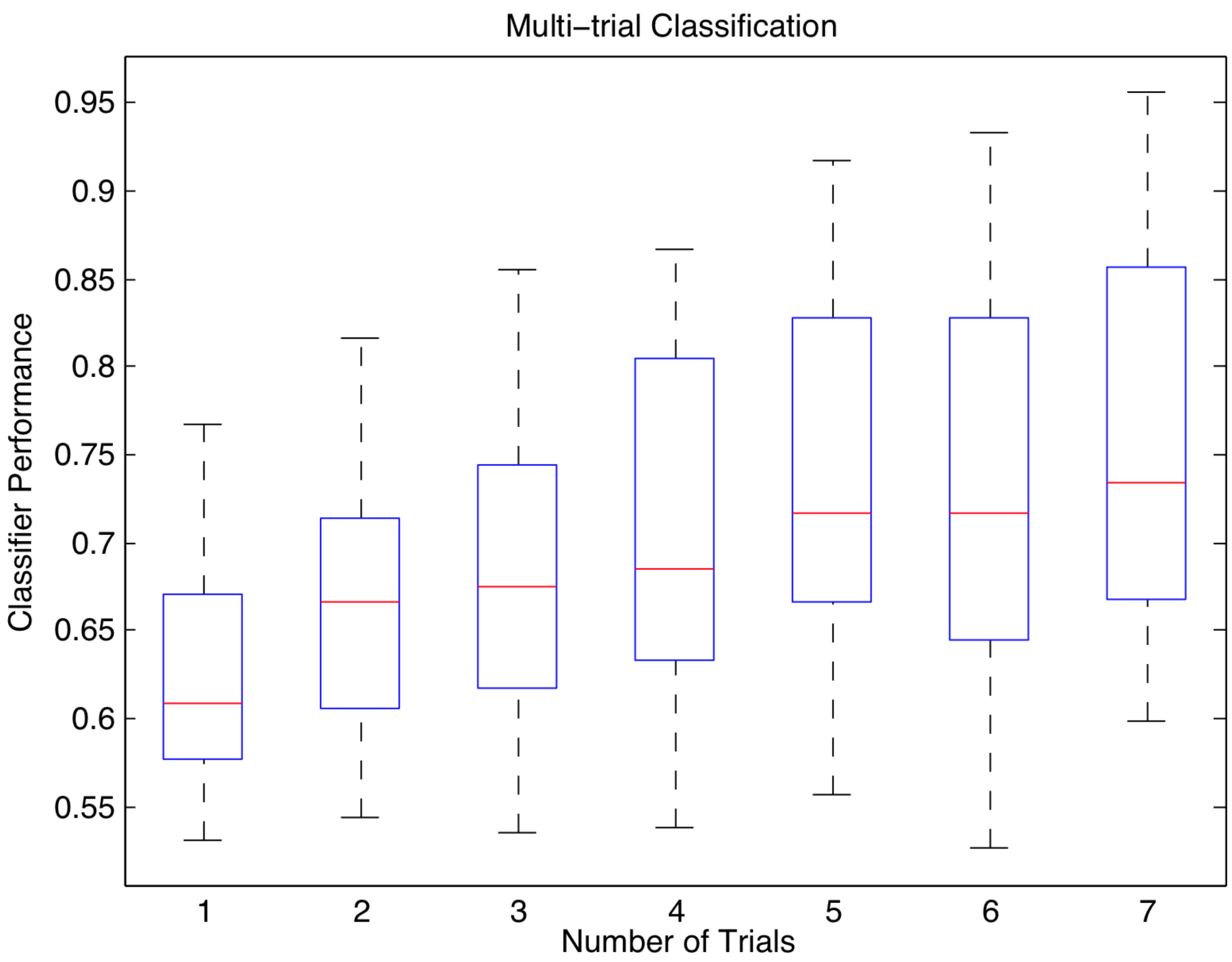


Results: Relationship to ERP and behavioral measures

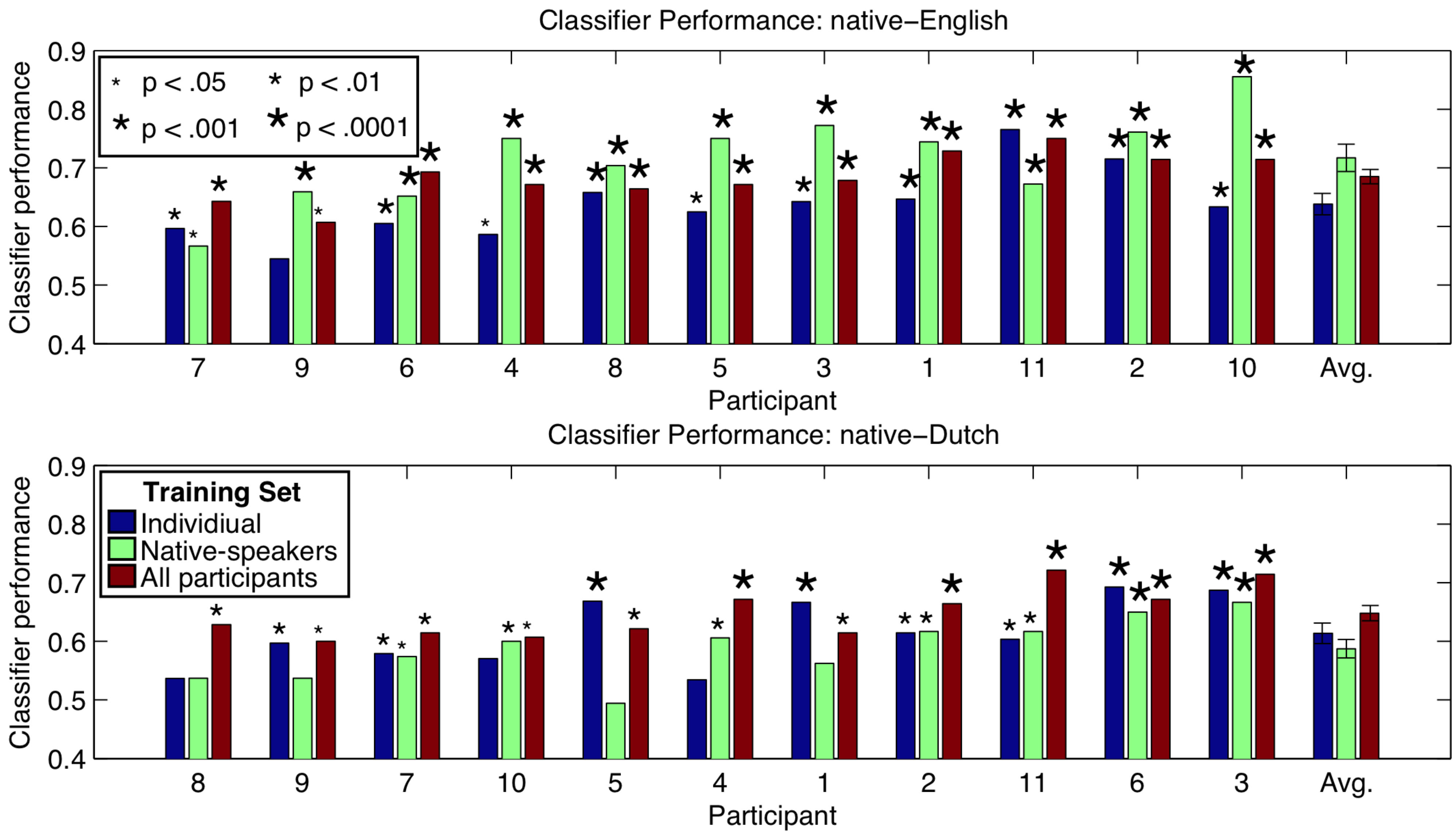




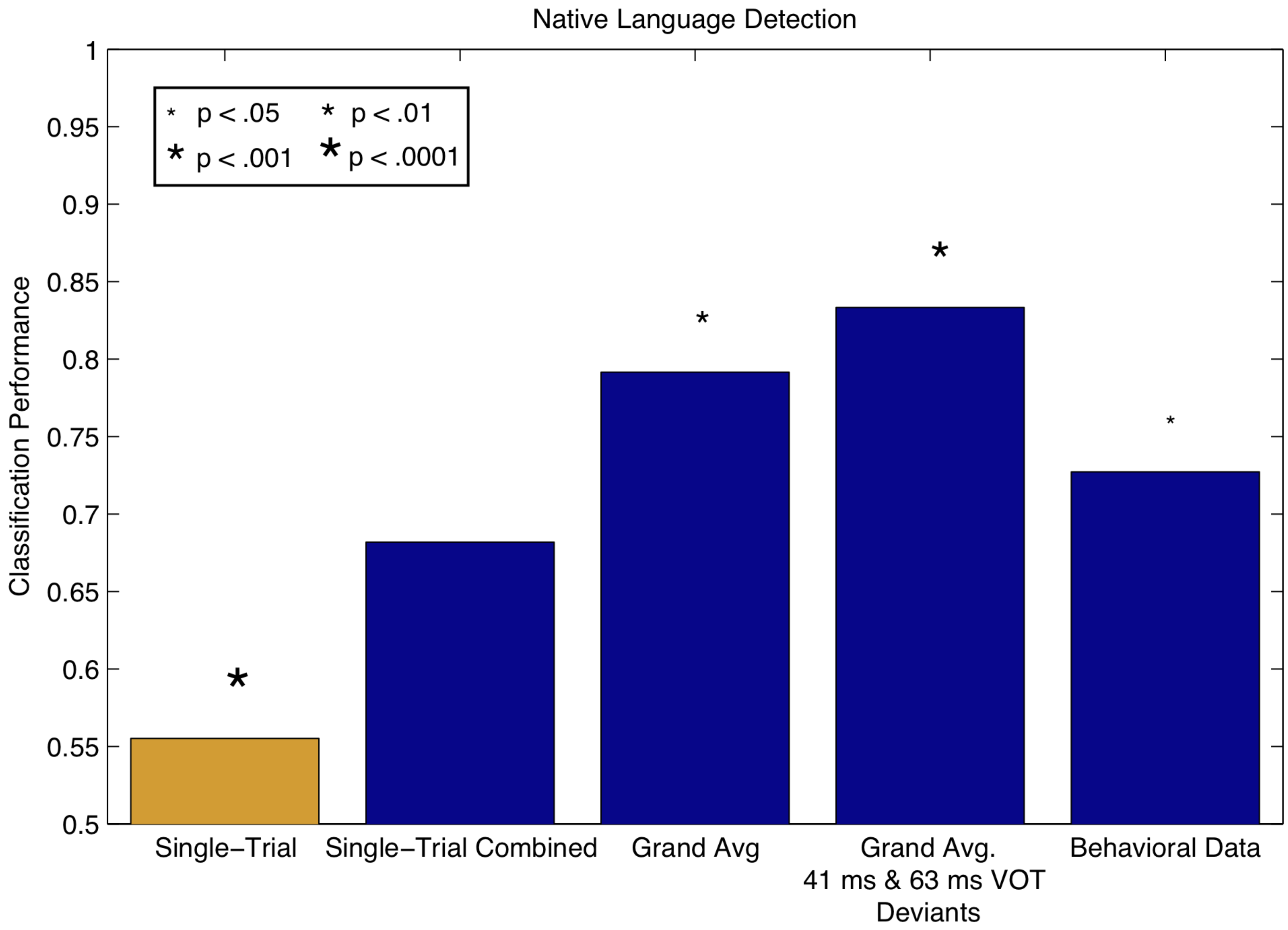
Results: Multi-trial classification performance



Results: Cross-participant classification analyses



sideline: Ba vs Pa Detection of native-language EN/NL

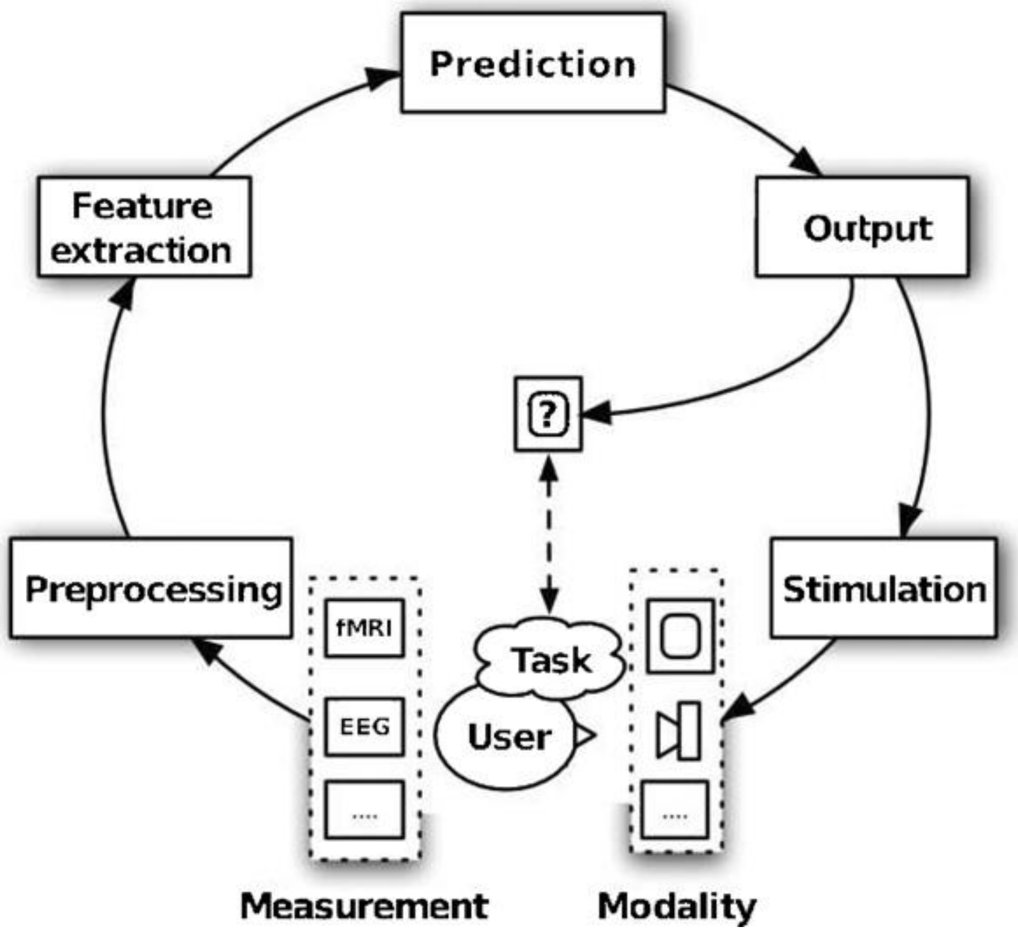


BCIs & Language Learning

1. Train classifier (spatial filter) on (MMN) response to large, already perceivable contrast

2. Use this classifier to identify (MMN) response for smaller contrasts

3. Based on classifier-performance (i.e. how well we can isolate an MMN/whether there is an MMN), adjust stimuli

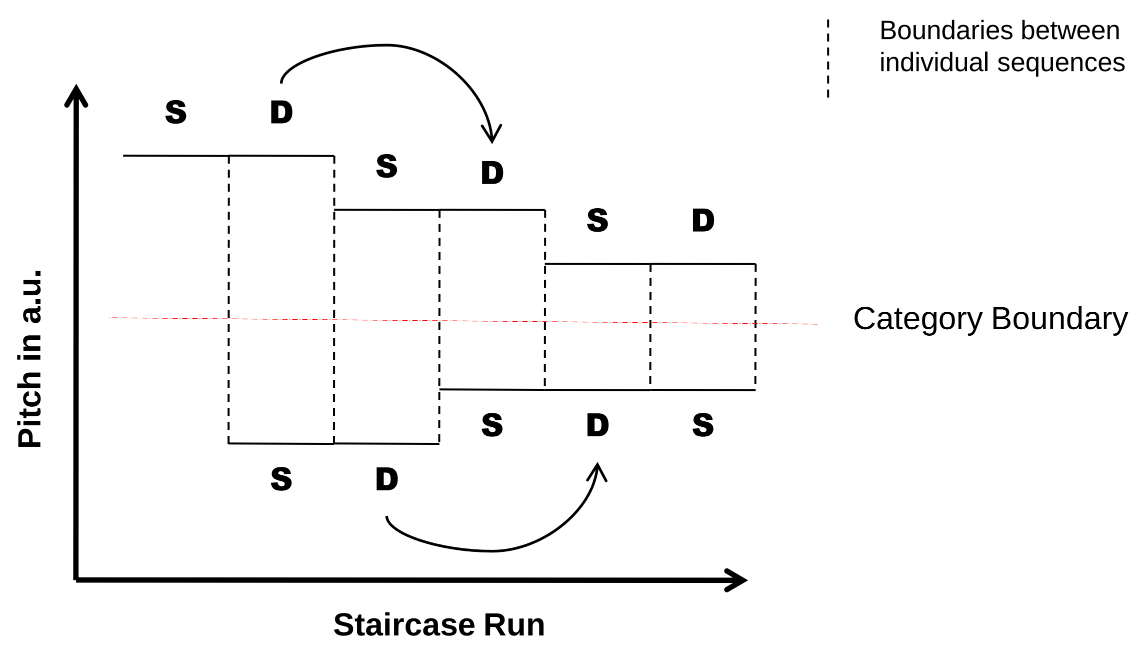


Use of feedback

- Reward (movie blur)
- Determine practice time on task, criterium
- Adapt stimuli

Visual feedback, movie blur

Methodology: staircase stimuli



Conclusion, on BCI for perceptual category learning

- (Speech) categories can be 'probed' with oddball MMN
- Single (few) trial detection is possible
- Even before/better than behavioral testing
- Use in online BCI setting for training (L2) categories
- **Is it more efficient than exposure or behavioral training?**
  - **3 studies running to test**

- Classify deviants and standards (within category discriminations)
  don't reward bias
- Discrimination (MMN) -> Identification (P300)

- Applicable to many domains (music !)

Conclusion

Adaptation is good

There are many ways to exploit it in the BCI cycle

Thanks

Galin Bajlekov
Alex Brandmeier
Jason Farquhar
Marcel van Gerven
Jeroen Geuze
Christian Hoffman
James McQueen
Makiko Sadakata
Loukianos Spyrou