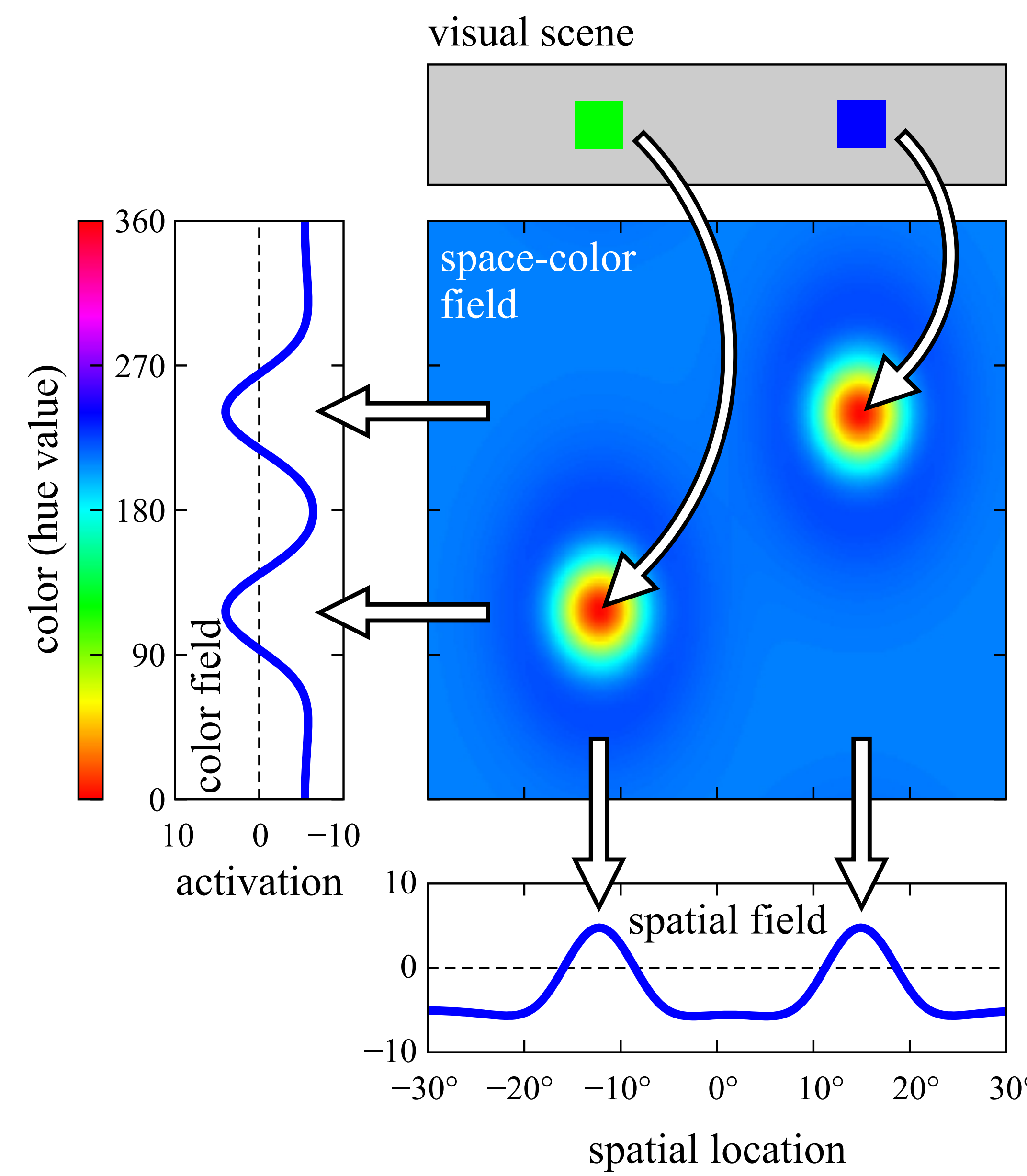
Grounding Spatial Language:

A case study in Dynamic Field Theory as a framework for neurally grounded architectures for higher cognition

Gregor Schöner
Institut für Neuroinformatik
Ruhr-Universität Bochum, Germany
gregor.schoener@rub.de

Previous lecture

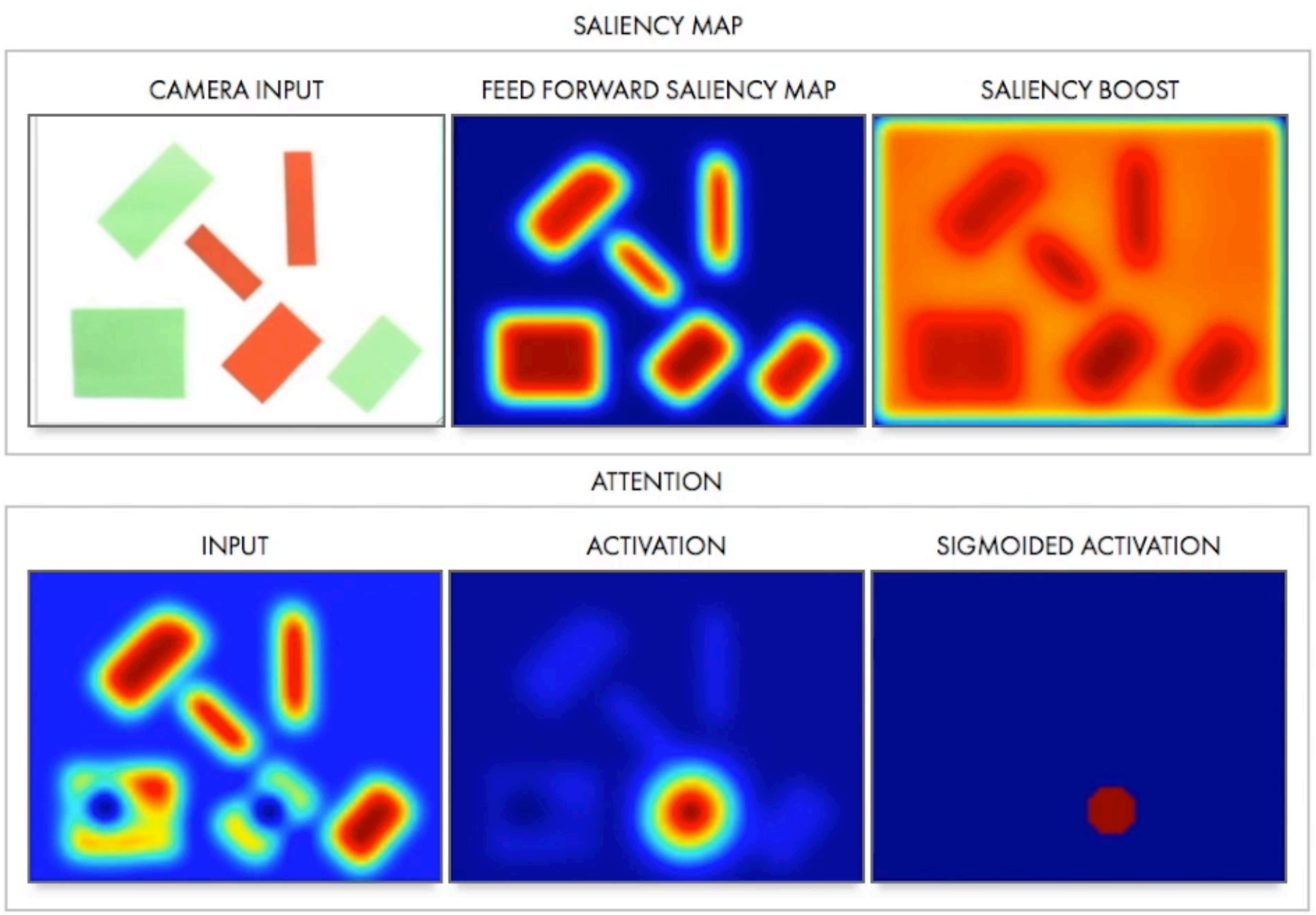
- higher-dimensional neural fields enable new functions
  - visual search… attentional selection
  - coordinate transforms…
  - binding different feature dimensions through space

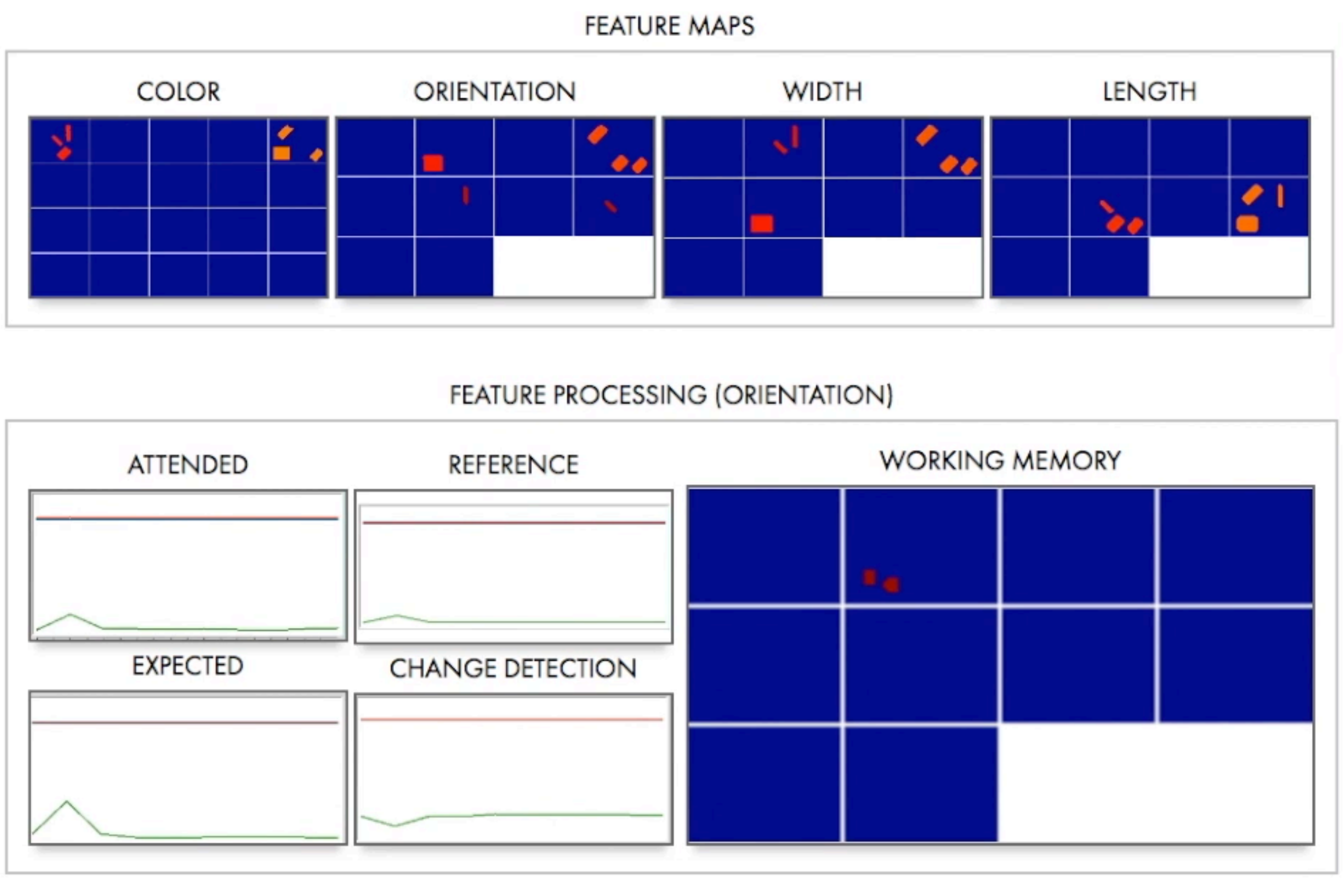


Previous lecture

- => scene representation

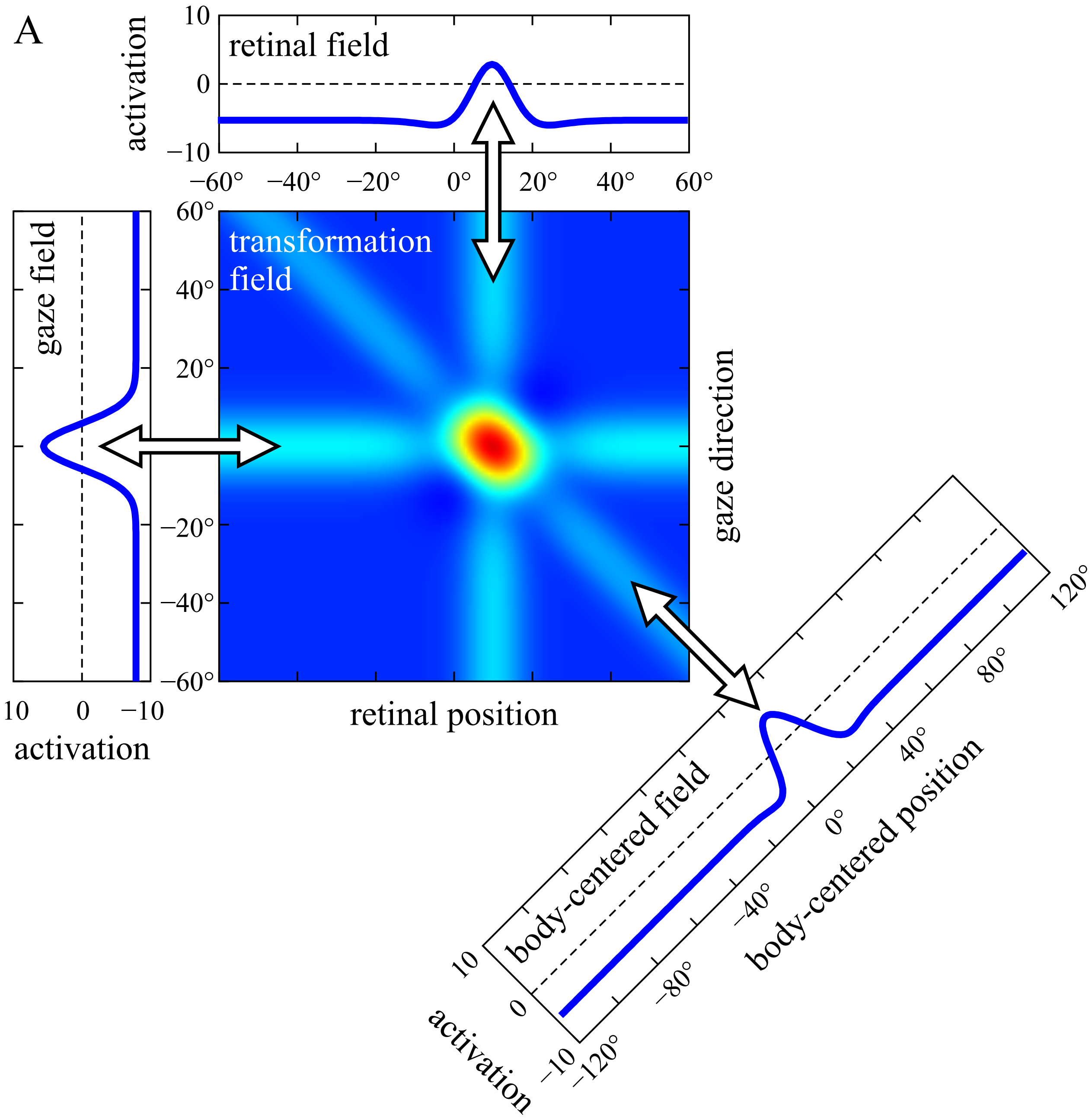
[Grieben et al, (Attention, Perception, & Psychophysics, in press)]





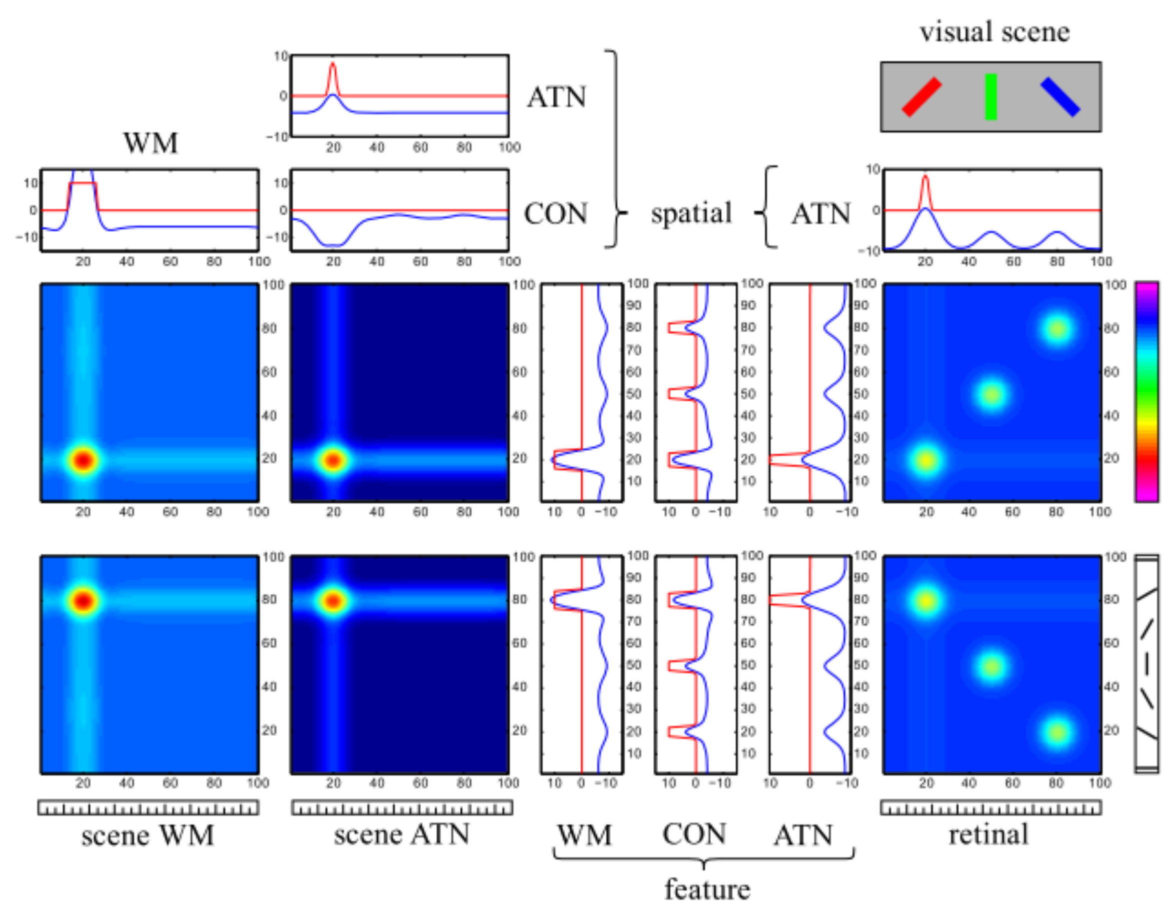
Previous lecture

- higher-dimensional neural fields enable new functions
  - visual search… attentional selectic
  - coordinate transforms…
  - binding different feature dimensior through space



Previous lecture

- higher-dimensional neural fields enable new functions
  - visual search… attentional selection
  - coordinate transforms…
  - binding different feature dimensions through space

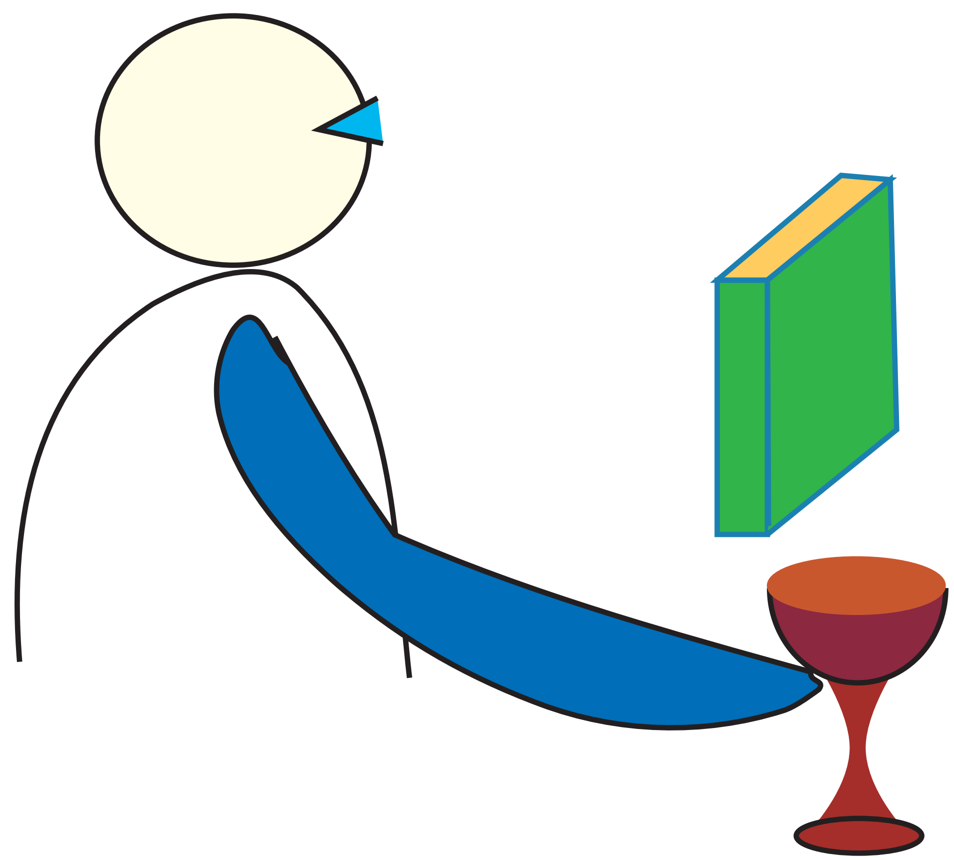


Today

- illustrate how these functions support lifting neural dynamics to higher cognition
- in the context of the perceptual grounding of concepts/language/relations

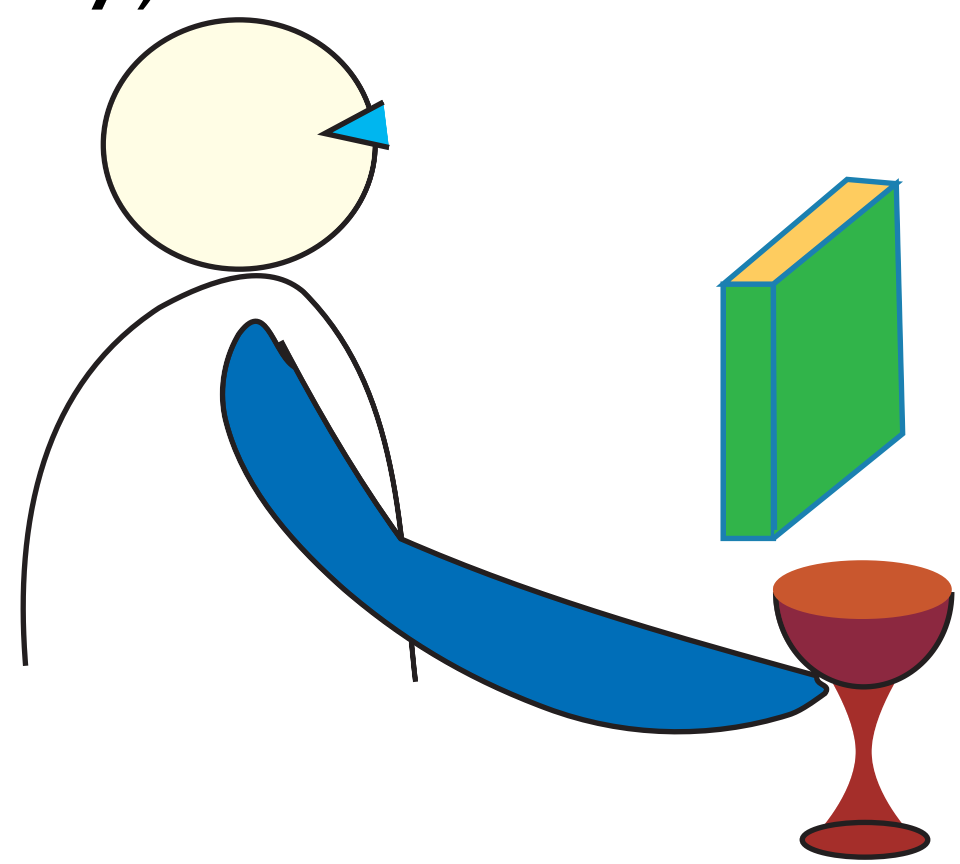
Perceptually grounding language

- human communication in its simplest form is about things that are our there in our environment, perceivable, reachable by action
- e.g., this cup is brown



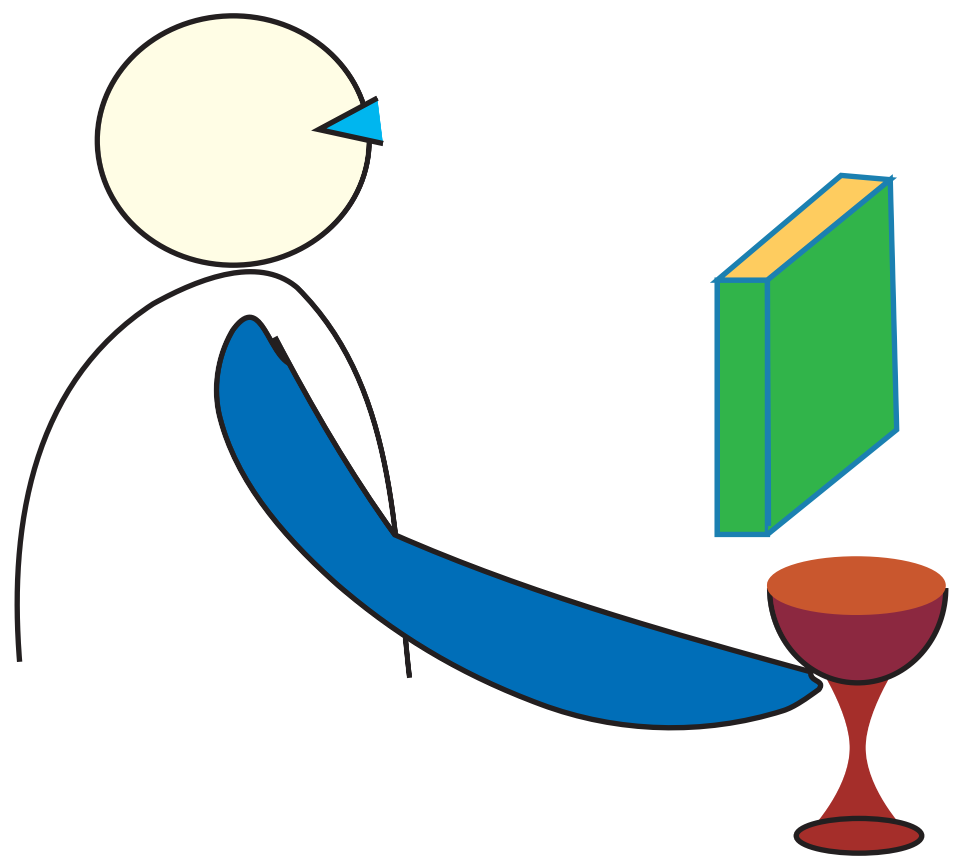
Perceptually grounding language

- this could be based by both the speaker and the listener looking at the scene and grounding the word "cup" by bringing an object of that category into the foreground
- also called "targetting" (Talmy)



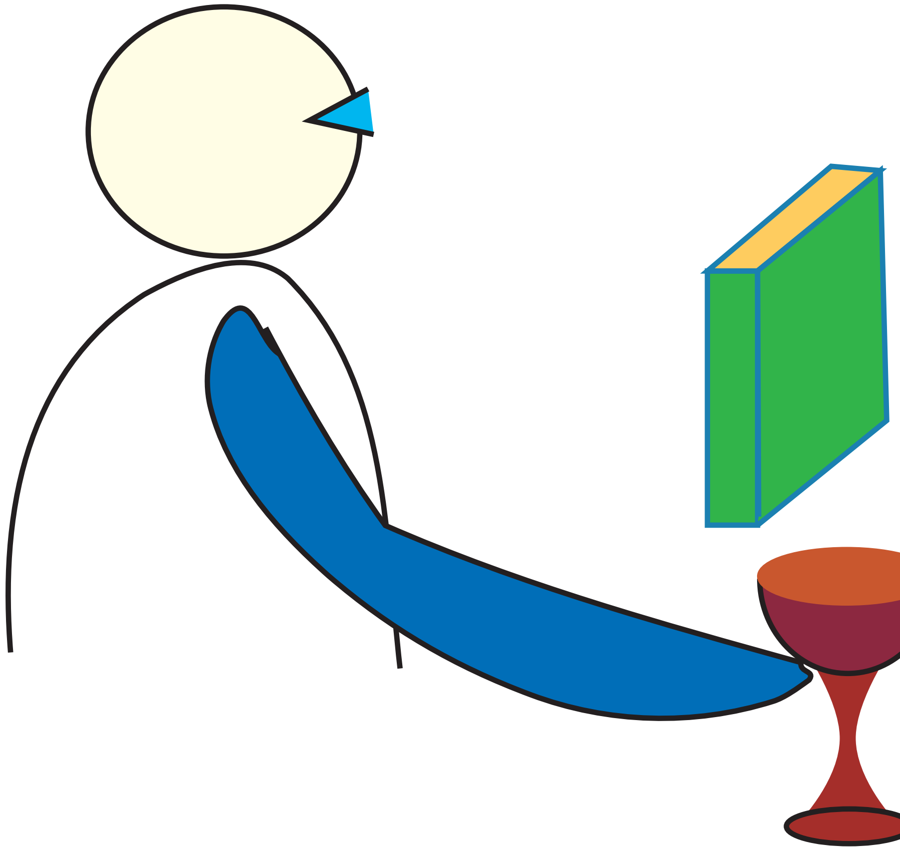
Perceptually grounding language

- that process could be mediated by other forms of communication, e.g., pointing (deictic code)



Perceptually grounding language

- that process could also be mediated by spatial language, e.g., “the cup to the right of the green book is brown” (spatial language)
- (which presupposes that the reference object “green book” is grounded for speaker and observer)



Perceptually grounding language vs. describing

- Perceptual grounding: understanding phrases by finding in the visual array the objects to which the phrase refers
- Describing: producing phrases that describe an observed scene or event

“what is to the right of the green object”

Spatial language

- such utterances as "to the left of", "on top of", "in", "in front of", "toward the south", "in front of" etc.
- a part of language that deep: evolves slowly in languages, with profound differences between languages and cultures, that is particularly challenging for "grounding"

Spatial language

- Examples:
  - some cultures use absolute directions “north”, “south” etc. even on a local scale (e.g, “the car north of the house” rather than “the car in front of the house”).
  - others have special spatial language referring to geographical landmarks (e.g., islanders who have a word for “toward the beach” vs. “away from the beach, toward the inland”)
  - “in front of” is used differently even in different indo-european languages

Grounding spatial language

- involves necessarily reference frames… there are 4 basic and commonly used reference frames

Grounding spatial language

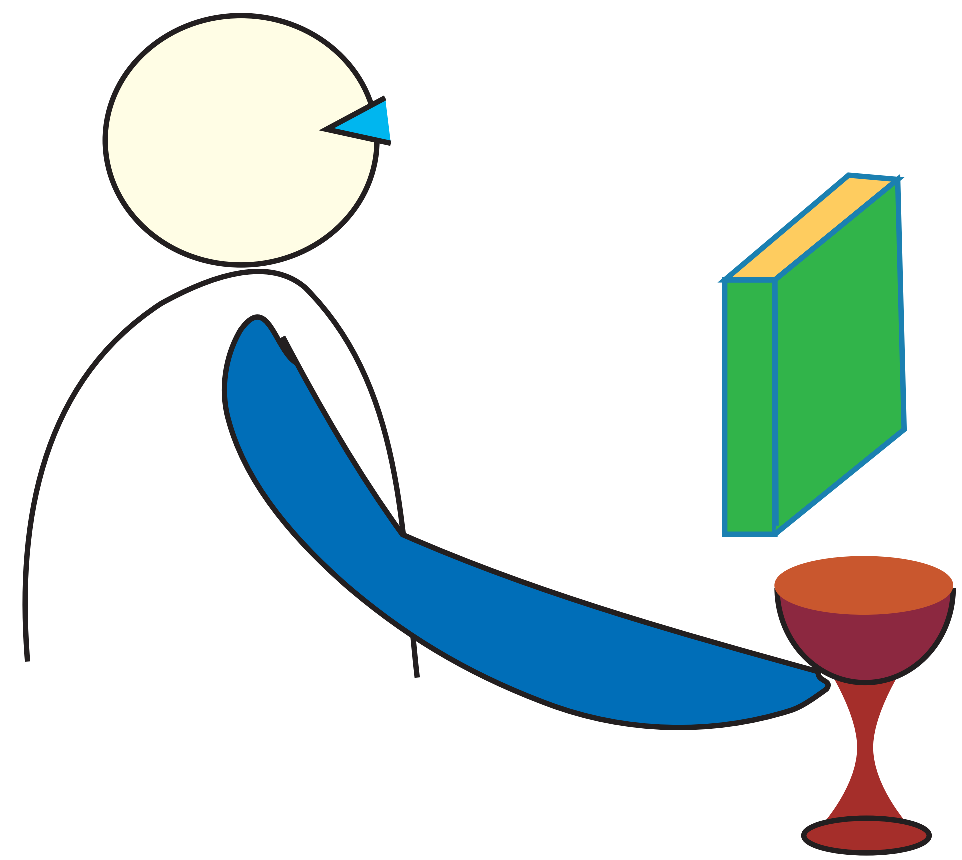
- orientation relative to speaker, position centered in speaker
  - “on my left”
- orientation relative to world/object, position centered in speaker:
  - “north”, “south...” or “leeward”, “windward” ...
- orientation relative to speaker, position centered in object
  - “the cup to the right of the bottle”
- orientation relative to object, position centered in object
  - “leave the train on the right hand side”

Grounding spatial language

- reference frames are subtle
  - Example: “in front of” can be in an ego-centric frame if the object has no special long axis and front end (e.g., “in front of the tree” meaning “between me and the tree”)
  - but can be in an object centered frame if the object has a long axis and front end (e.g. “in front of the car” meaning “on the side of the car in the direction in which its front end points”)
    - (and on this count different languages differ)

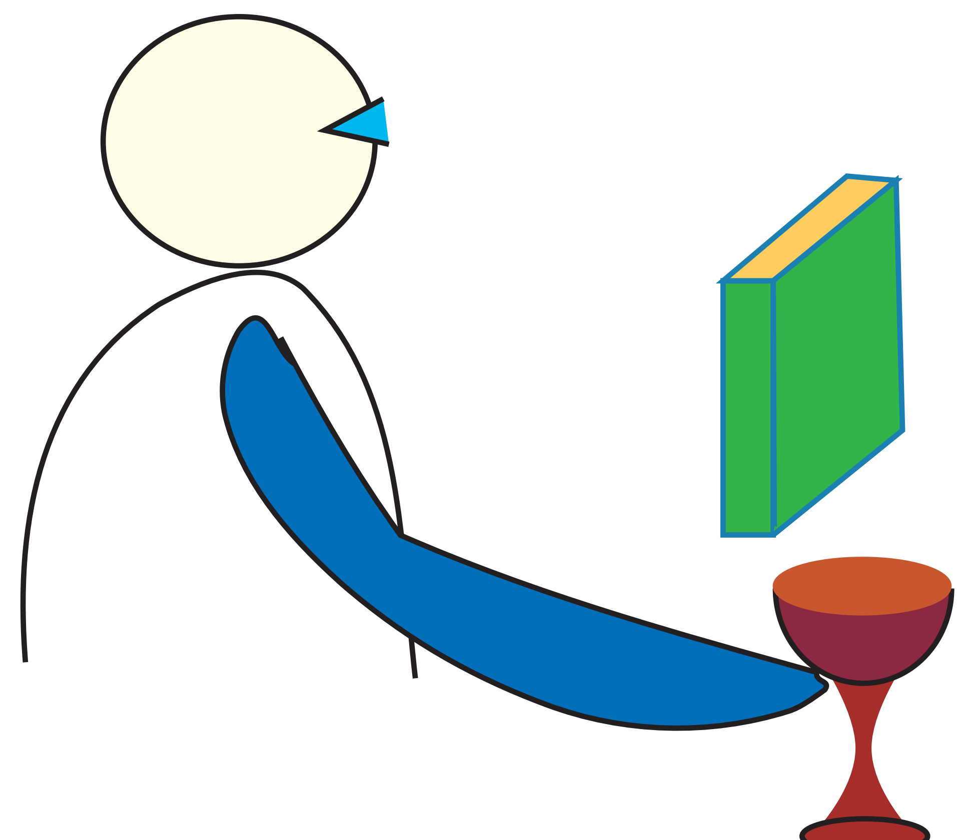
Grounding spatial language

- spatial language often involves reference objects
  - Example: “to the right of the green book”: this is a statement in an ego-centric reference frame for direction but that is spatially centered in an object



Grounding spatial language

- spatial language often involves coordinate transforms
  - e.g., "to the right of the green book": coordinate transformation: from the speaker/observer centered reference frame into a frame centered in the reference object
  - e.g., "to my right" requires the listener to transform the reference frame from his or her own view to the directional and positional frame of the speaker



Operations involved in grounding spatial language

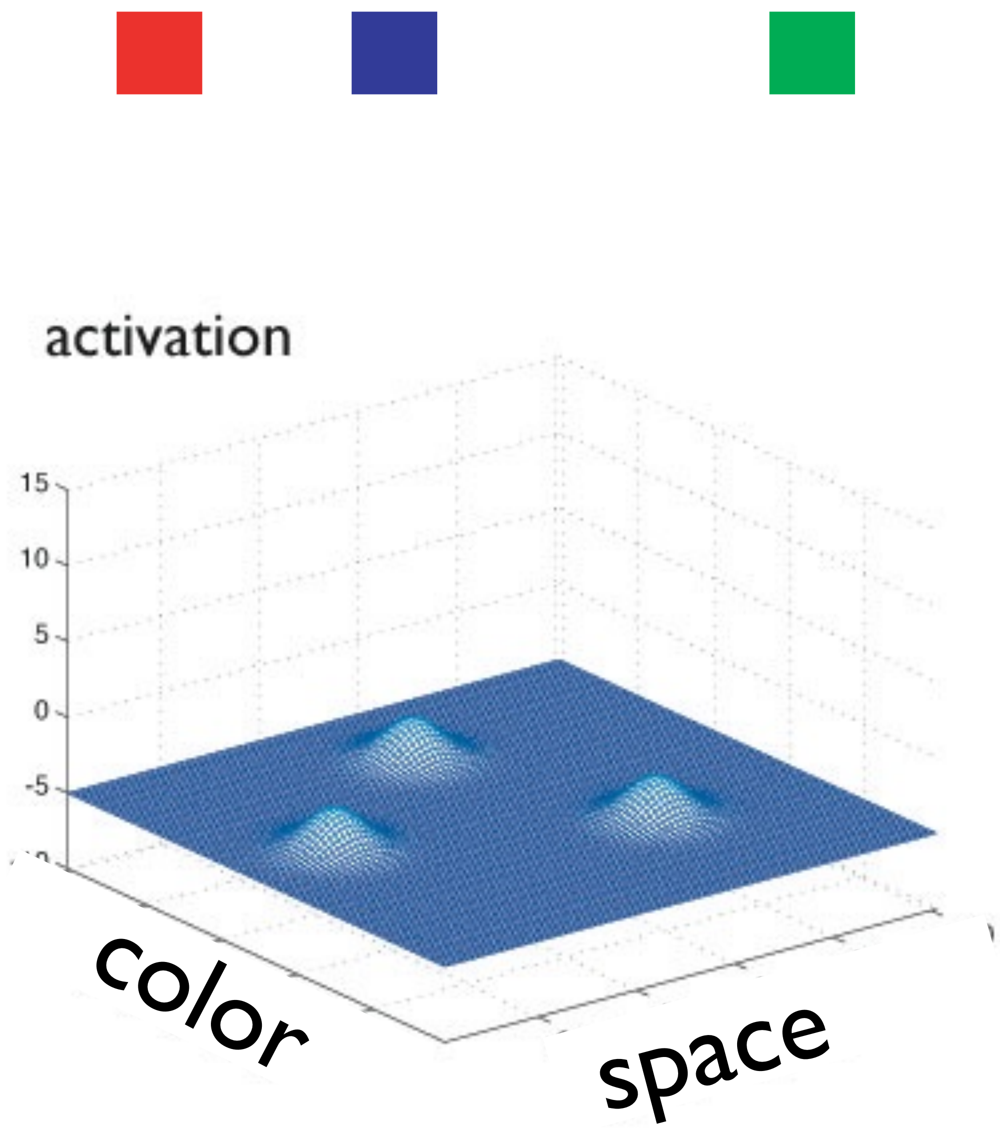
- bring objects (target and reference) into the perceptual foreground (visually find them)
- make coordinate transformation
- apply comparison operators

DFT approach to bringing a perceptual object into the foreground

- => lecture on higher-dimensional fields

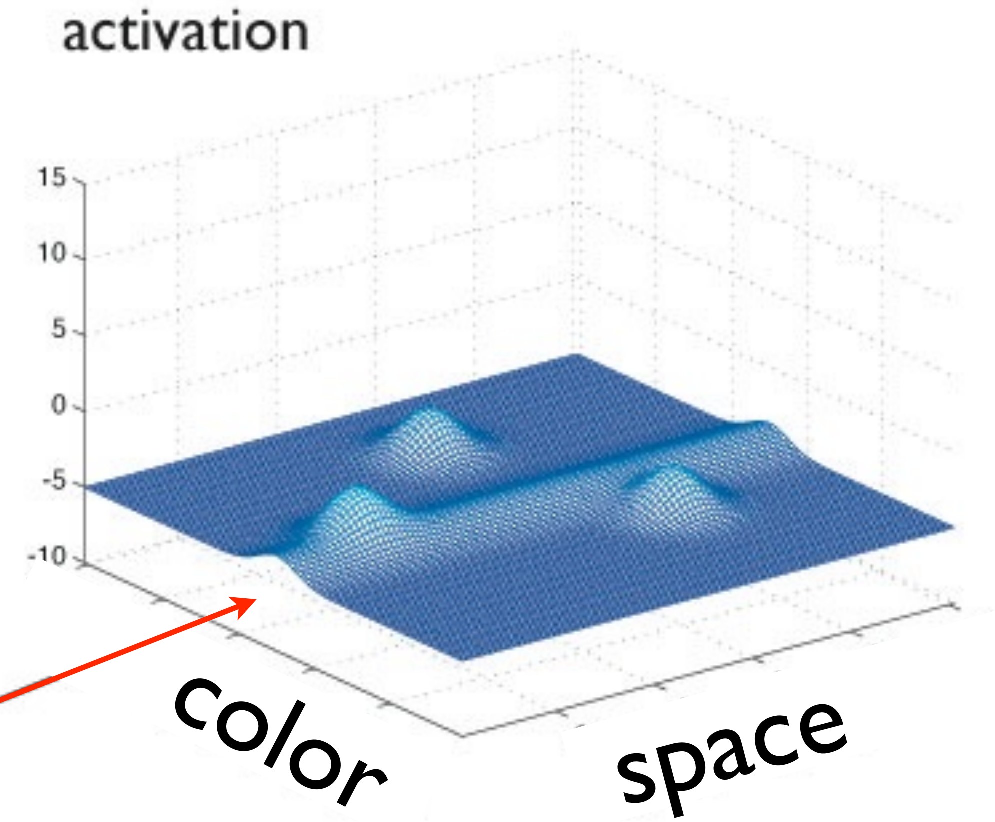
Bringing an object to the foreground

- visual search: "where is the red object"?



Bringing an object to the foreground

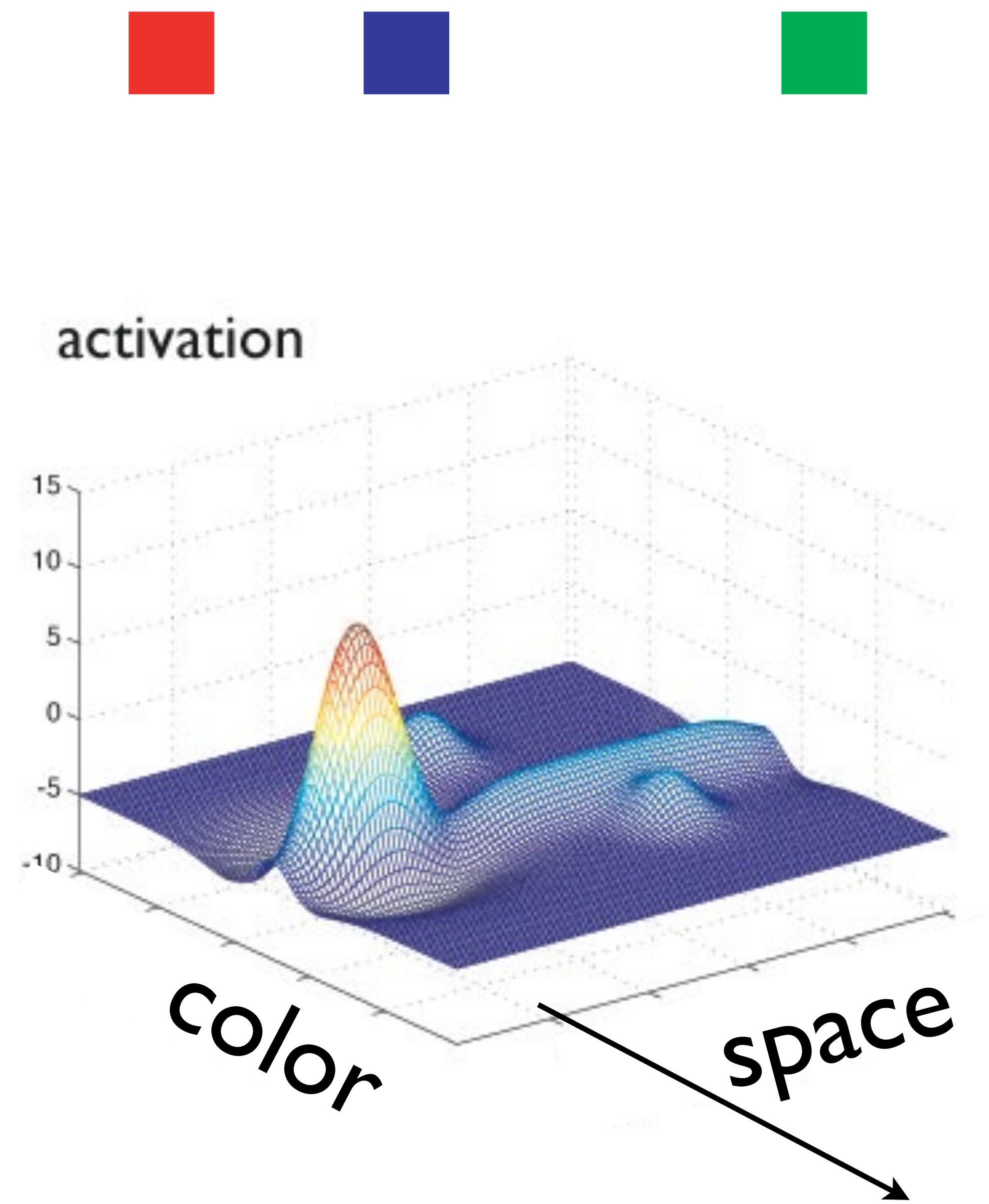
- visual search: "where is the red object"?



ridge specifying red

Bringing an object to the foreground

- visual search: "where is the red object"?

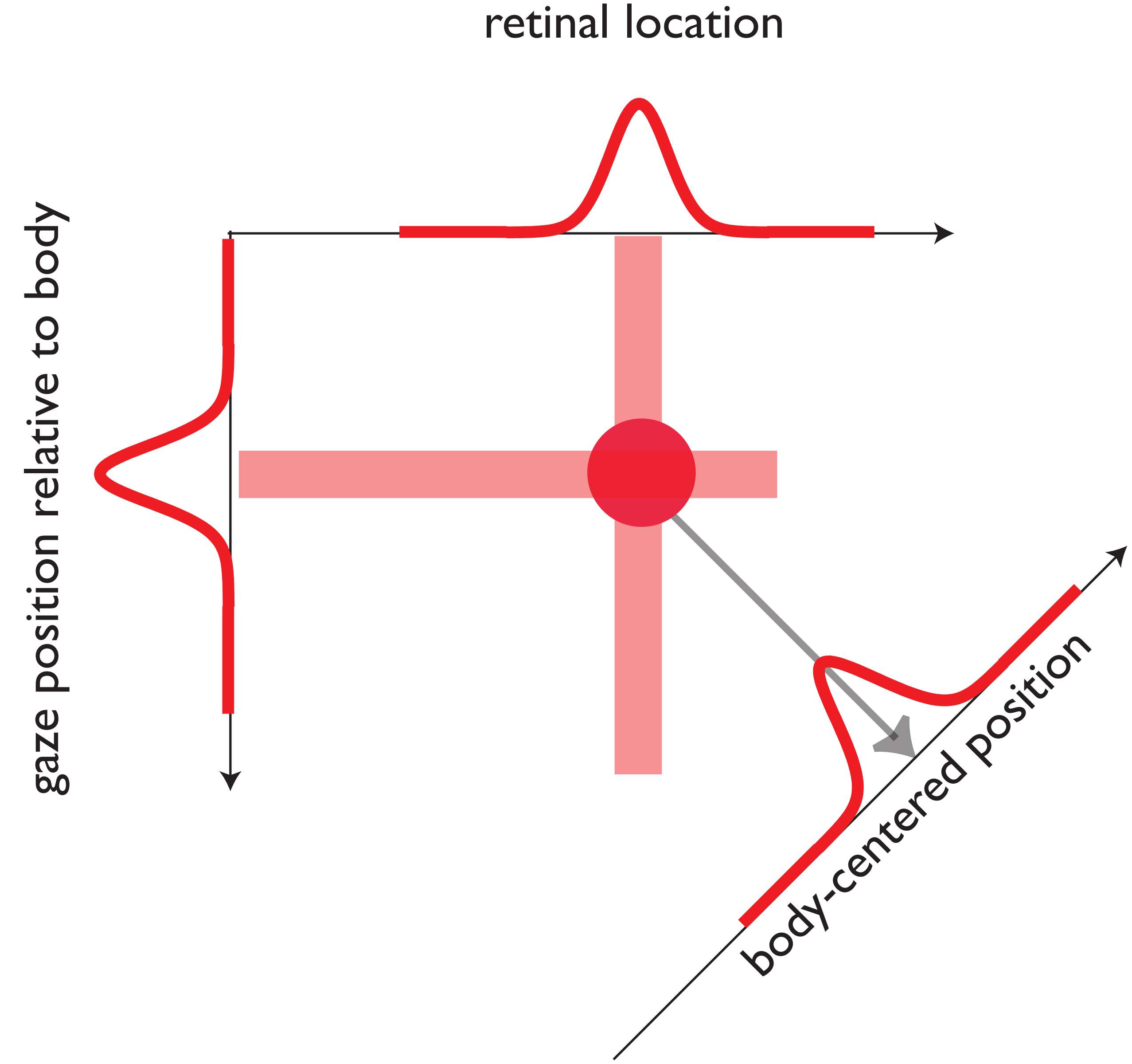


read out spatial location
of red object

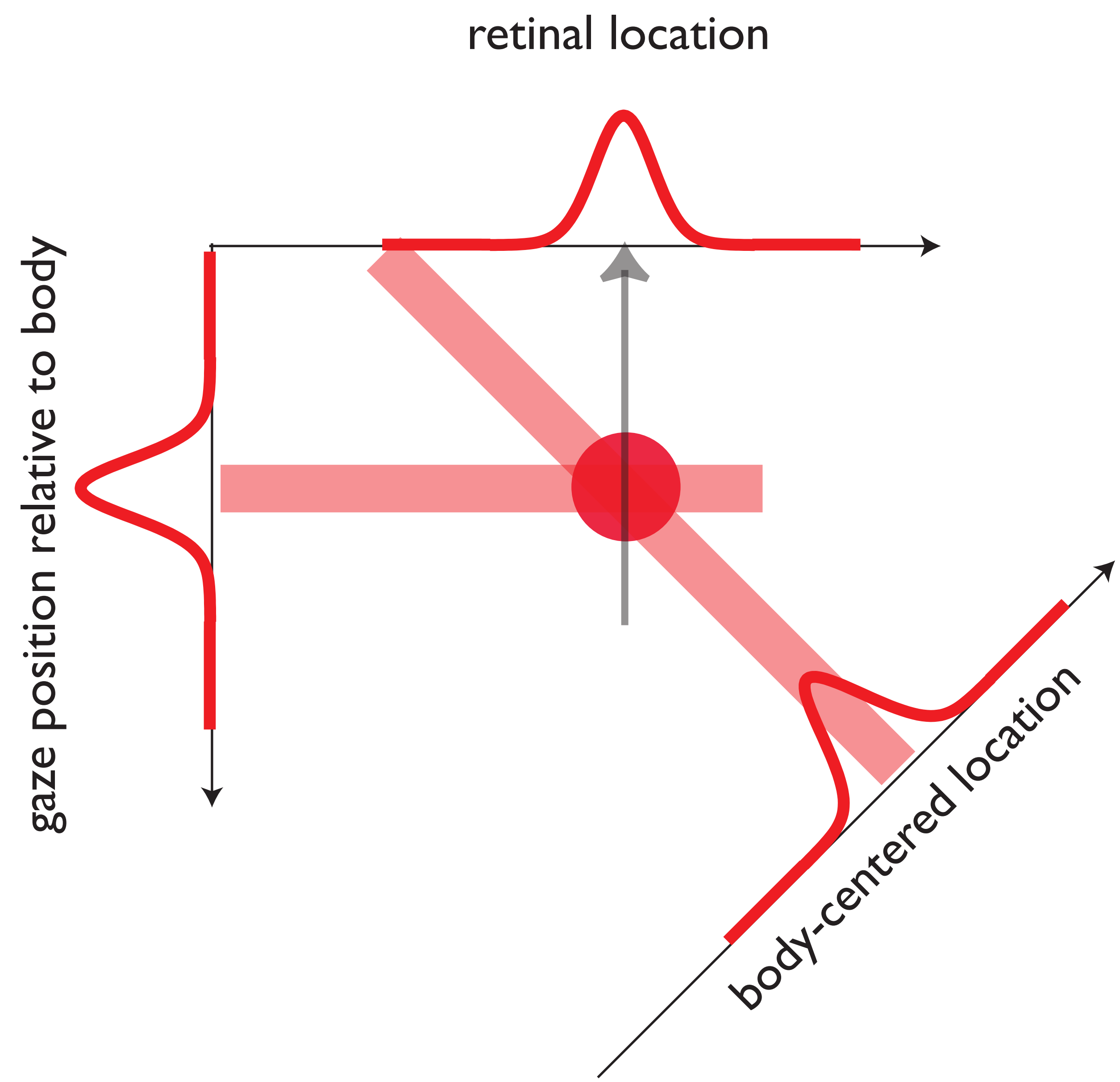
DFT approach to coordinate transforms

- => lecture on higher-dimensional fields

Coordinate transformations

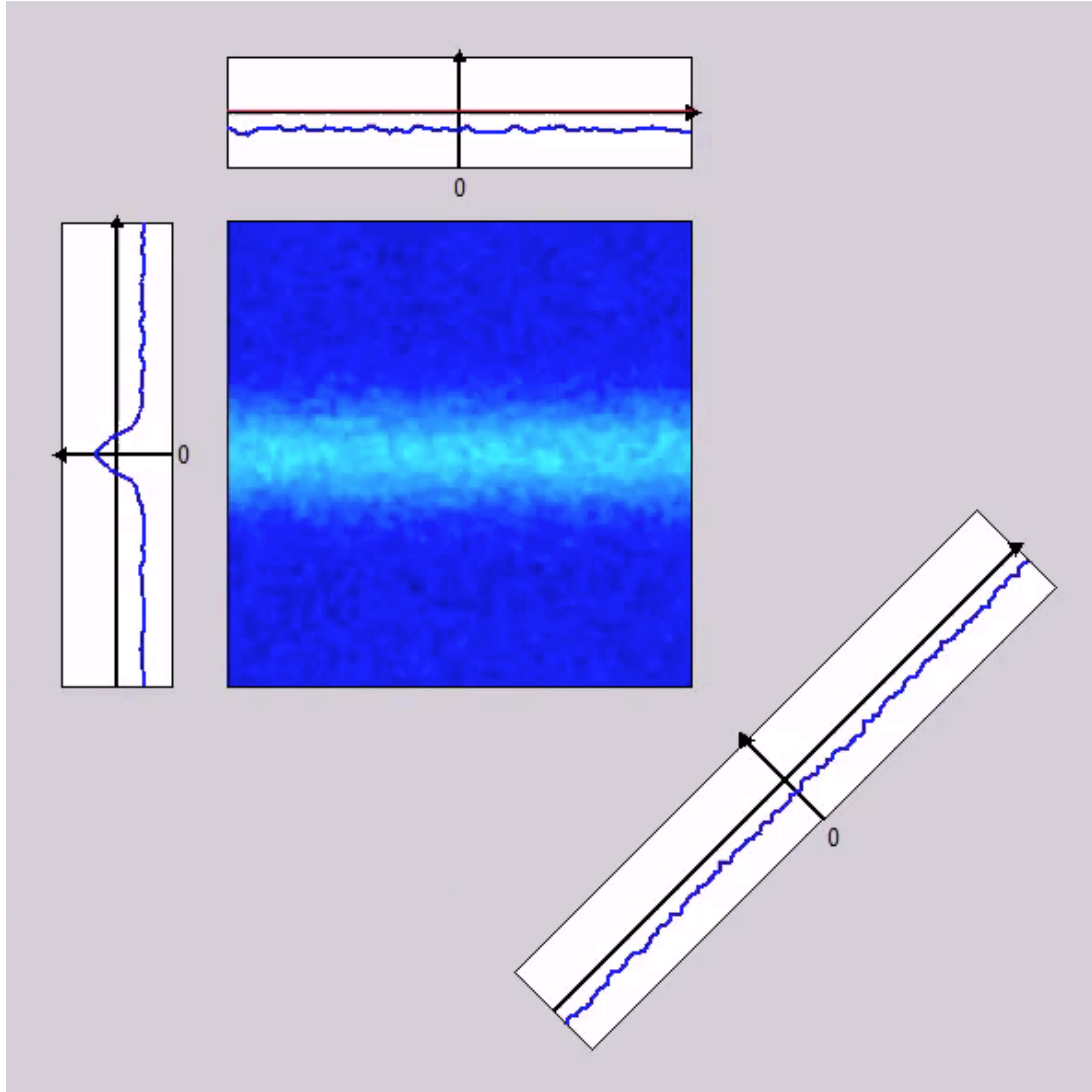


Coordinate transformations



Coordinate transformations

- predict retinal location following gaze shift

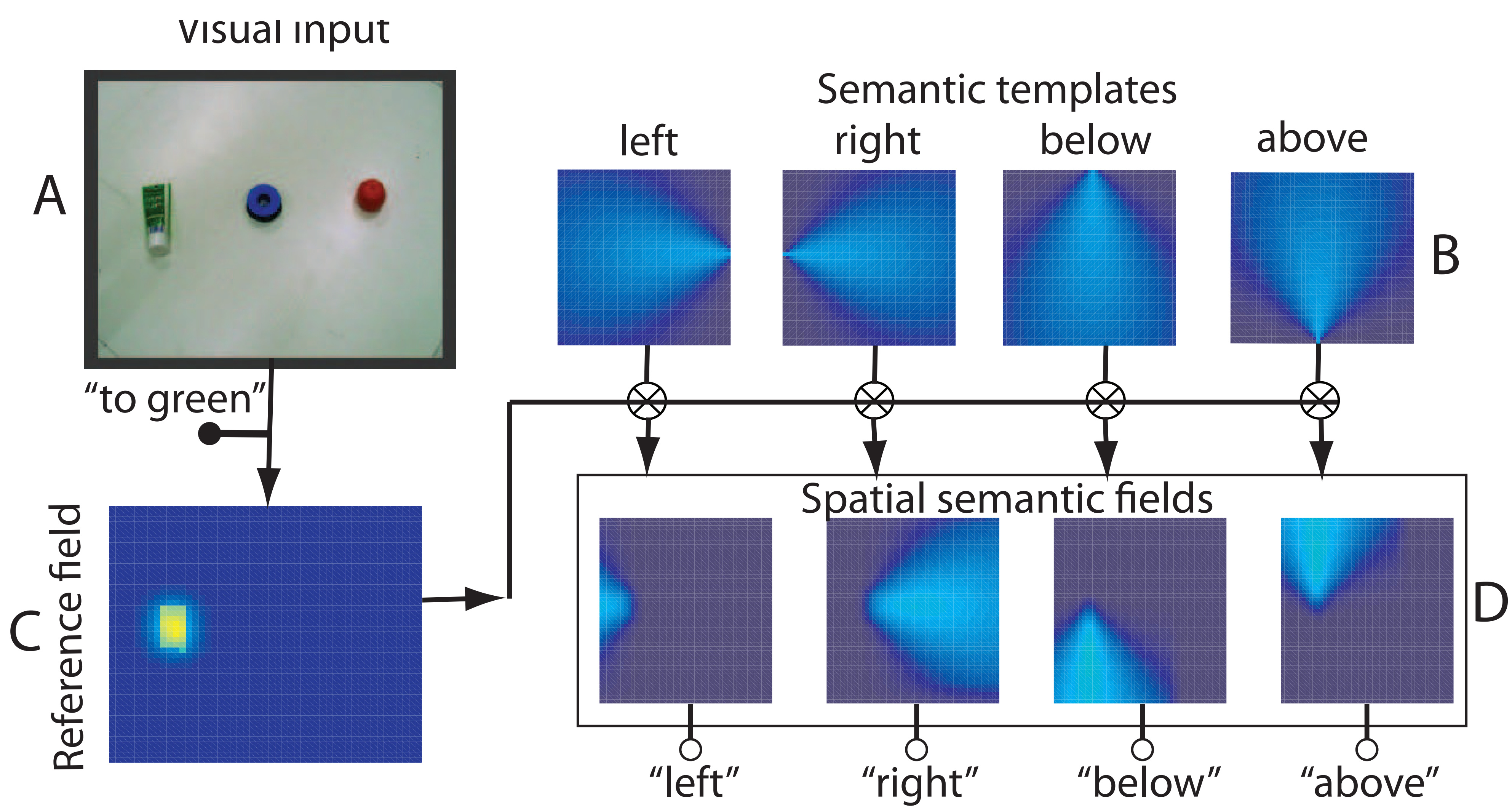


[Schneegans, Schöner, 2012]

DFT approach to applying operators

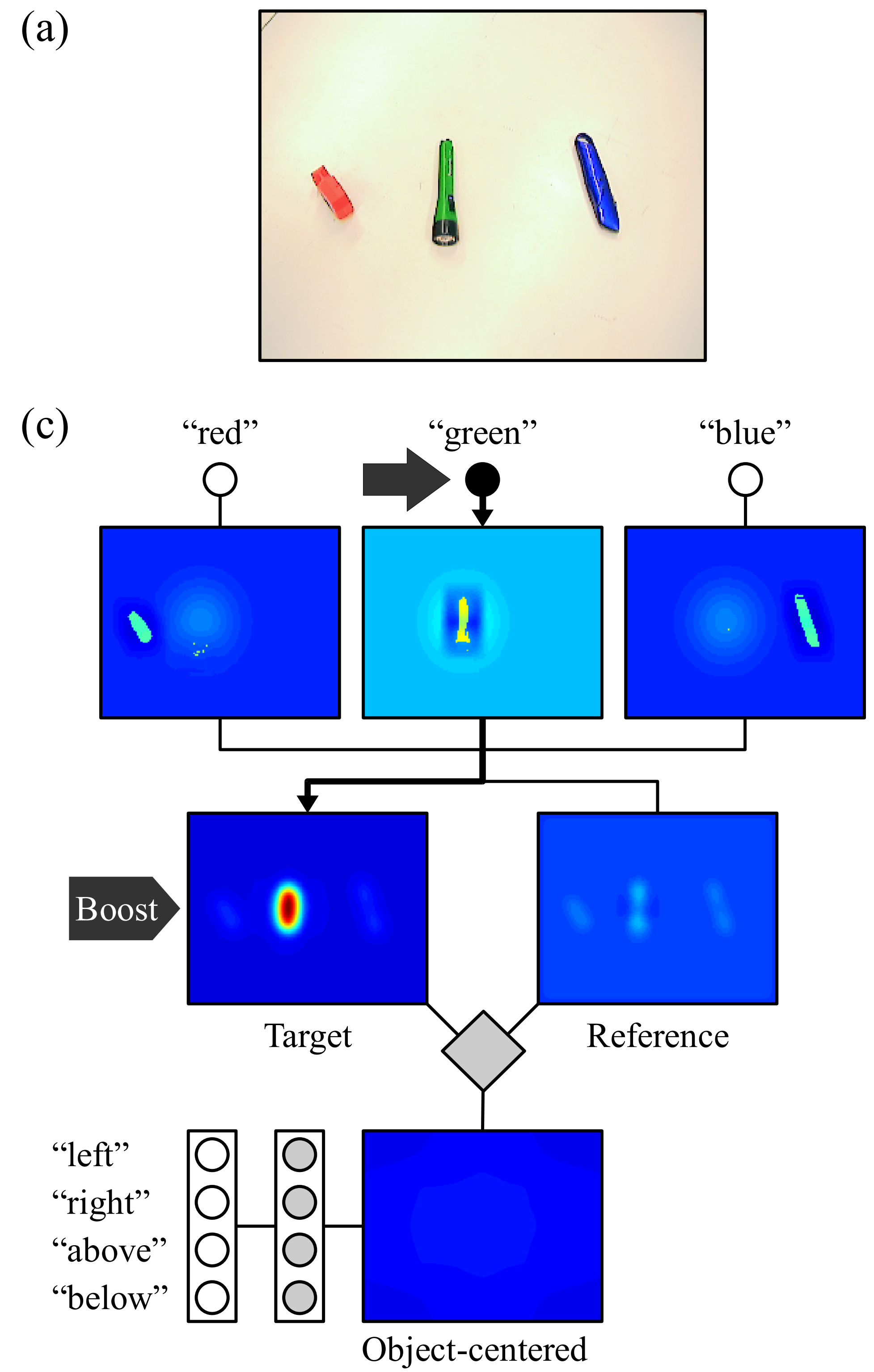
DFT approach to applying operators

- based on convolution of fields with kernels



[from: Lipinski, Sandamirskaya, Schöner, 2009]

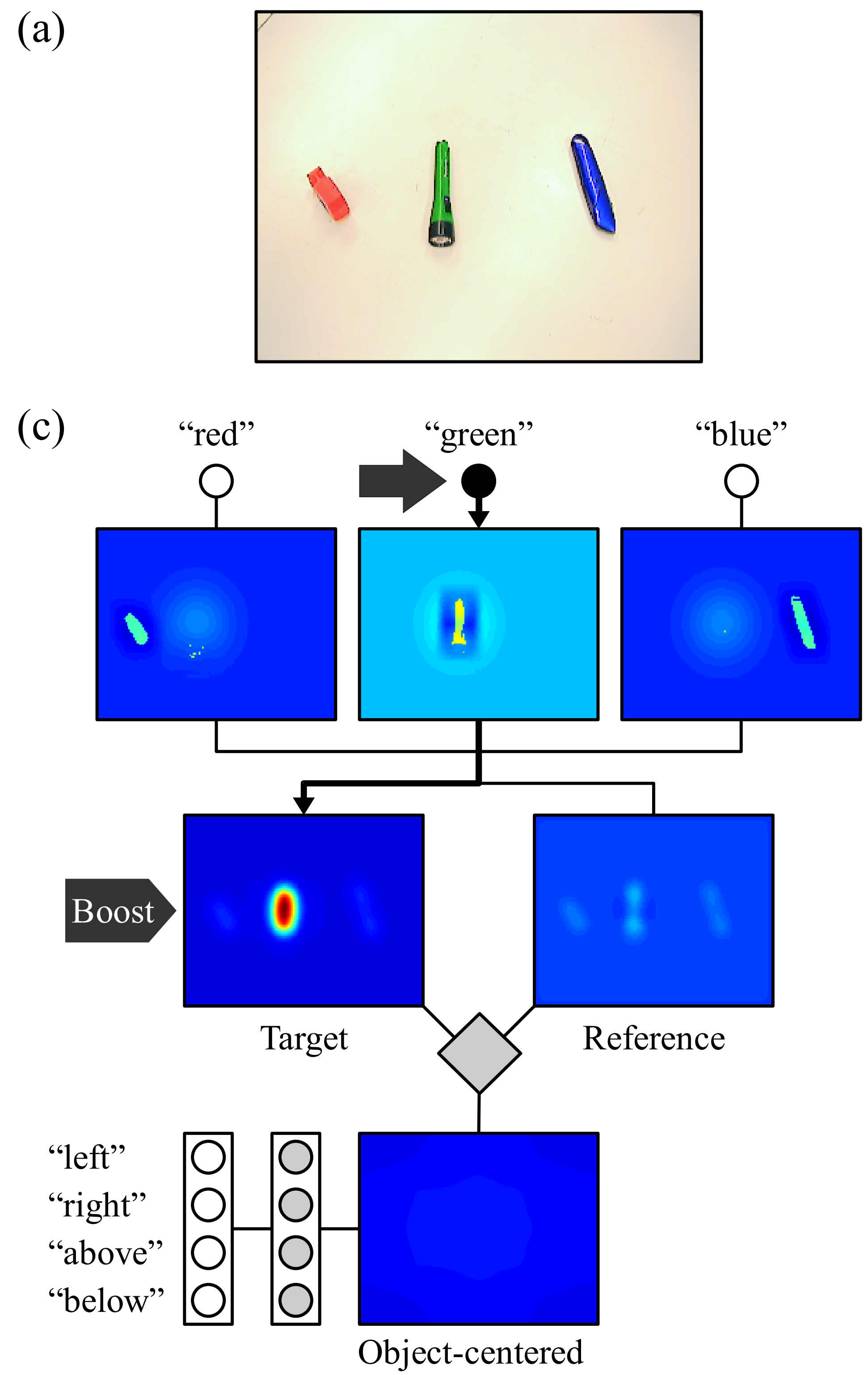
A cognitive architecture for grounded spatial language in DFT



[Lipinski et al: JEP:LMC (2011)]

Spatial comparison in DFT

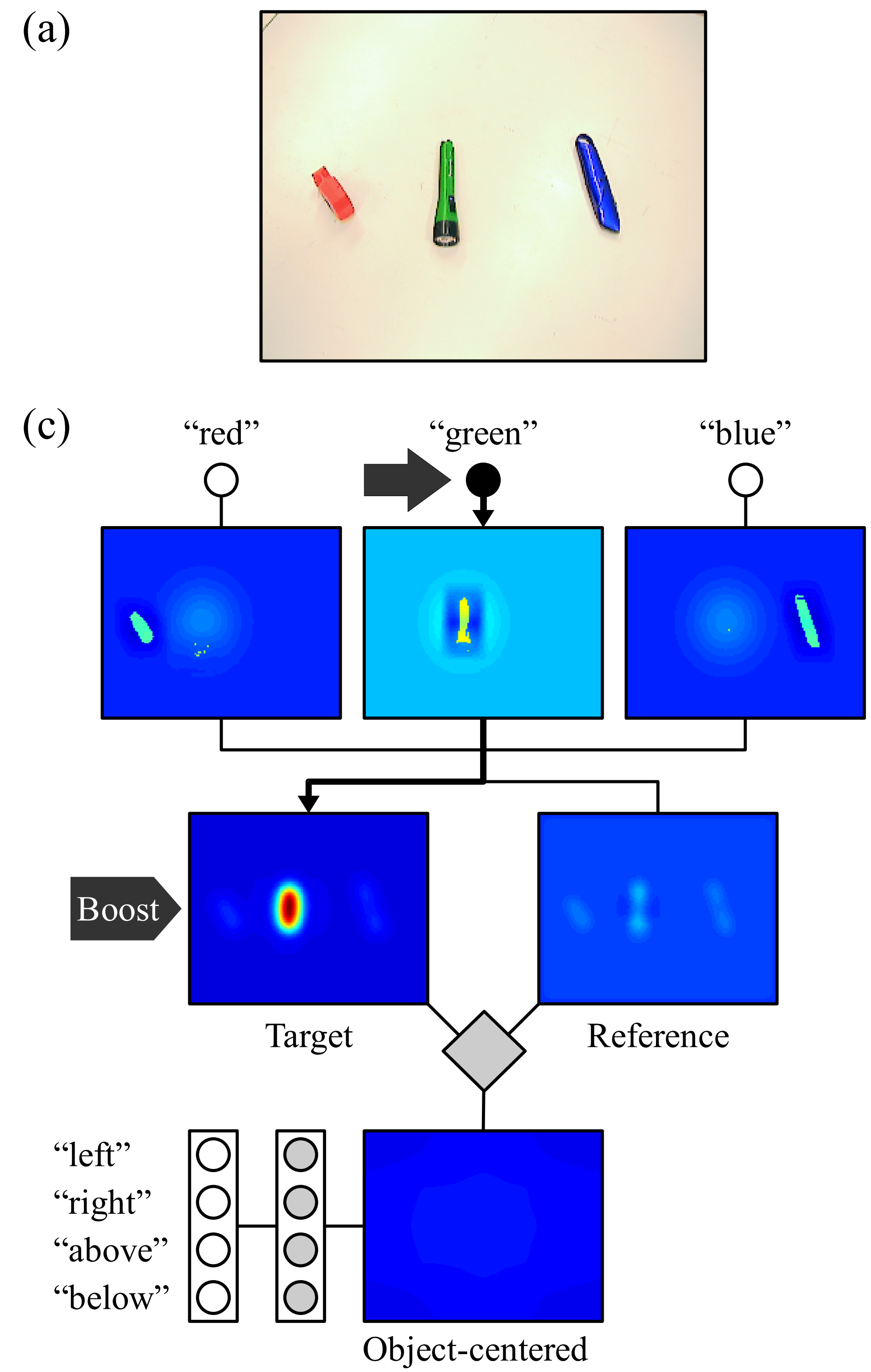
- bring objects into foreground
- make coordinate transformation
- apply comparison operators



[Lipinski et al: JEP:LMC (2011)]

Spatial comparison in DFT

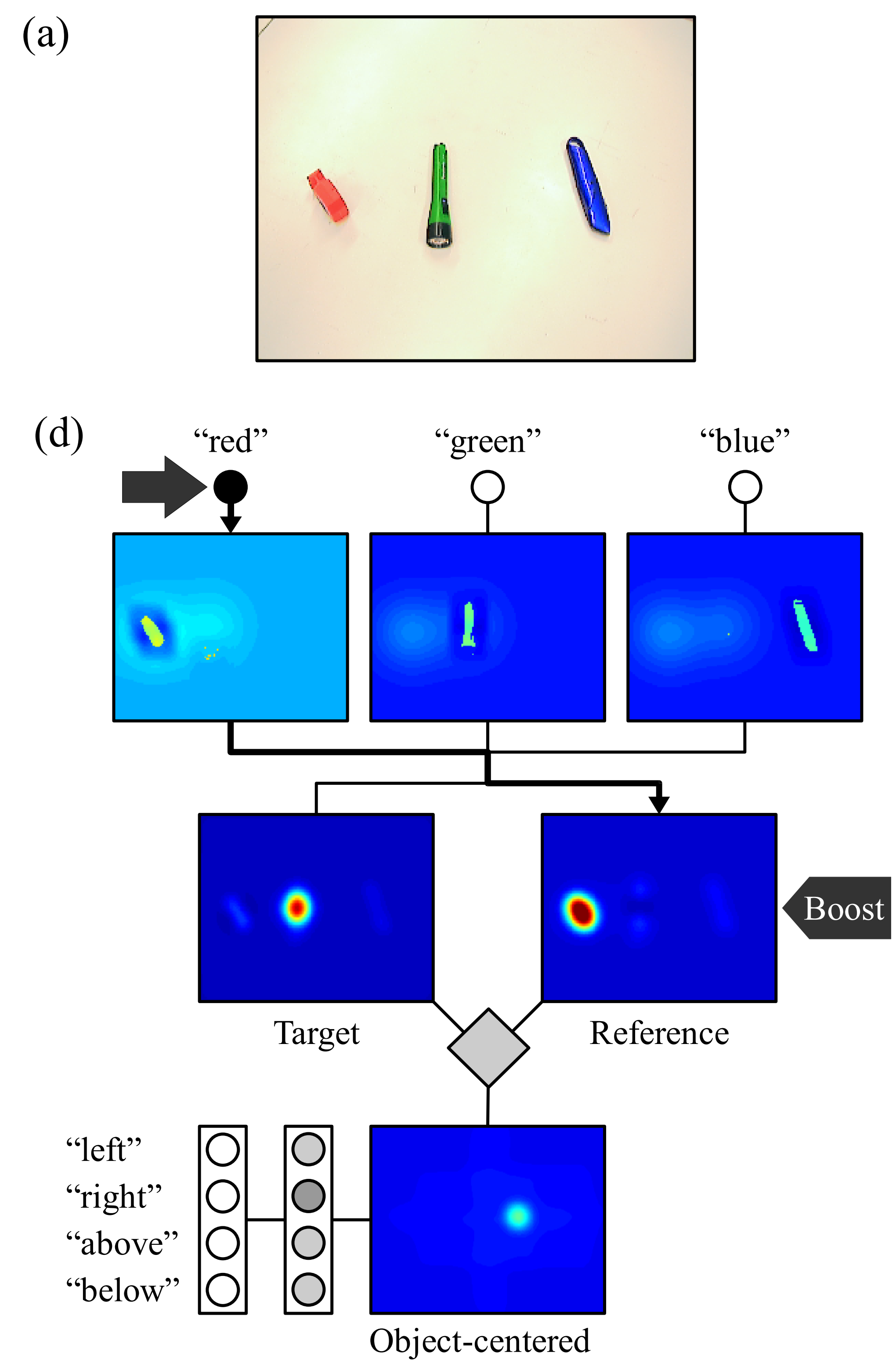
- bring objects into foreground
- make coordinate transformation
- apply comparison operators



[Lipinski et al: JEP:LMC (2011)]

Spatial comparison in DFT

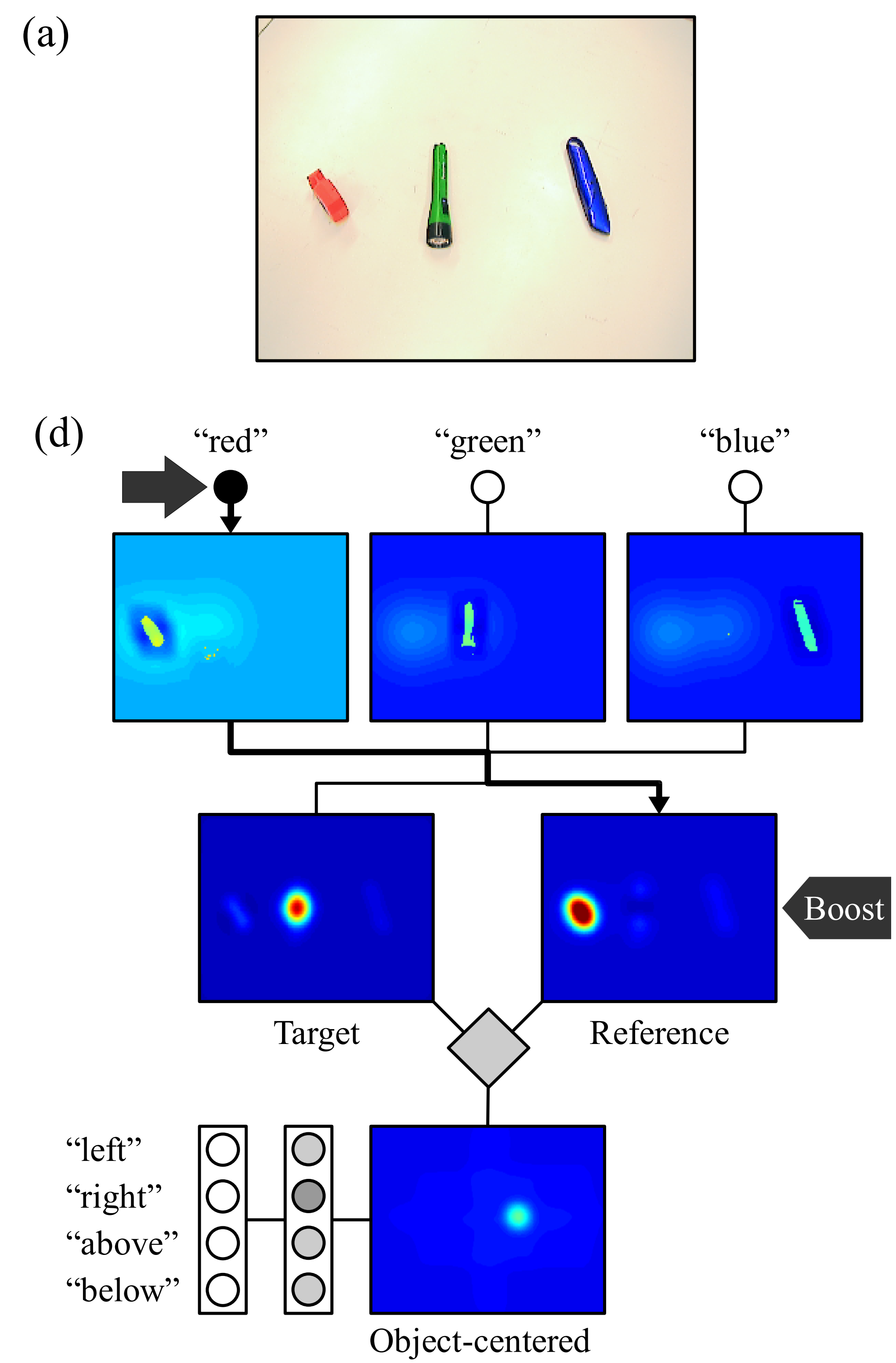
- bring objects into foreground
- make coordinate transformation
- apply comparison operators



[Lipinski et al: JEP:LMC (2011)]

Spatial comparison in DFT

- bring objects into foreground
- make coordinate transformation
- apply comparison operators

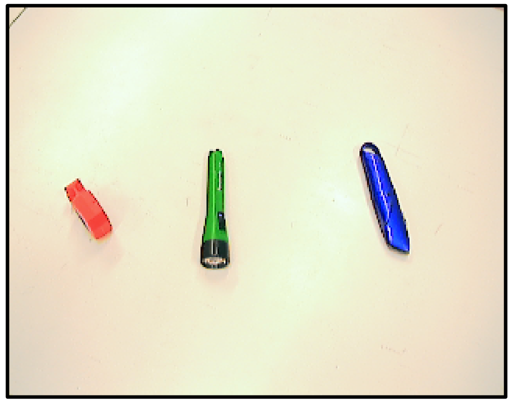


[Lipinski et al: JEP:LMC (2011)]

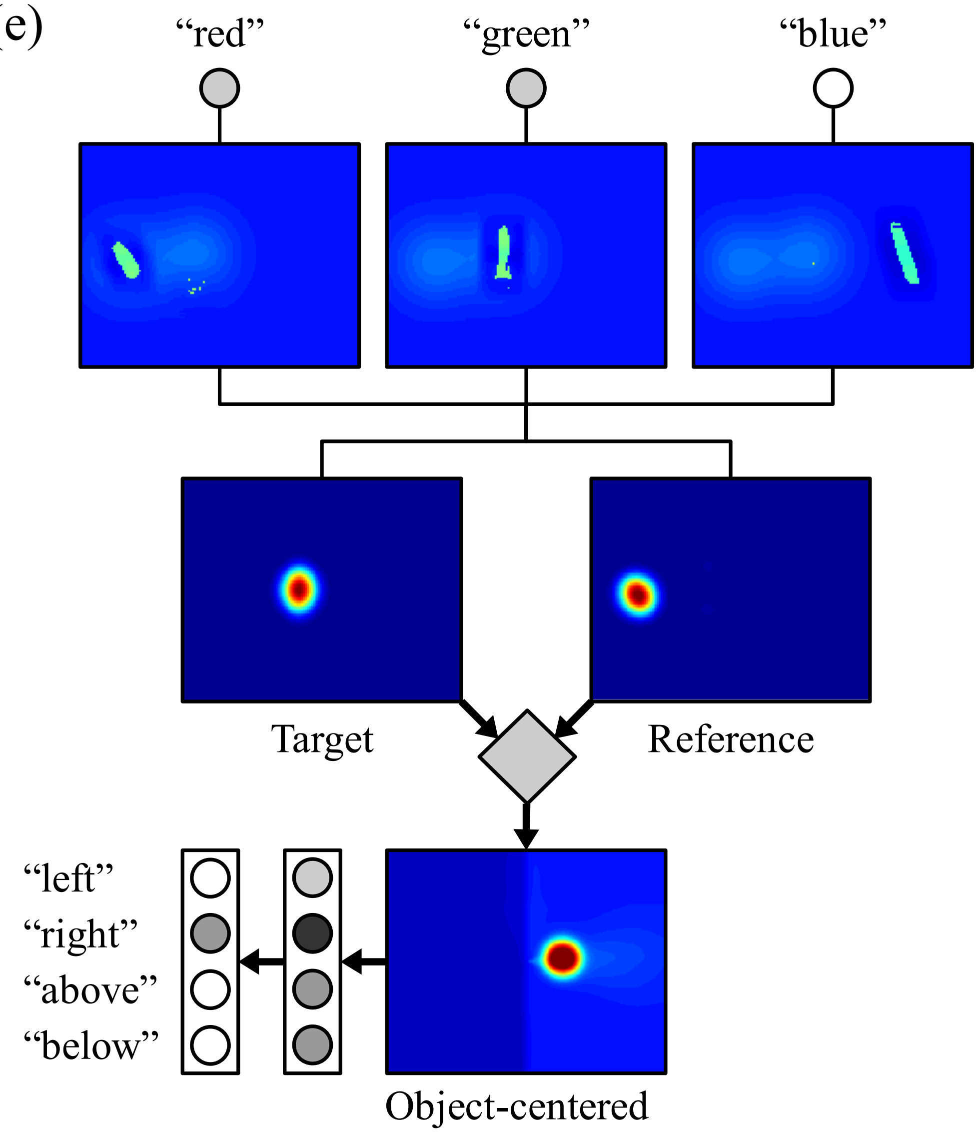
Spatial comparison in DFT

- bring objects into foreground
- make coordinate transformation
- apply comparison operators

(a)

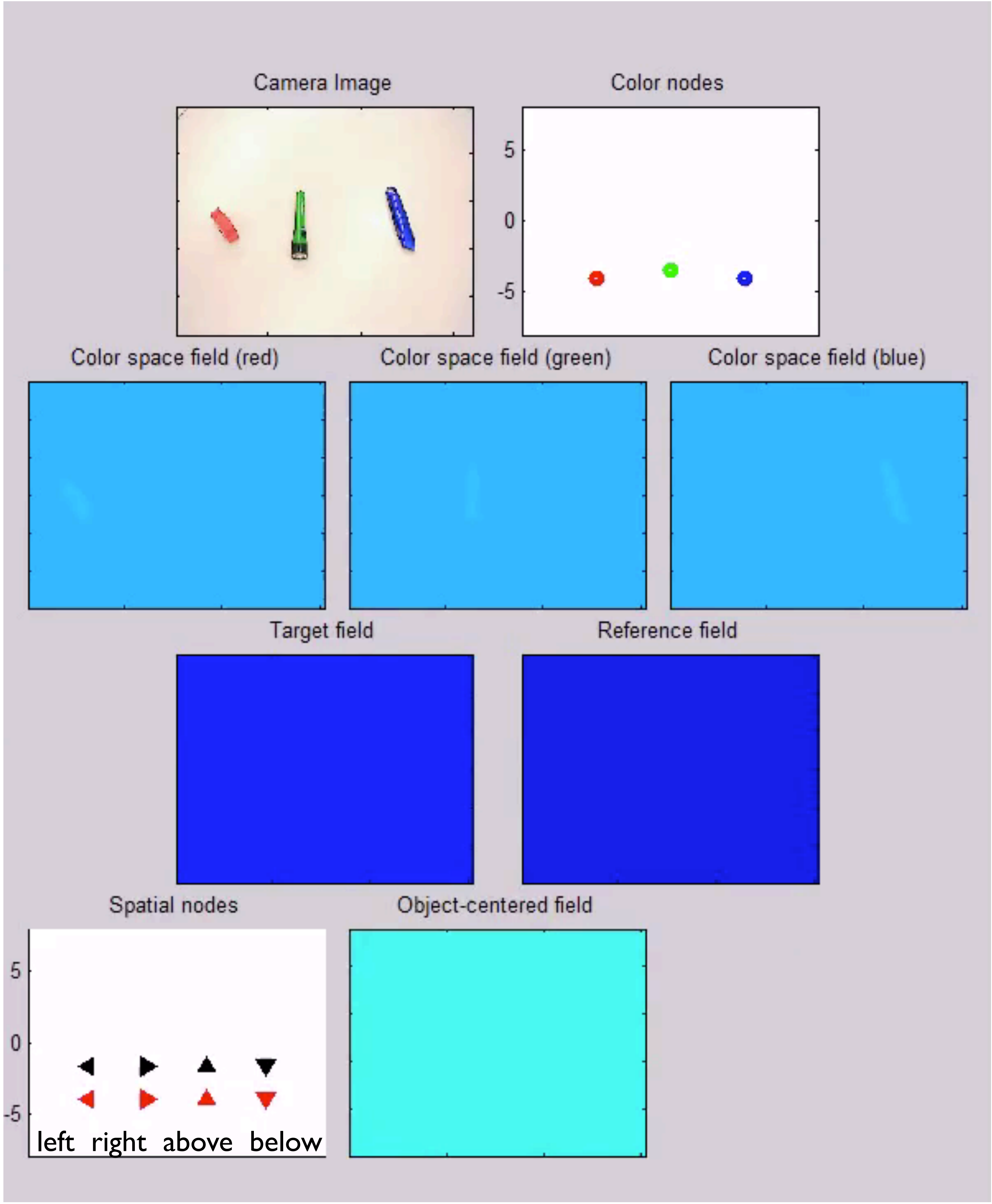


(e)



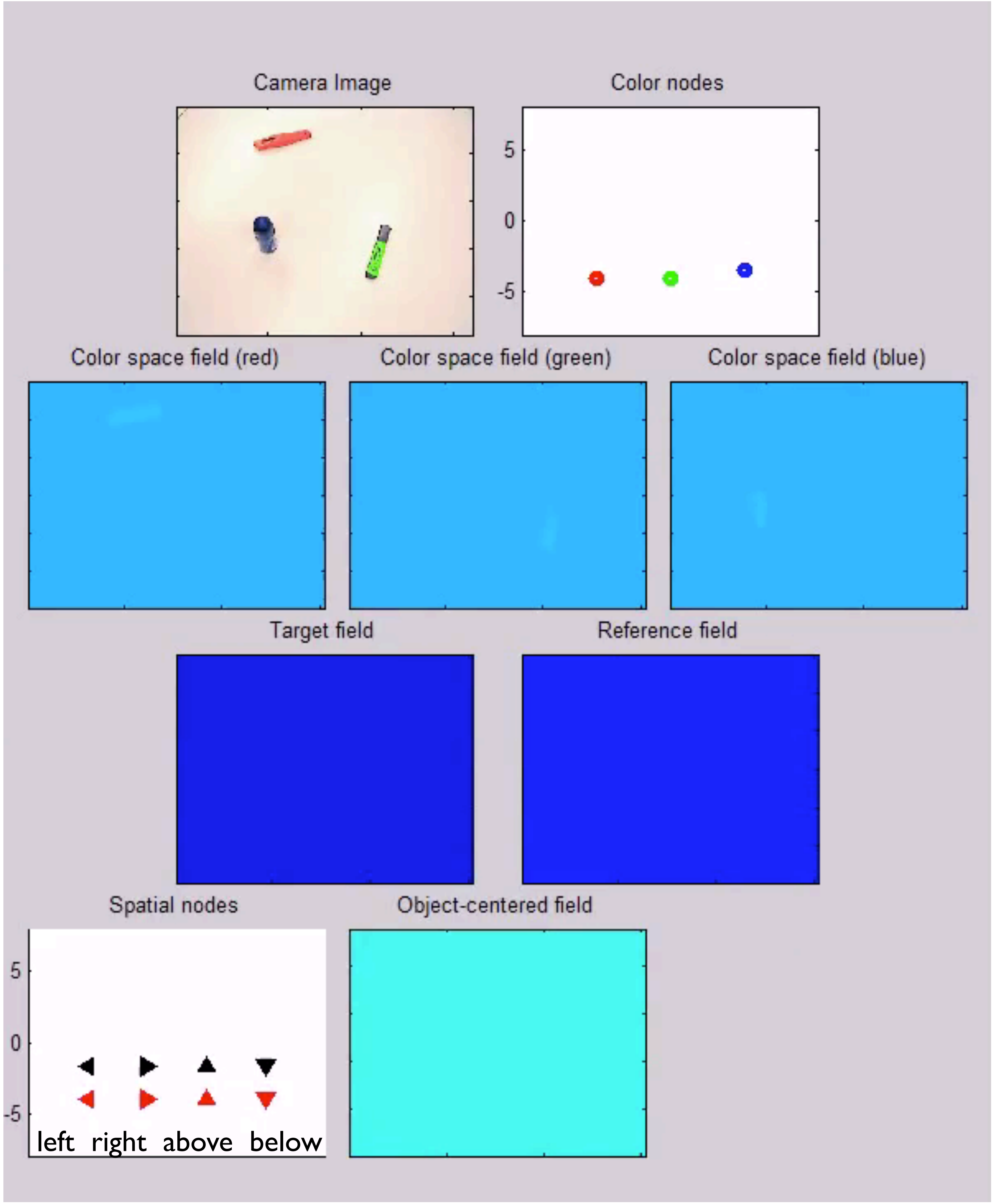
[Lipinski et al: JEP:LMC (2011)]

- “where is the green object relative to the red object?”



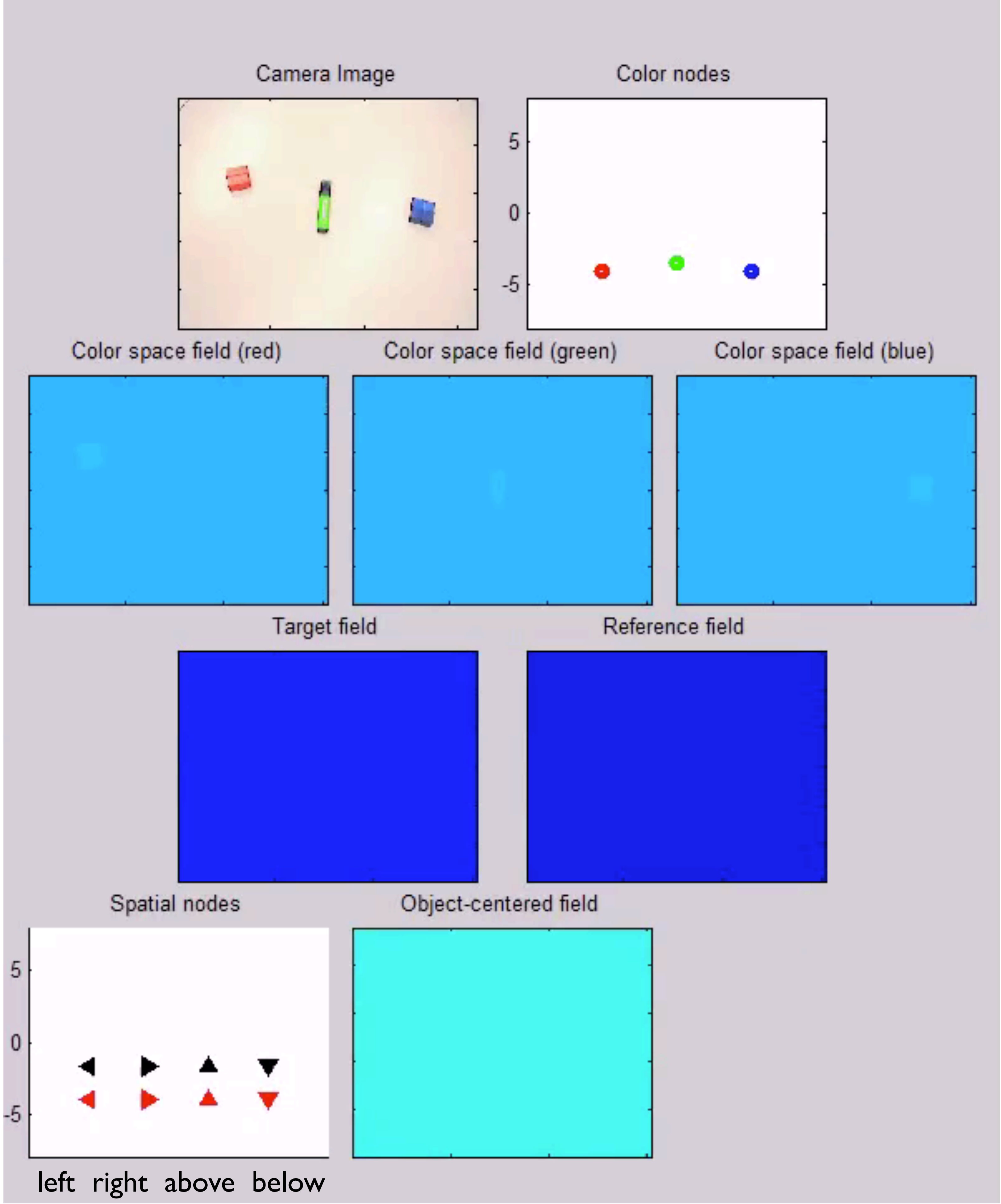
[Lipinski et al: JEP:LMC (2011)]

- “which object is above the blue object?”



[Lipinski et al: JEP:LMC (2011)]

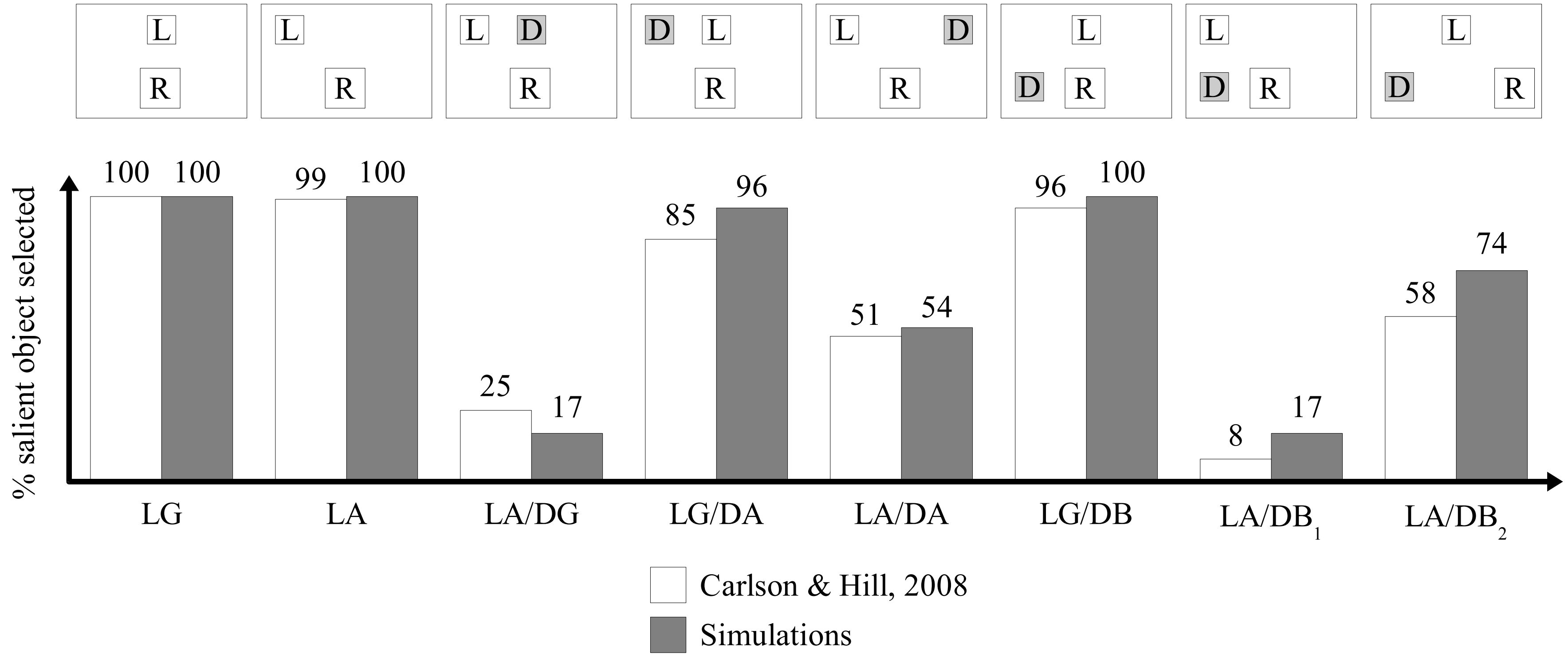
- “where is the green object?”



[Lipinski et al: JEP:LMC (2011)]

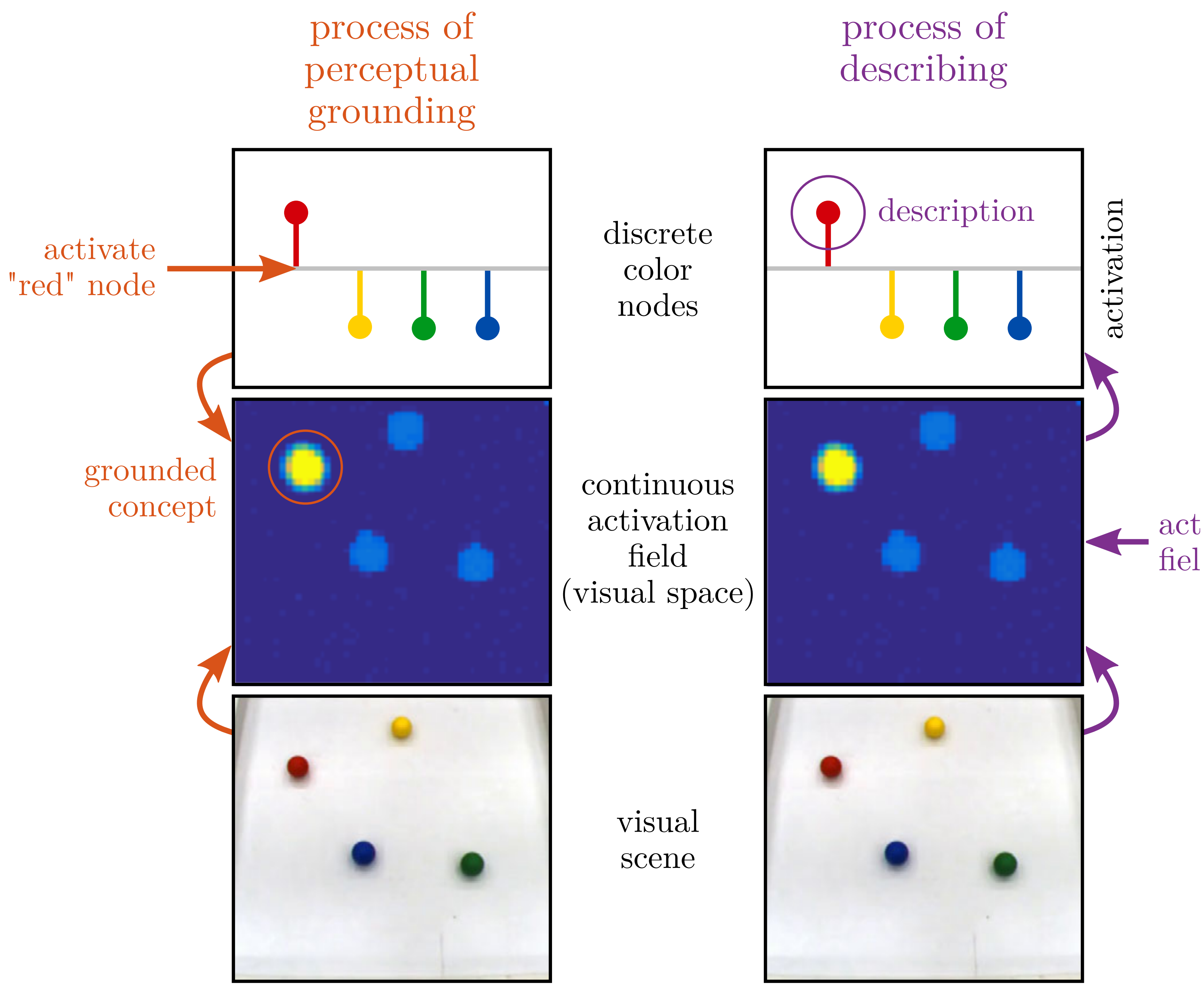
Spatial comparison in DFT

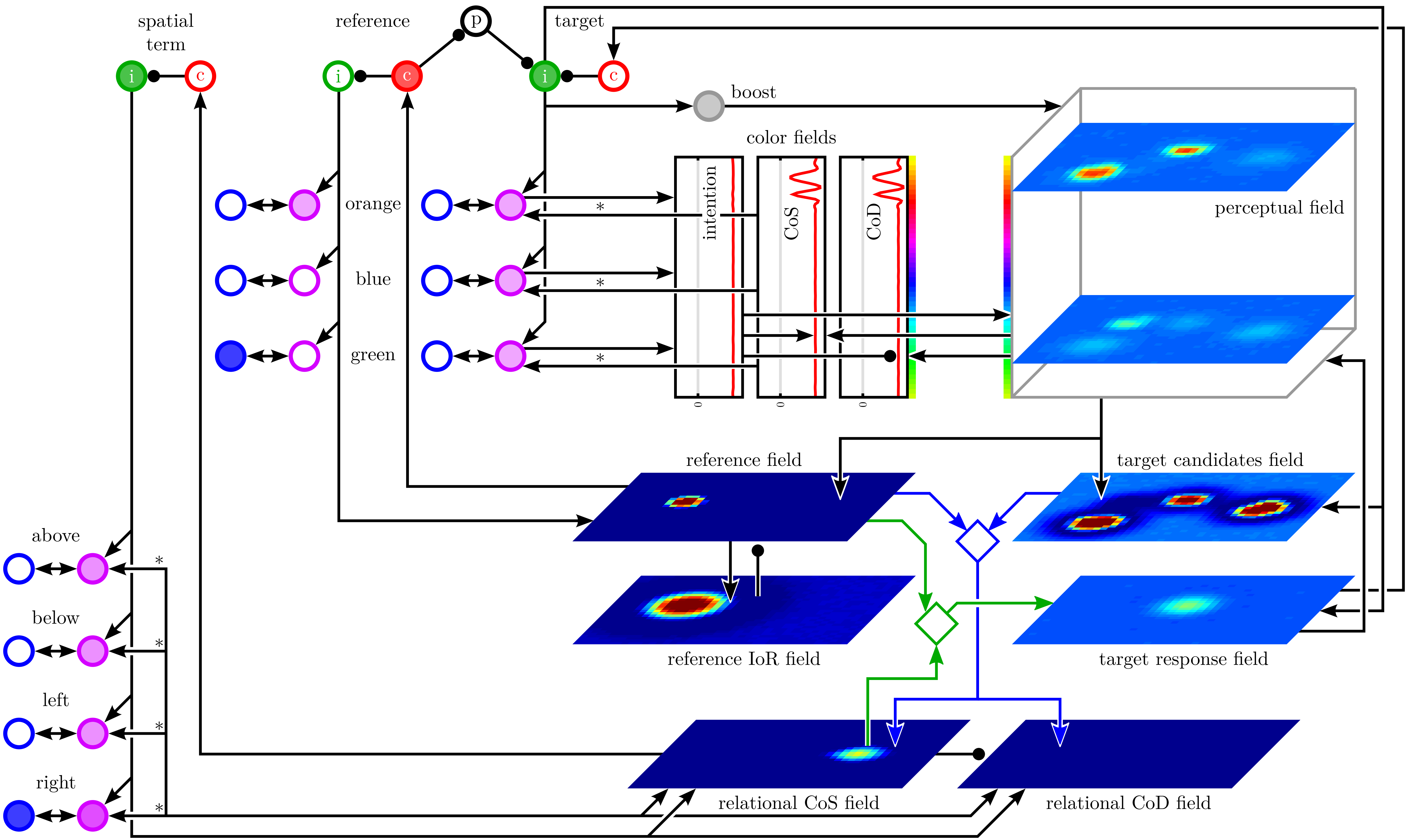
- accounts for human data



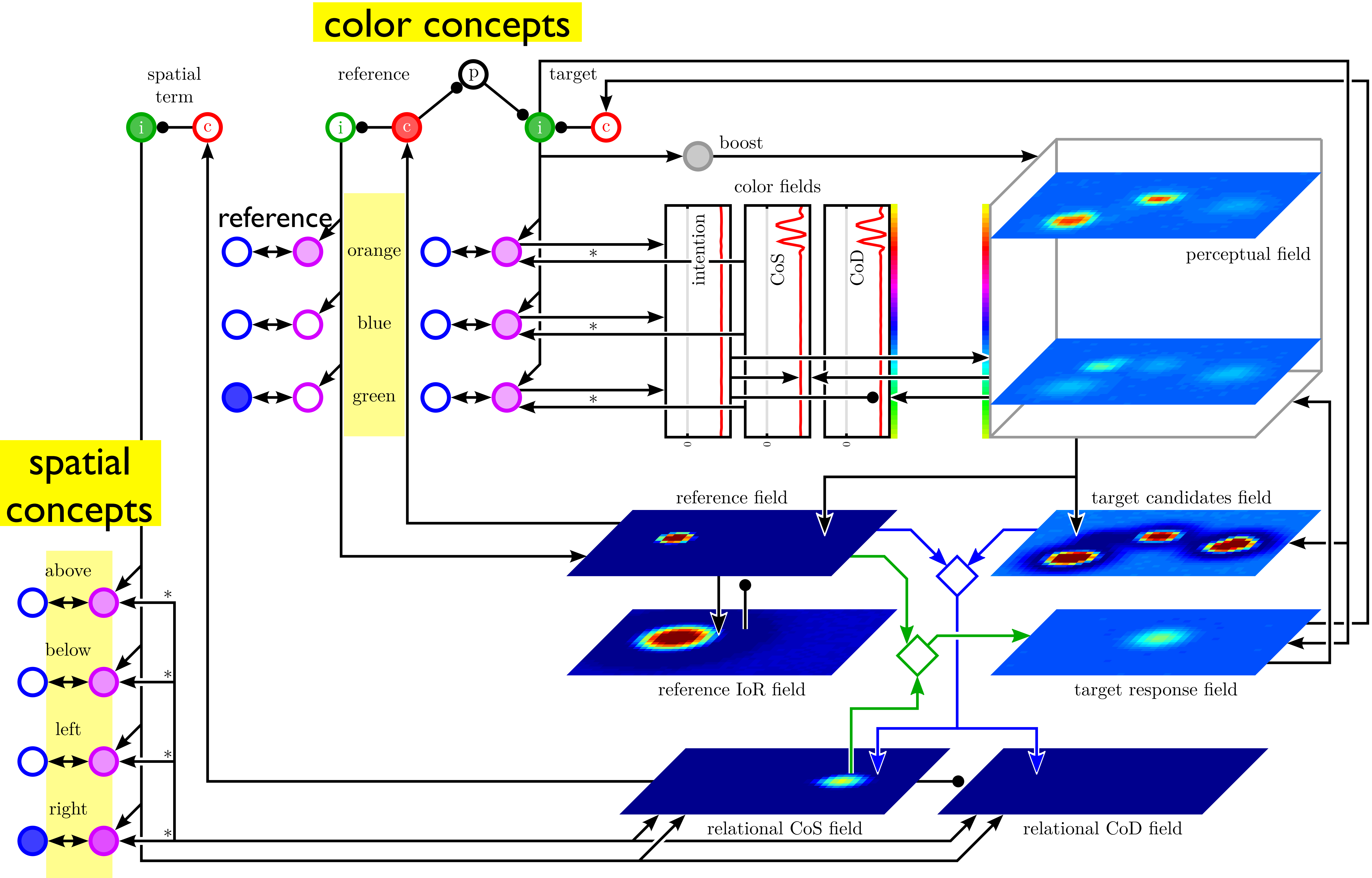
[Lipinski et al: JEP:LMC (2011)]

A DFT architecture that does both grounding and describing

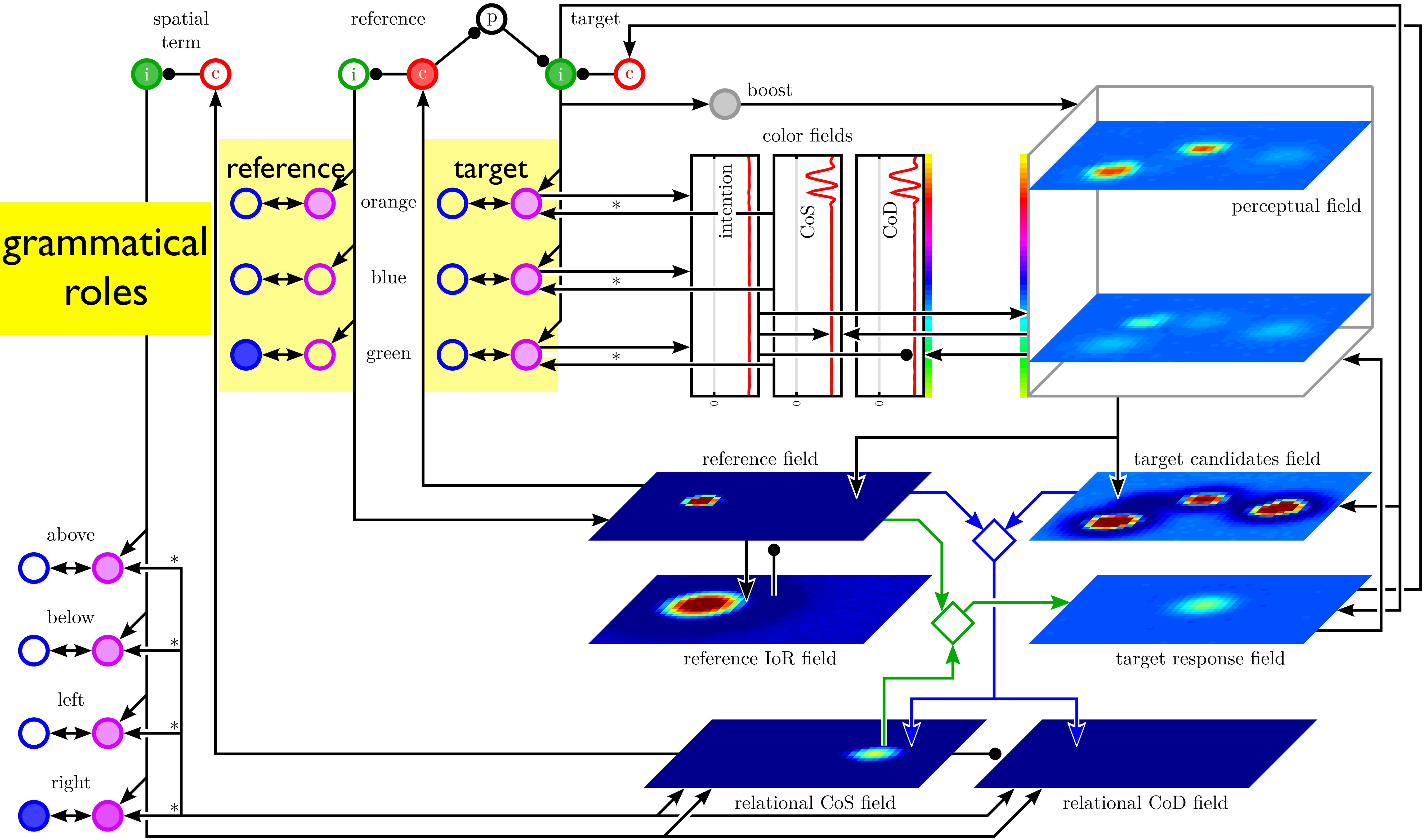




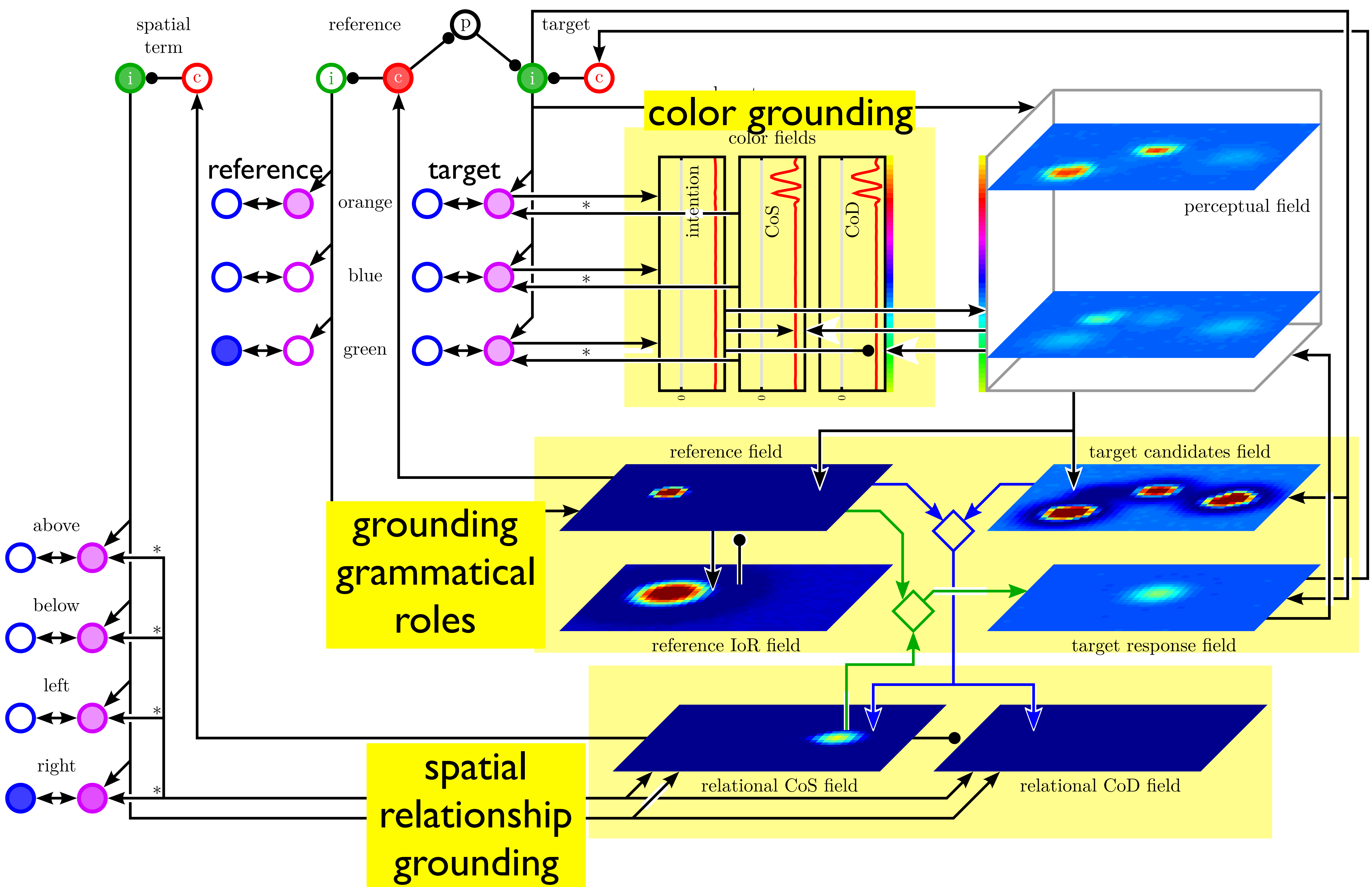
[Richter, Lins et al. ICANN 2014]



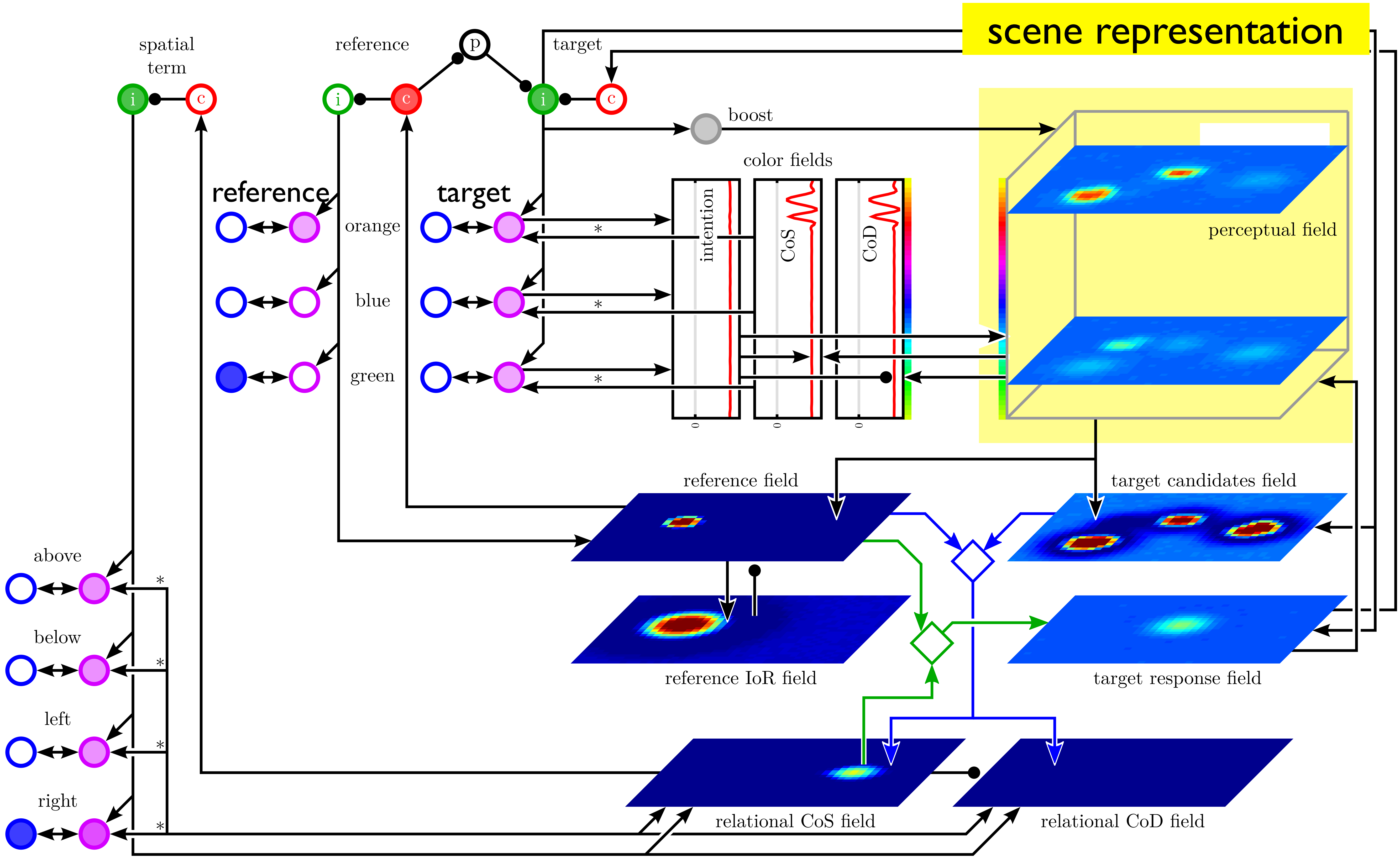
[Richter, Lins et al. ICANN 2014]



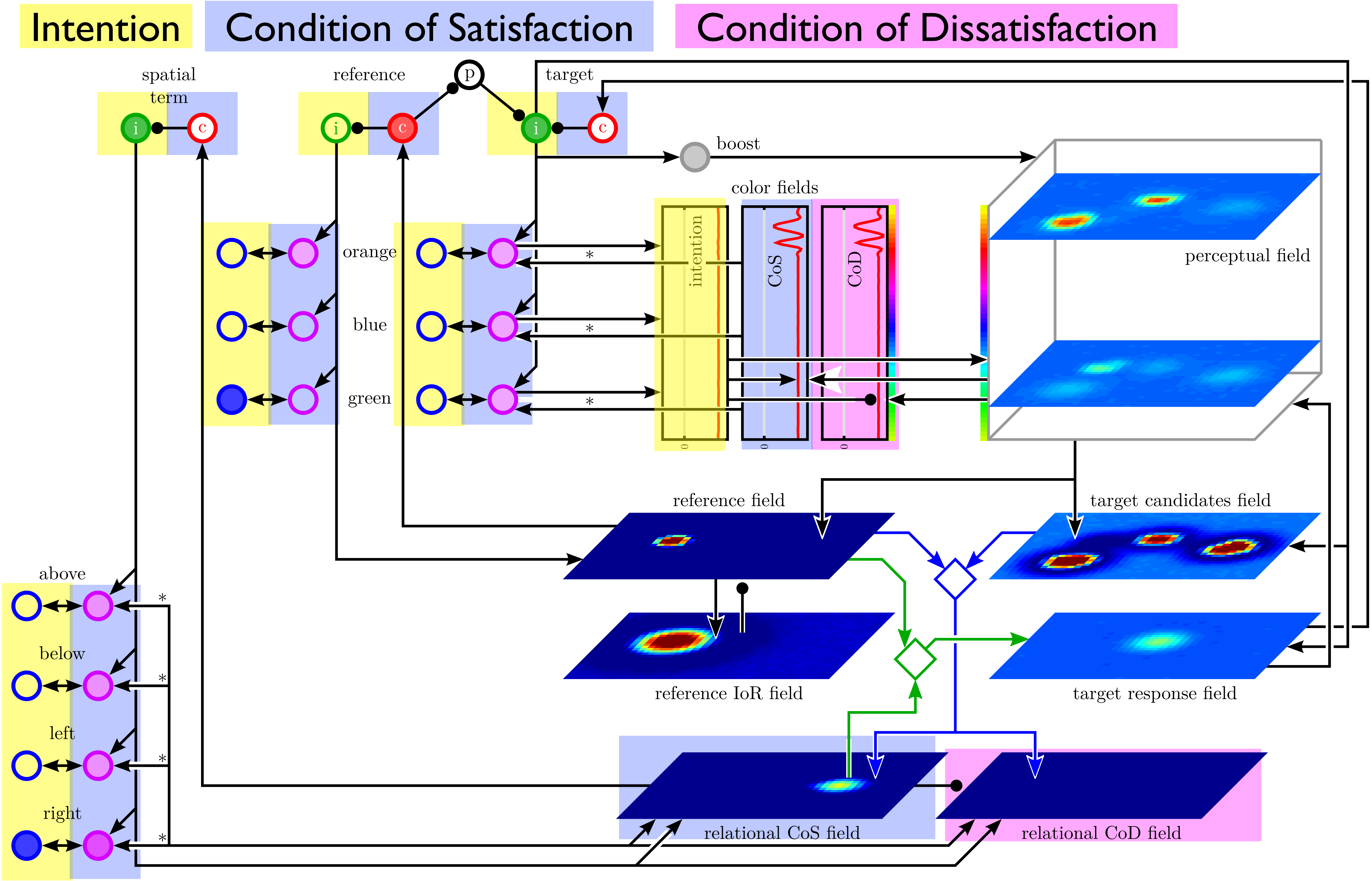
[Richter, Lins et al. ICANN 2014]



[Richter, Lins et al. ICANN 2014]

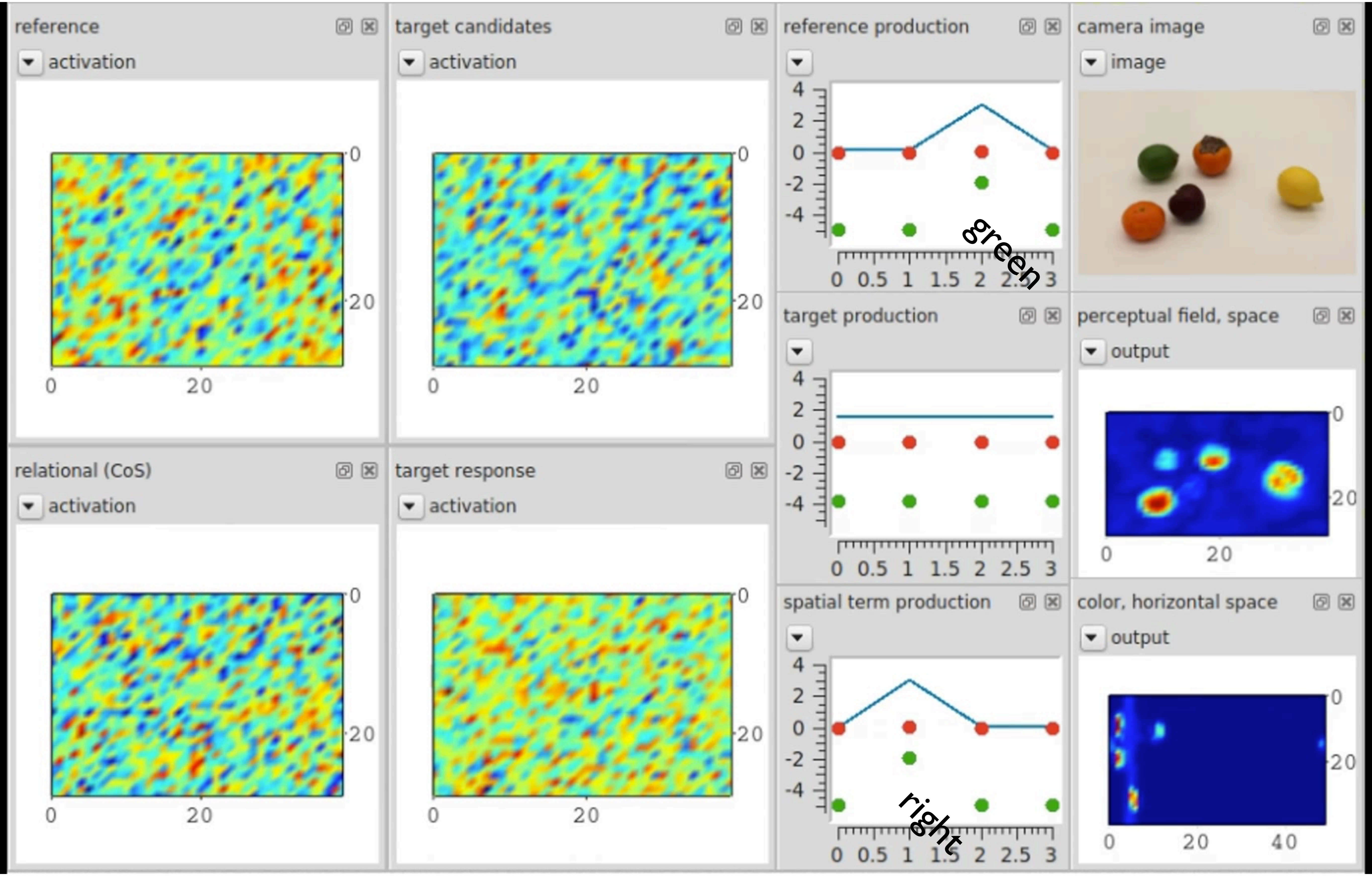


[Richter, Lins et al. ICANN 2014]

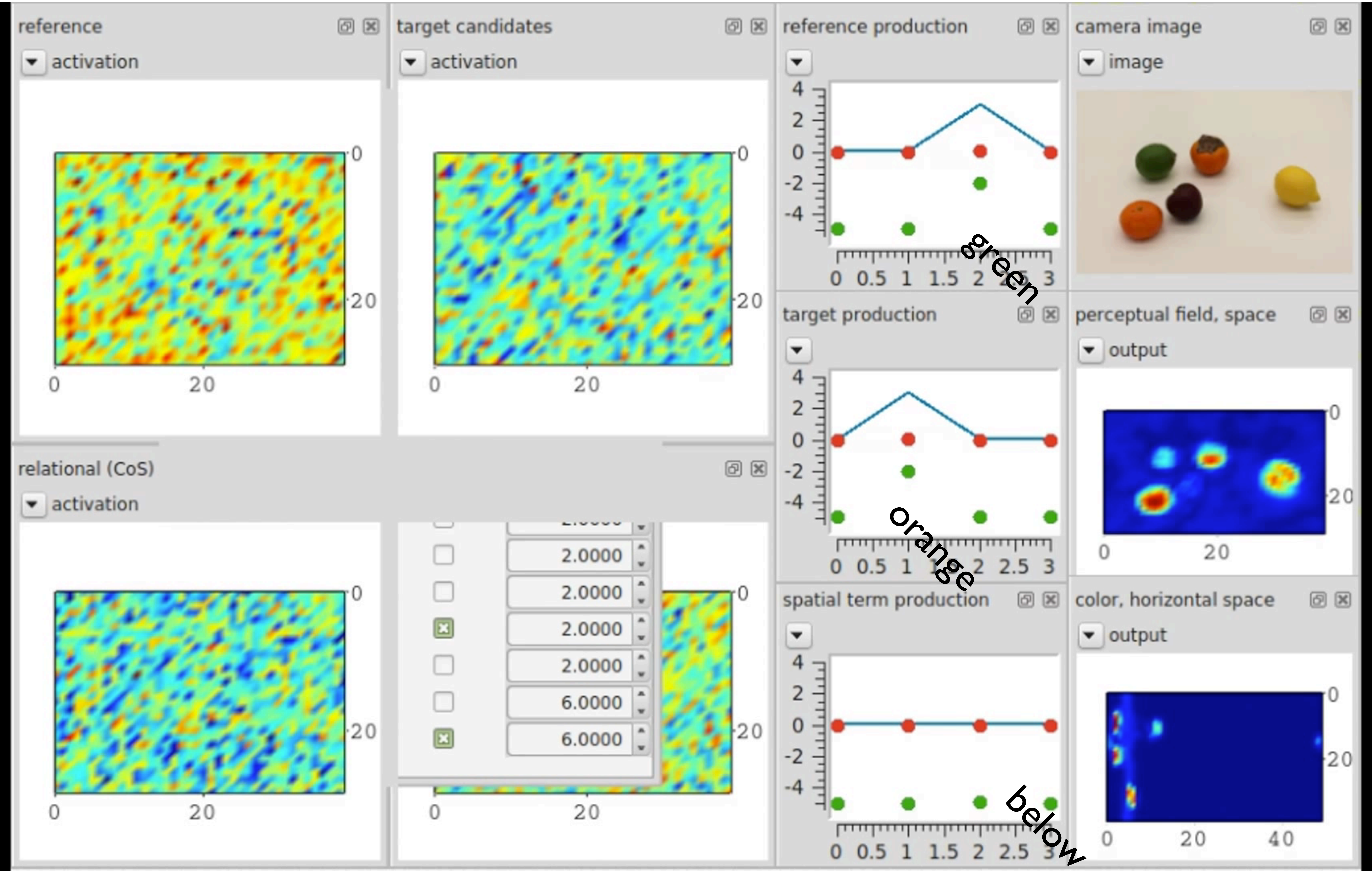


[Richter, Lins et al. ICANN 2014]

what is to the right of the green object?



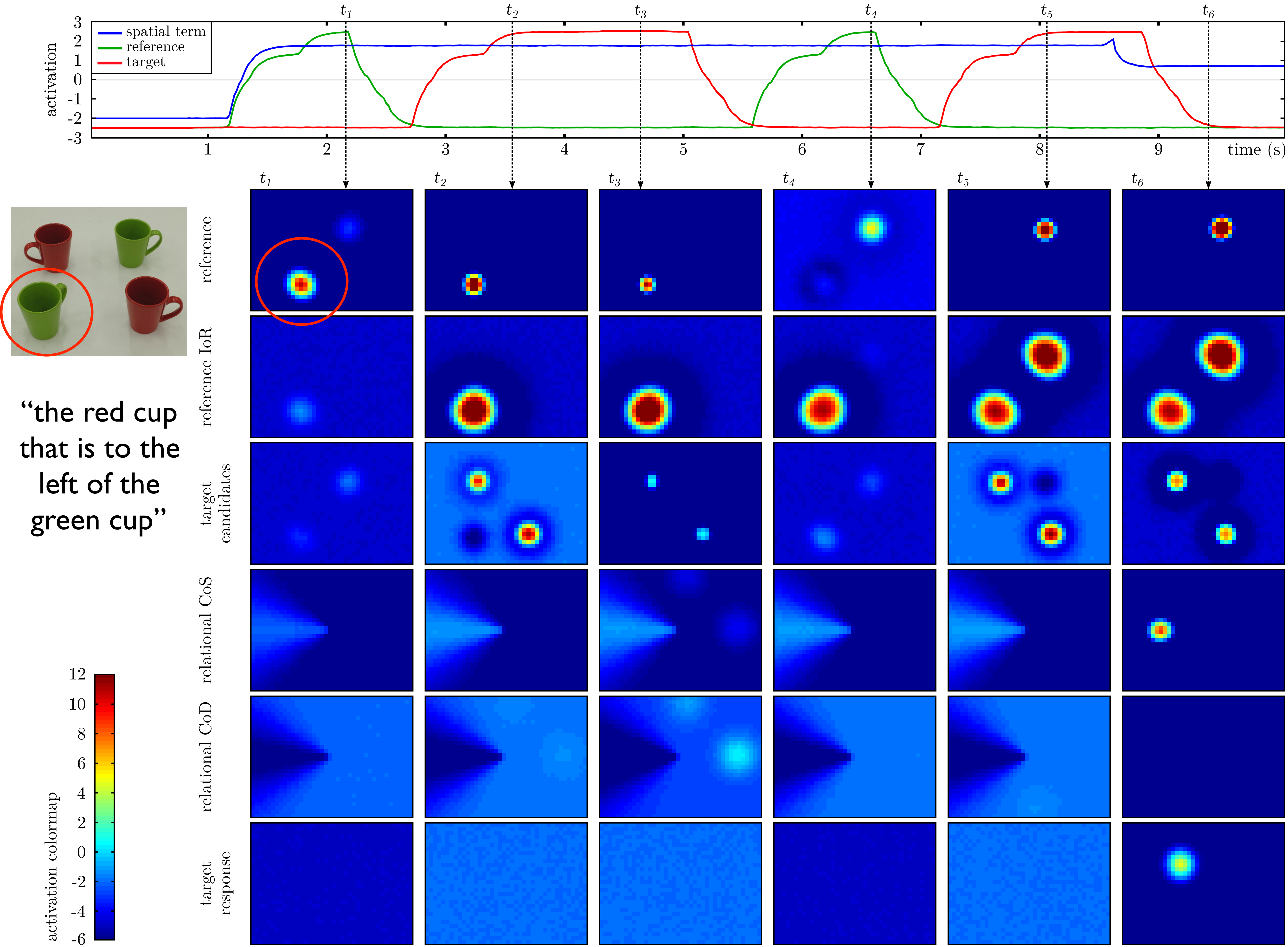
where is the orange relative to the green object



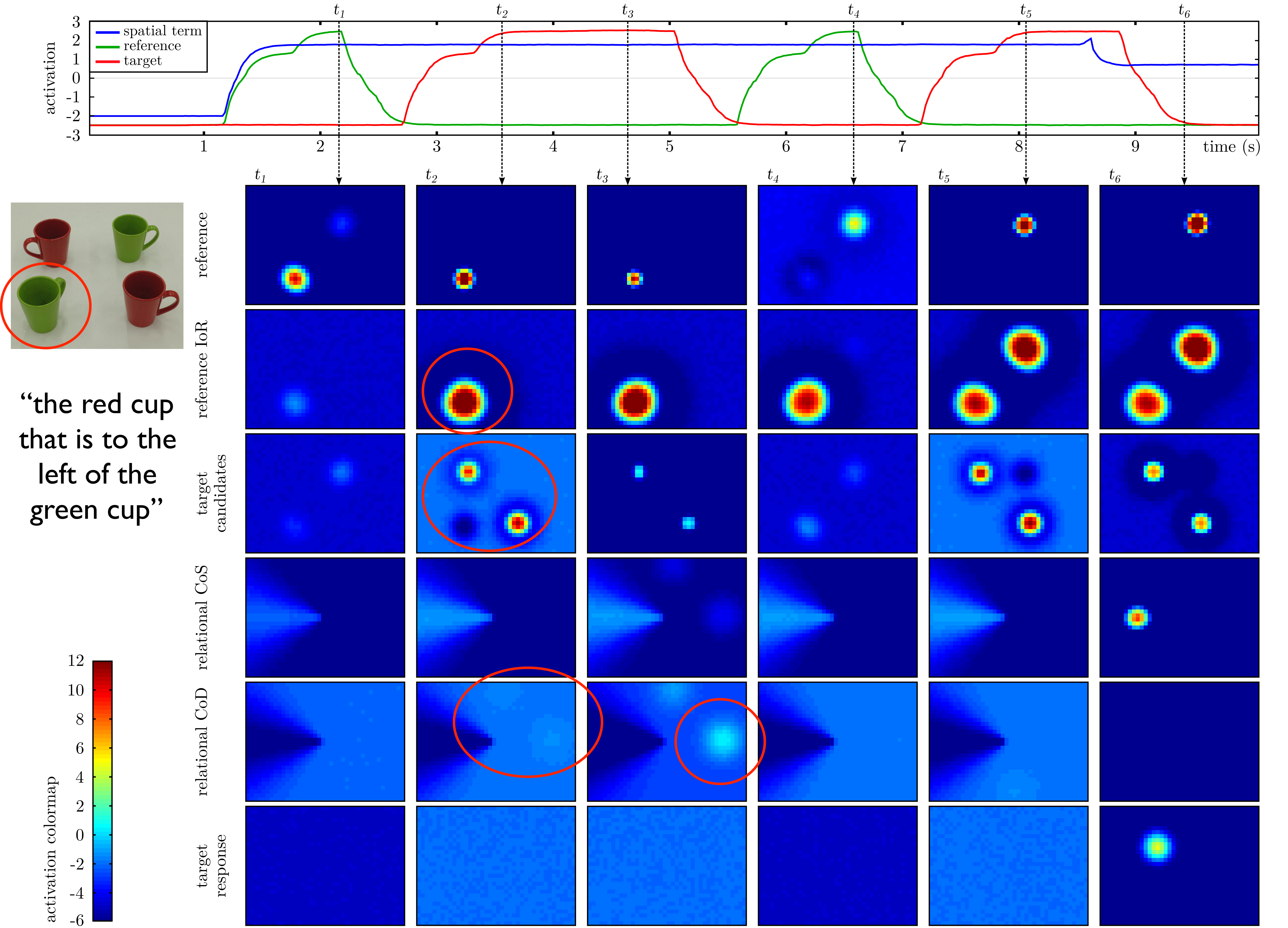
Autonomous hypothesis testing

"the red cup that is to the left of the green cup"

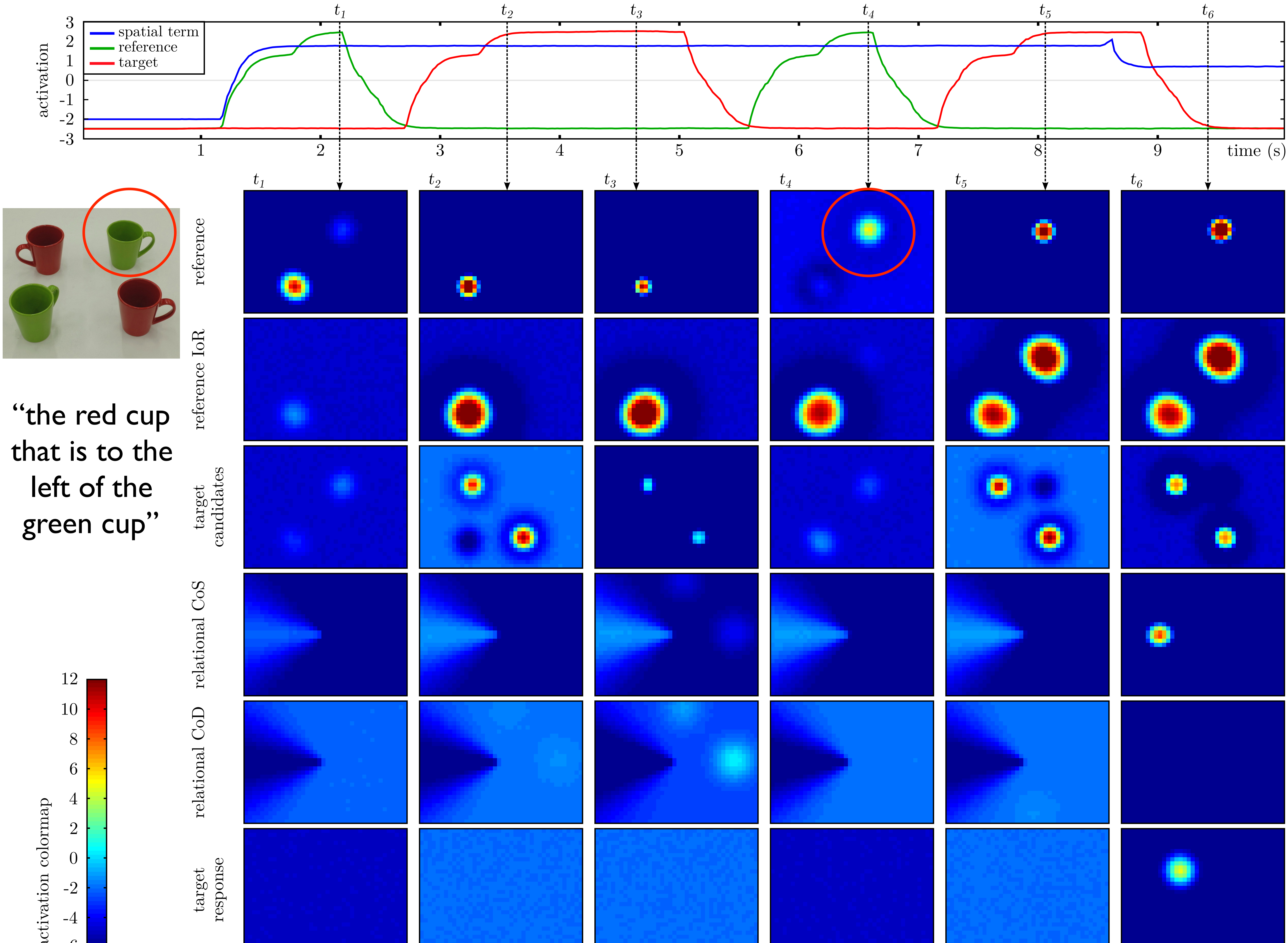
[Richter, Lins et al, CogSci 2014]



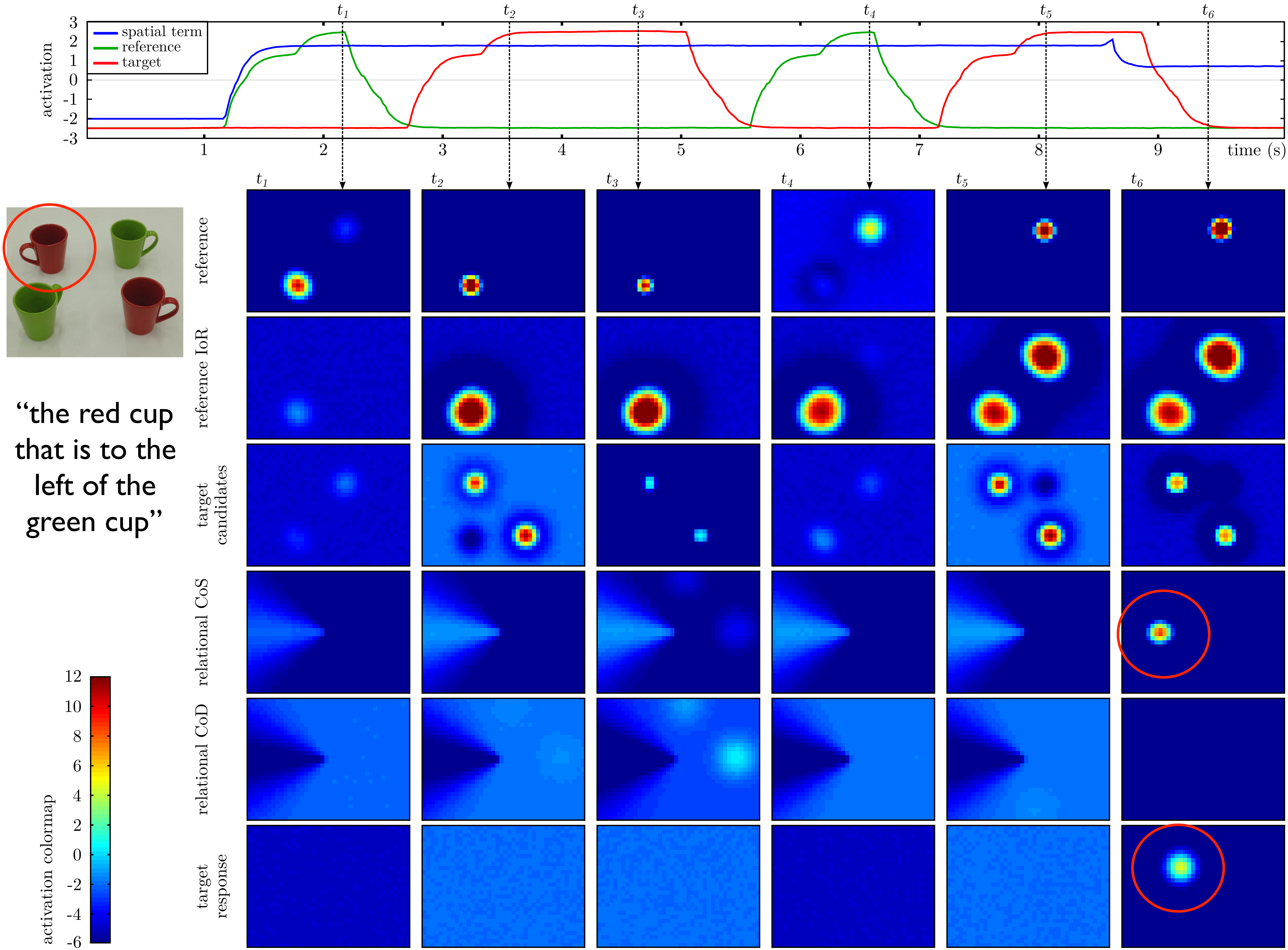
activation
spatial term
reference
target
time (s)
t_1 t_2 t_3 t_4 t_5 t_6
reference
reference IoR
target candidates
relational CoS
relational CoD
target response
"the red cup that is to the left of the green cup"
activation colormap



spatial term
reference
target
activation
time (s)
t_1 t_2 t_3 t_4 t_5 t_6
reference
reference IoR
target candidates
relational CoS
relational CoD
target response
"the red cup that is to the left of the green cup"
activation colormap

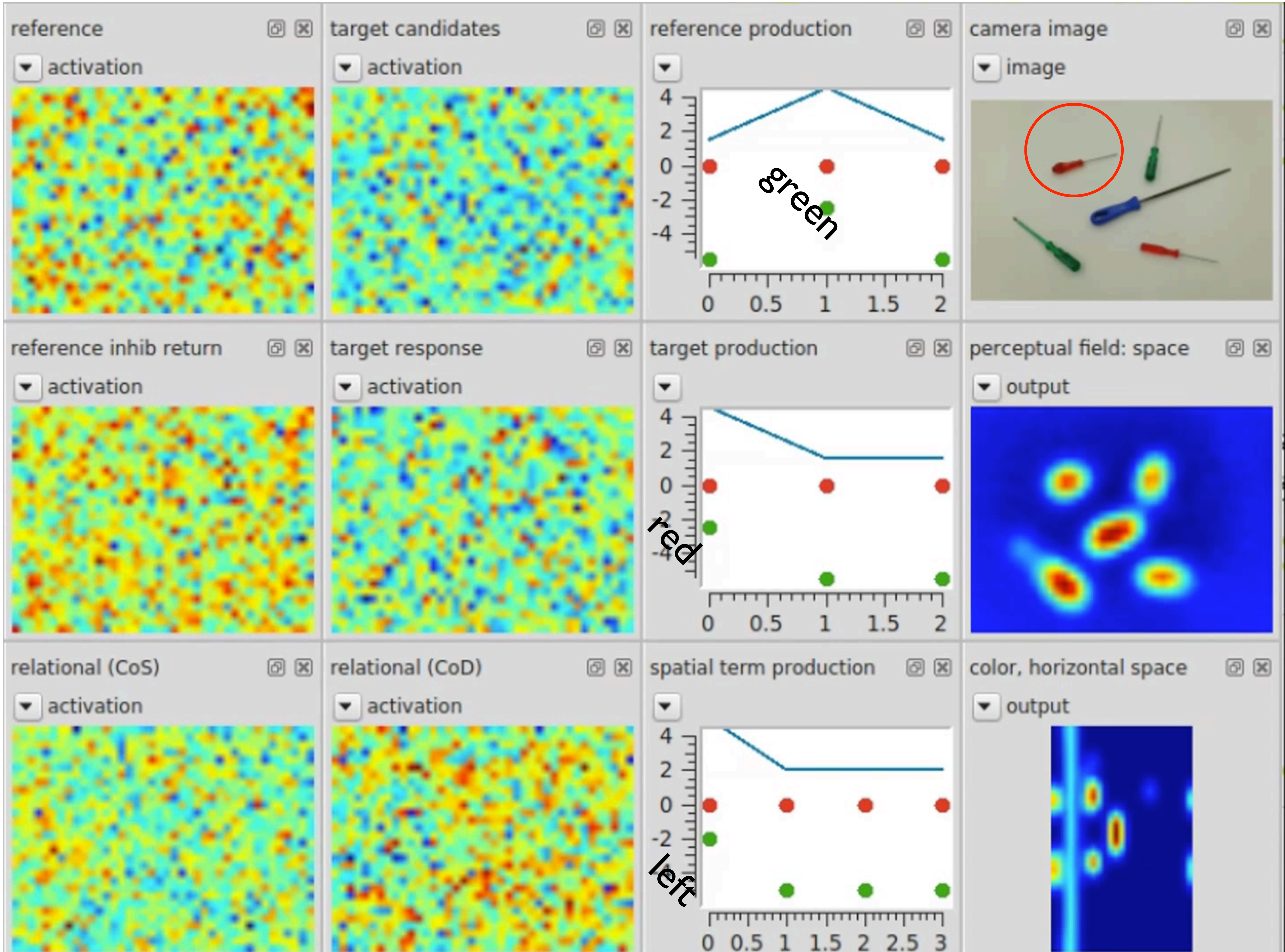


spatial term
reference
target
activation
time (s)
t1 t2 t3 t4 t5 t6
"the red cup that is to the left of the green cup"
reference
reference IoR
target candidates
relational CoS
relational CoD
target response
activation colormap

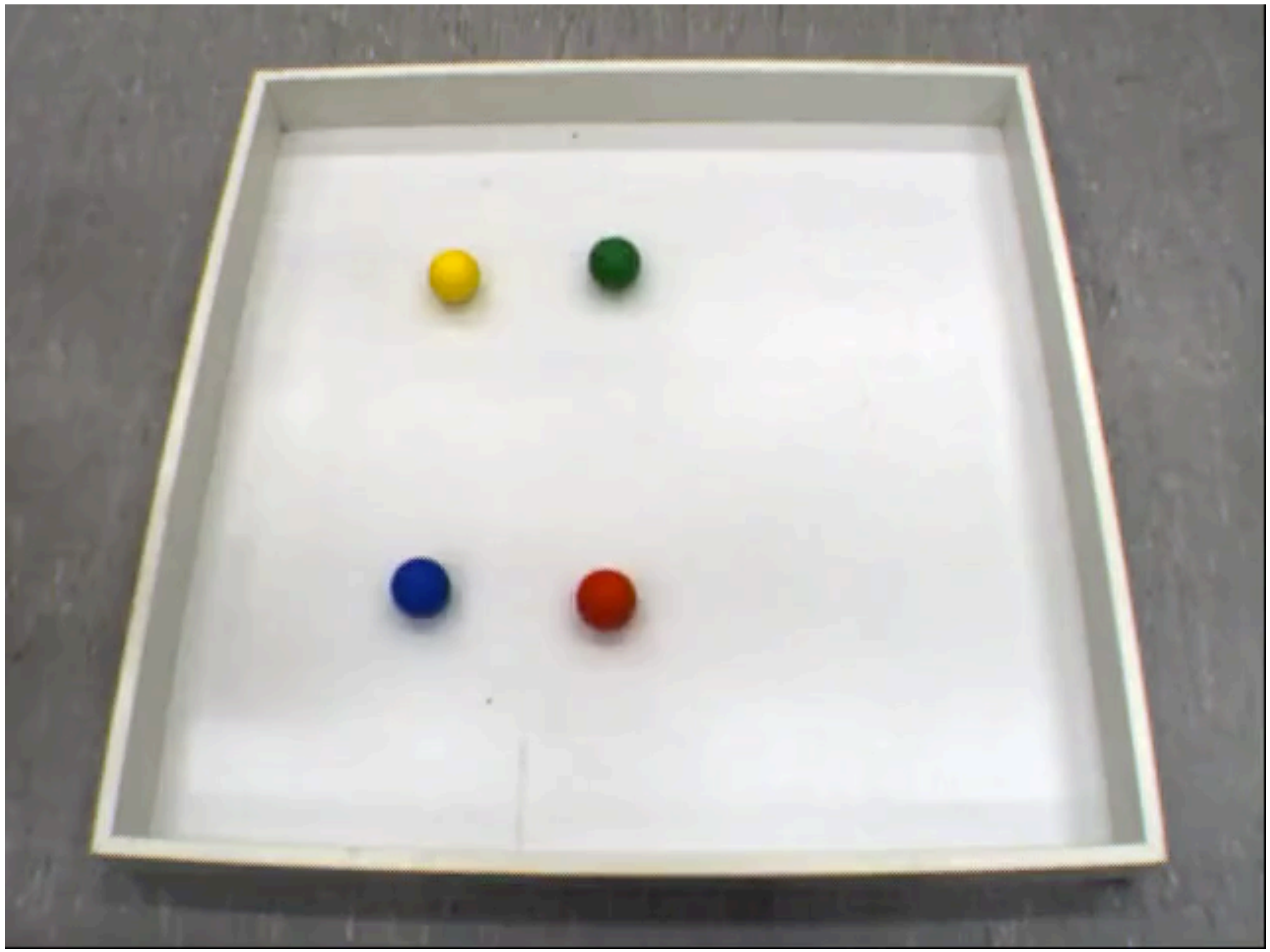


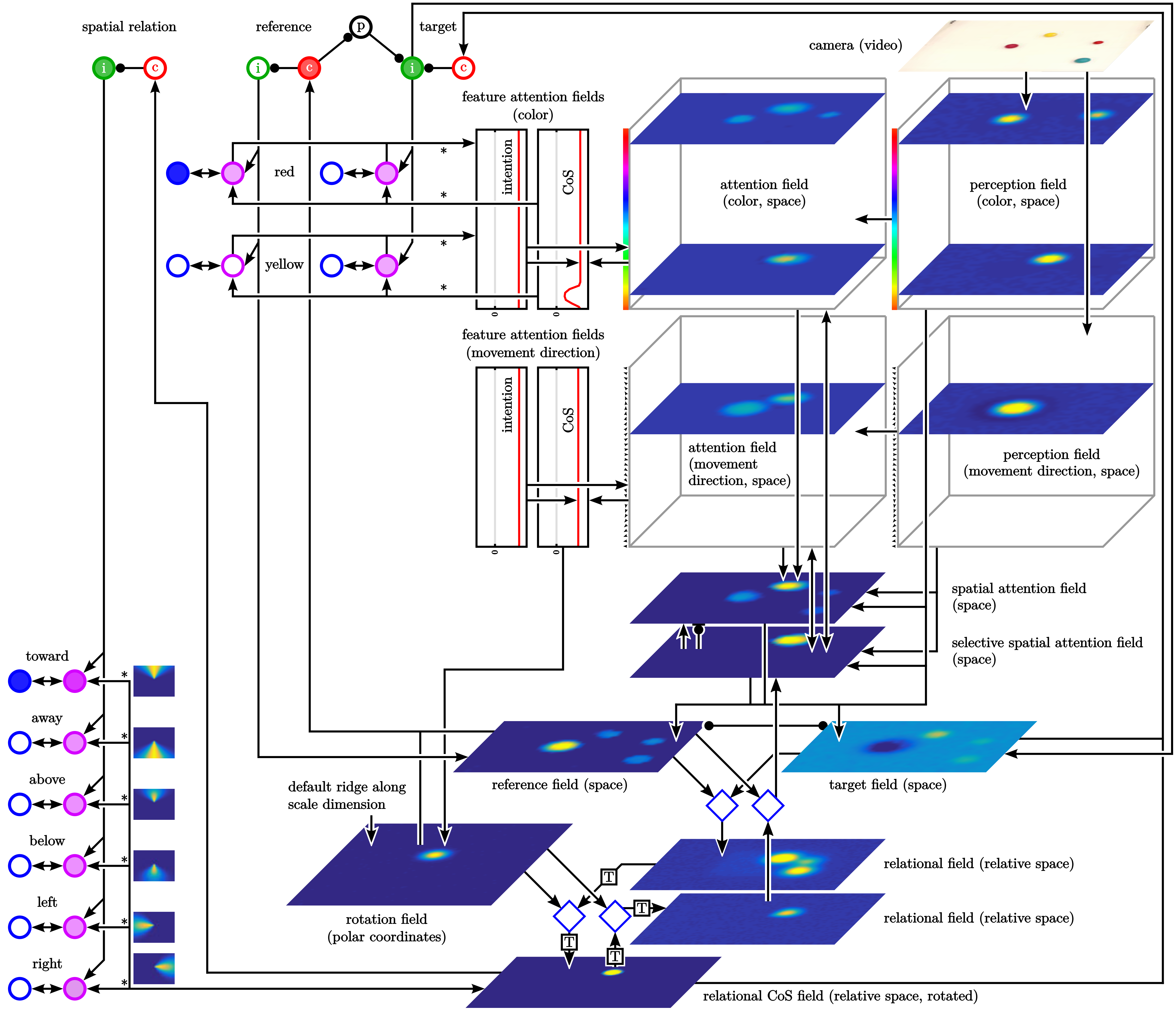
spatial term
reference
target
activation
time (s)
t_1
t_2
t_3
t_4
t_5
t_6
reference
reference IoR
target candidates
relational CoS
relational CoD
target response
"the red cup that is to the left of the green cup"
activation colormap

"the red to the left of the green"

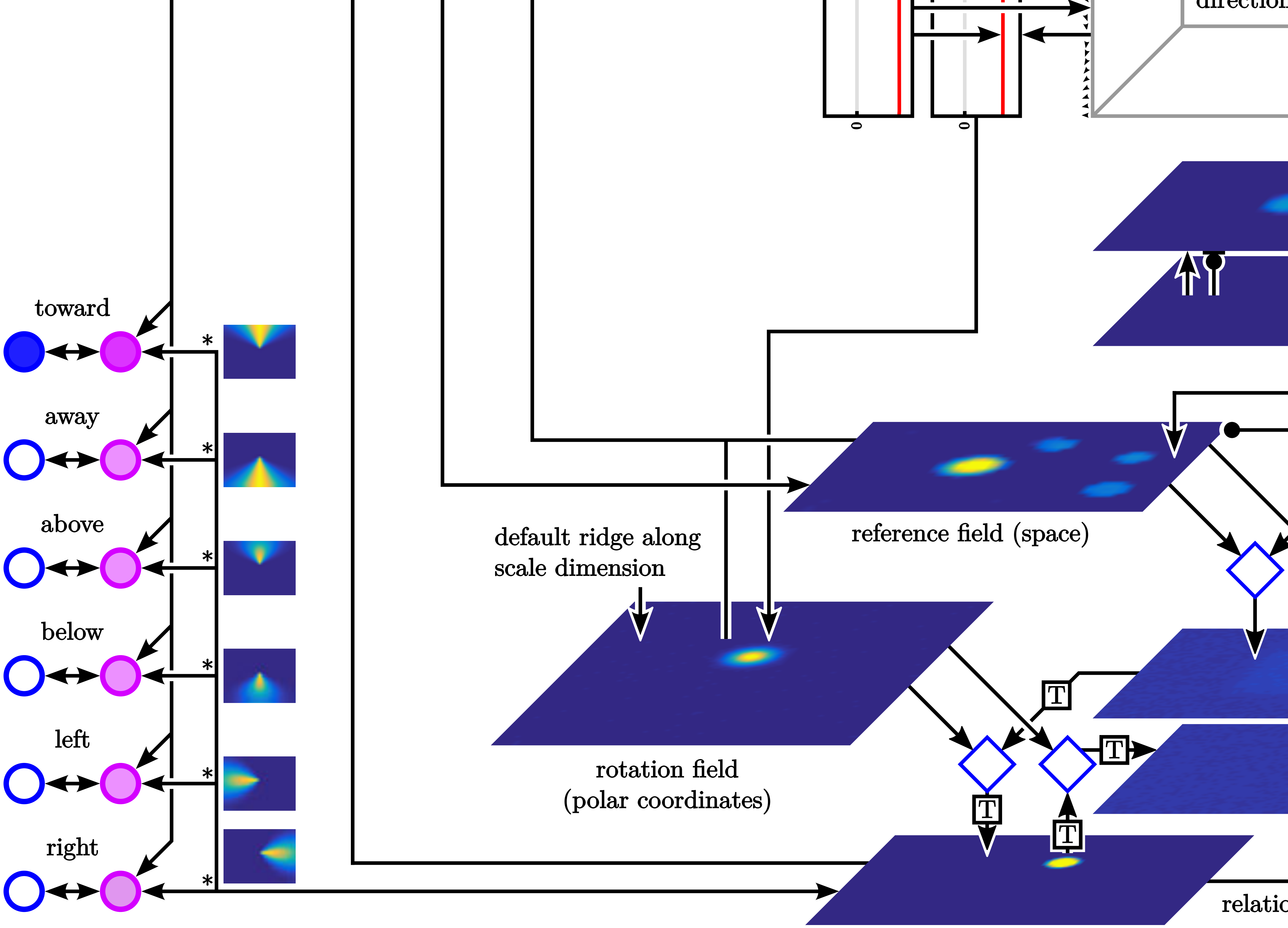


Grounding movement relations





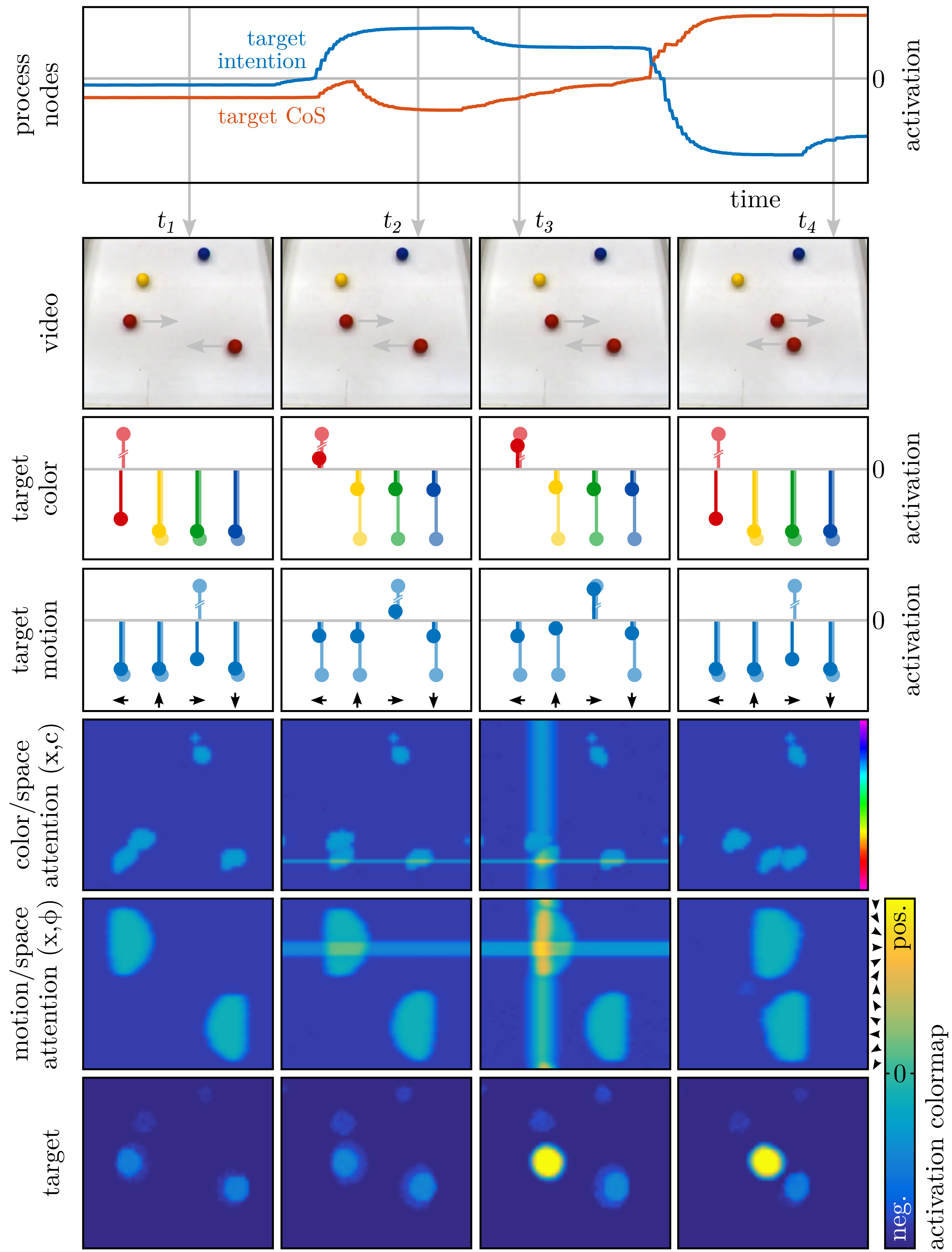
[Richter, Lins, Schöner, ToPiC (2017)]



[Richter, Lins, Schöner, ToPiC (2017)]

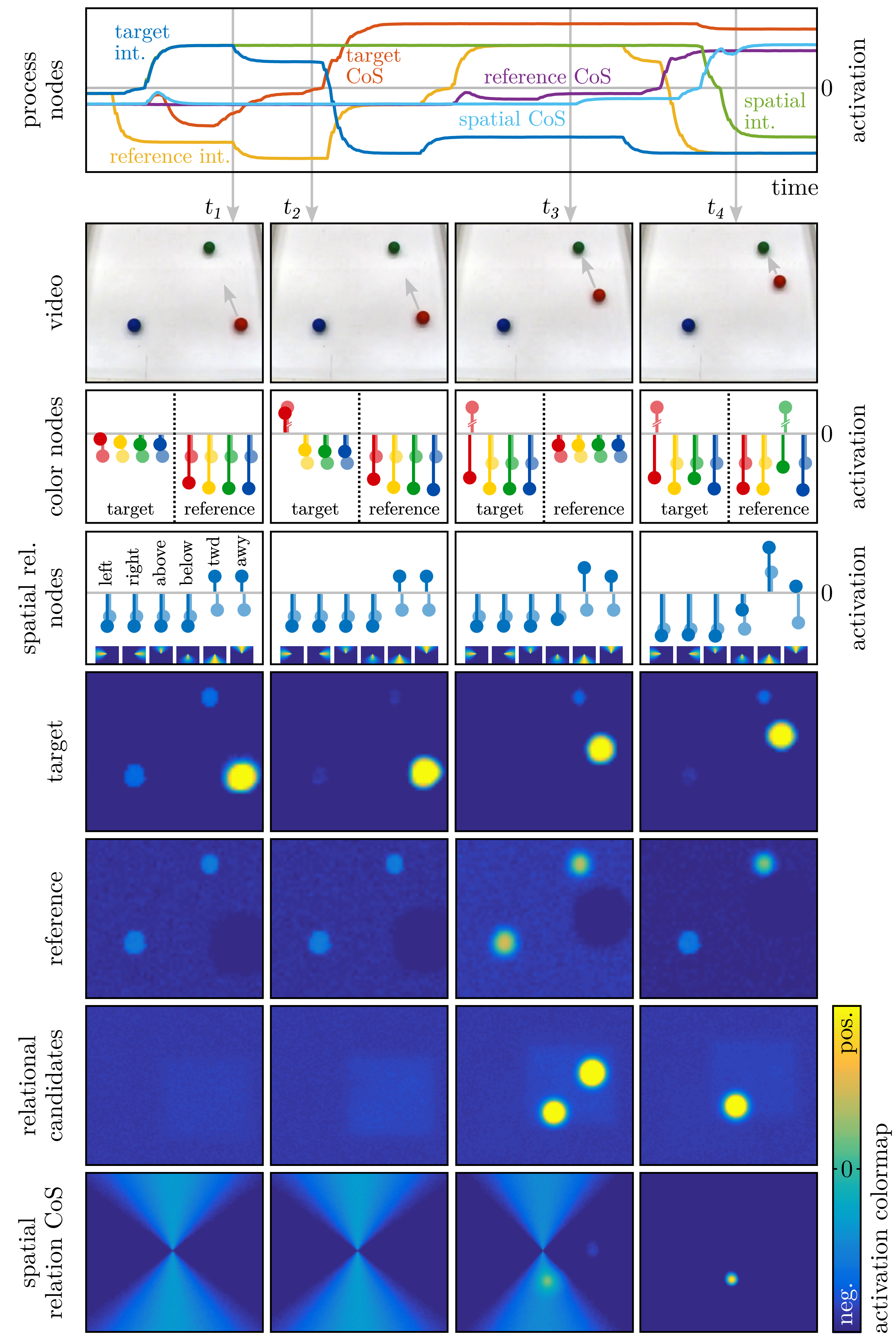
“the red moving to the right”

perceptual grounding



[Richter et al]

description



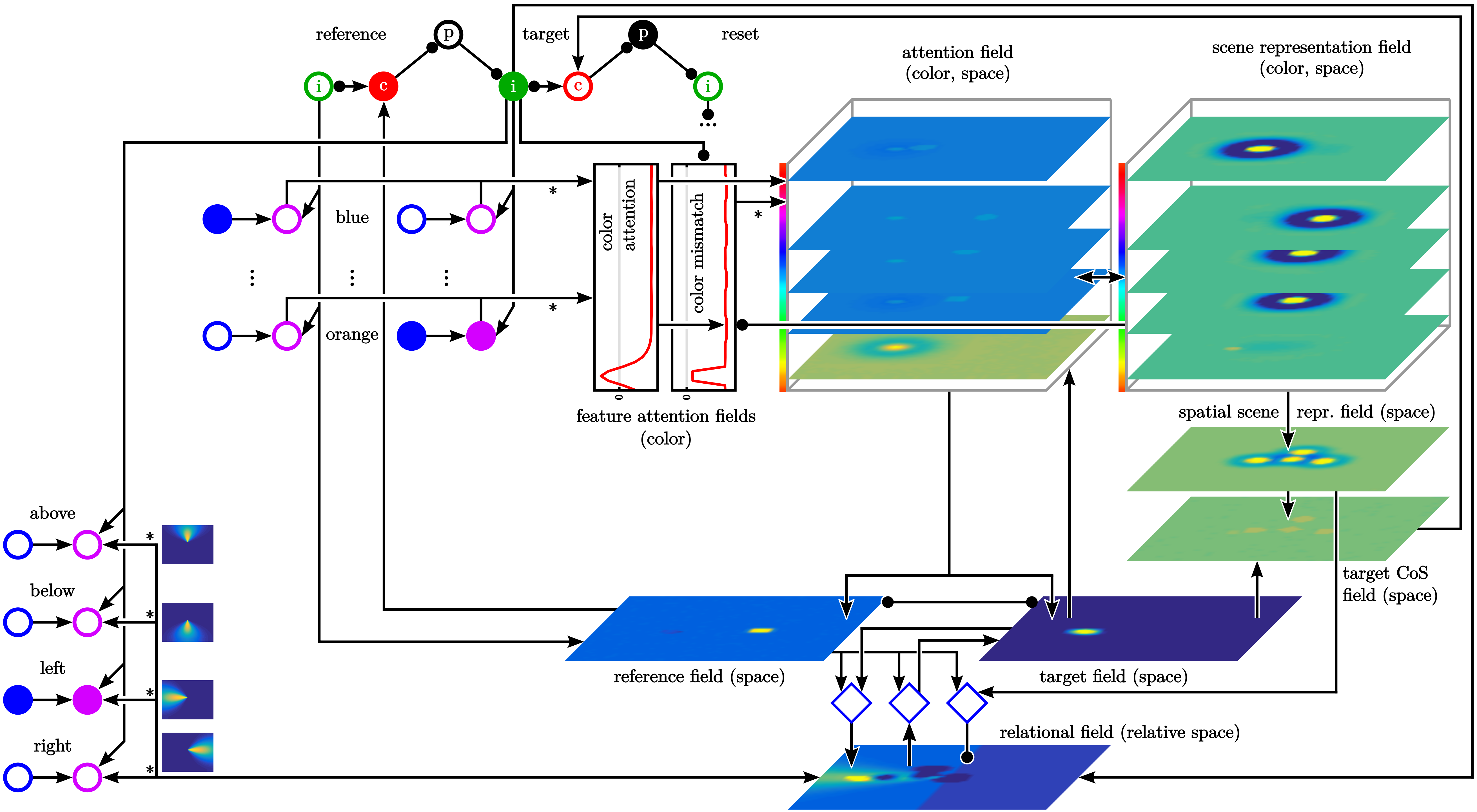
[Richter et al]

Mental mapping and inference

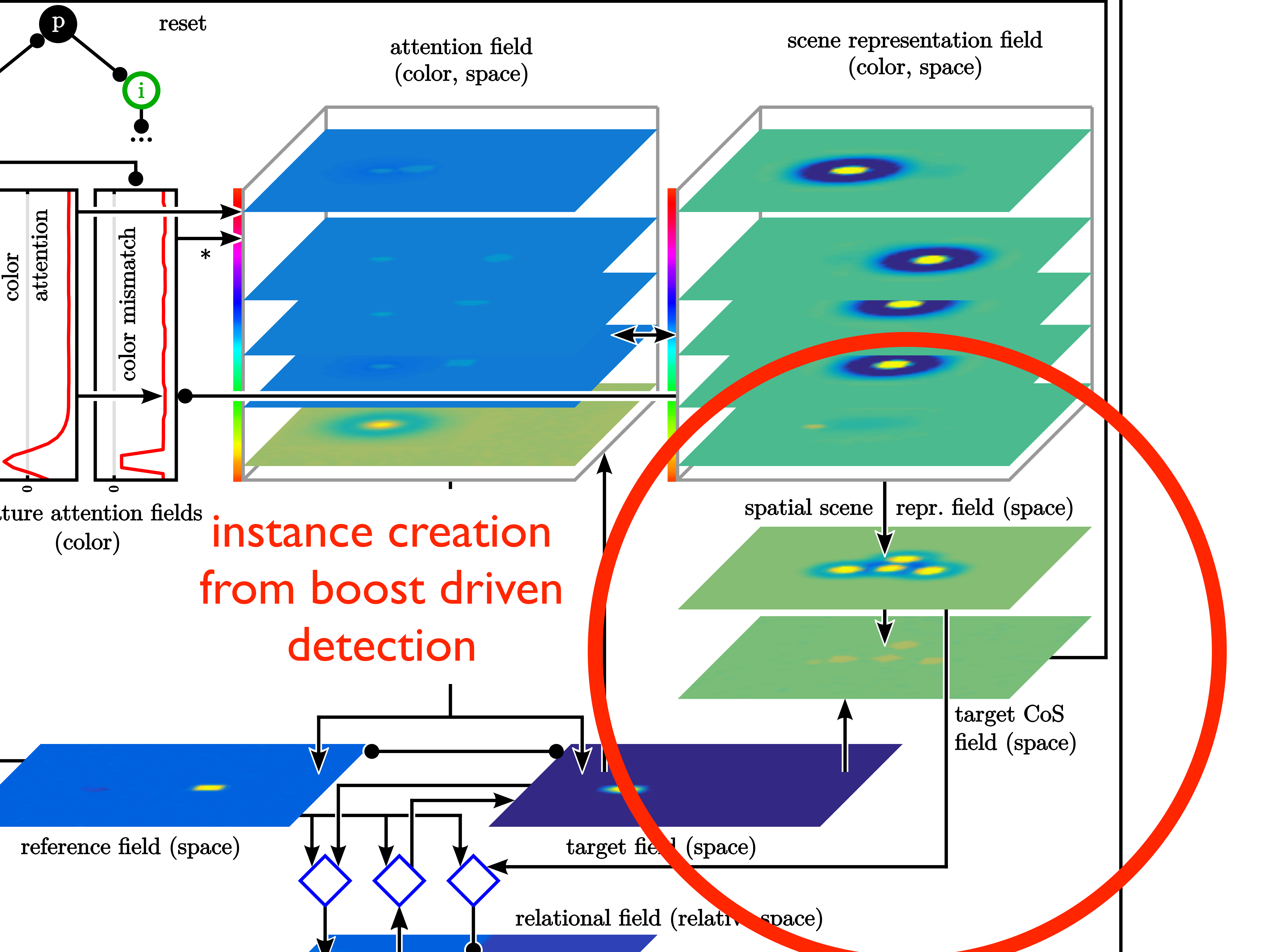
- making sense of propositions (about spatial relations) purely mentally, without any perception to ground in
- and operating on such “sense” by drawing inferences…

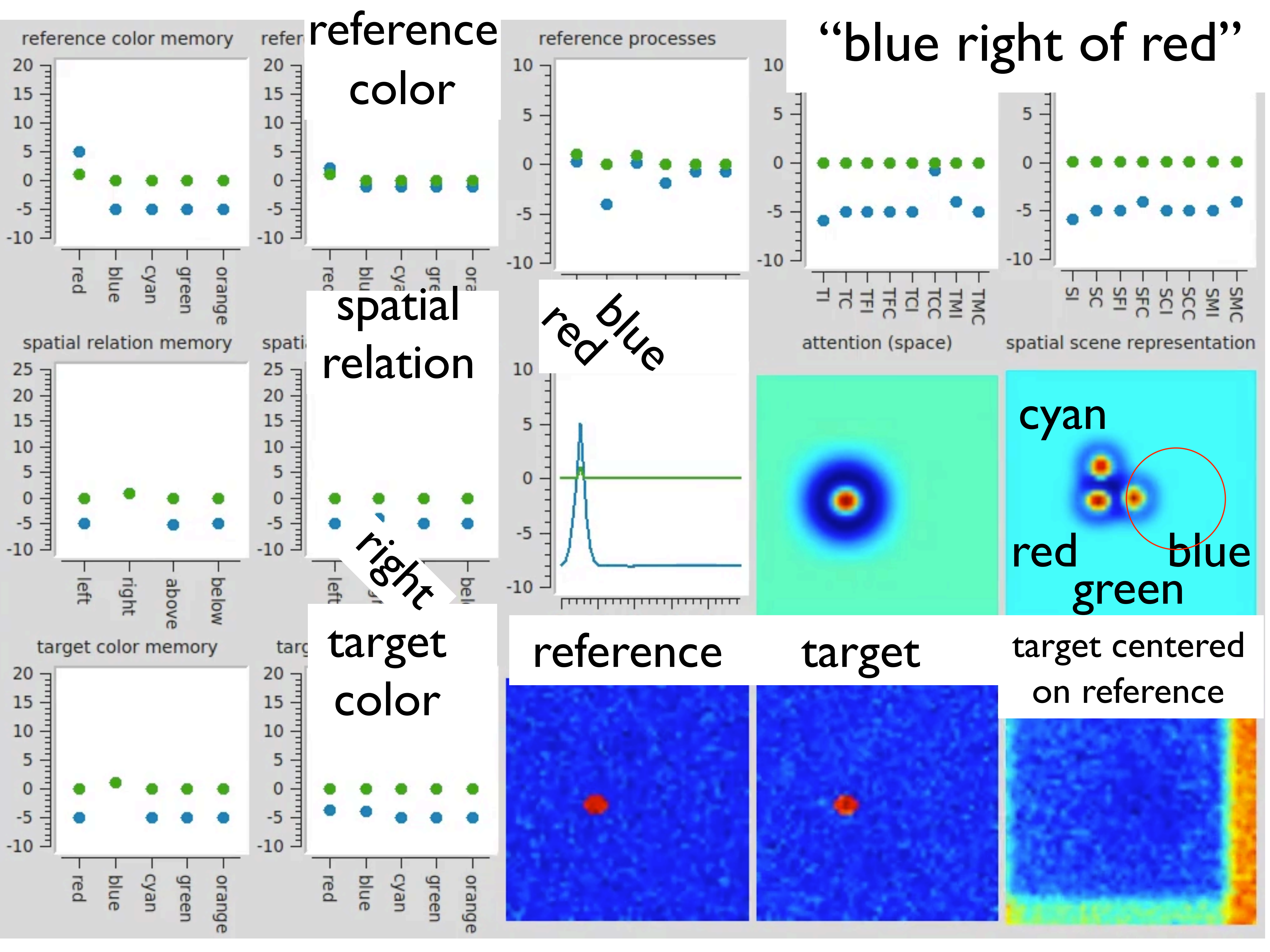
Mental mapping and inference

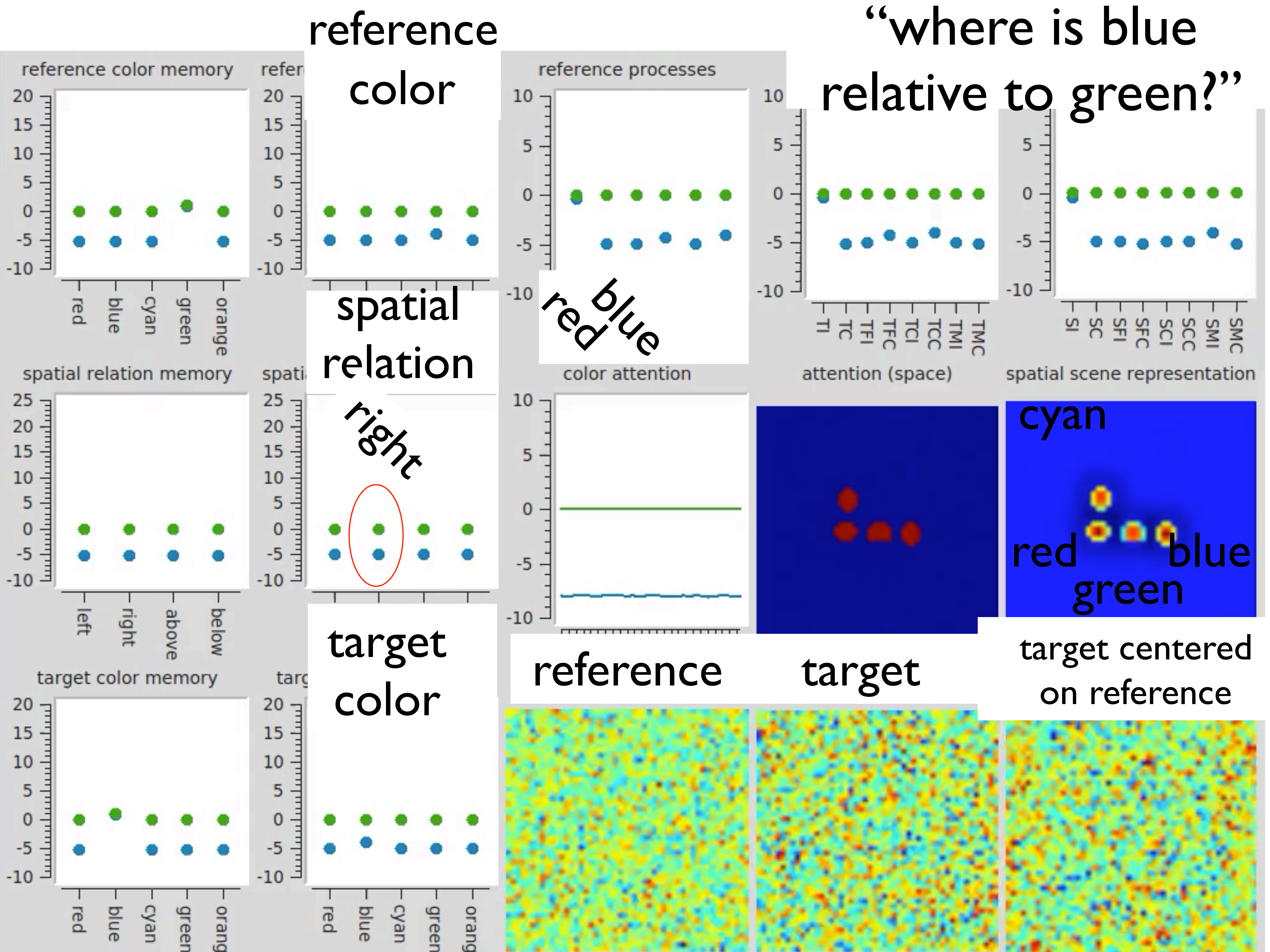
- mental map formation from propositions
  - “There is a cyan object above a green object.”
  - “There is a red object to the left of the green object.”
  - “There is a blue object to the right of the red object.”
  - “There is an orange object to the left of the blue object.”
- inference
  - “Where is the blue object relative to the red object?”



[Kounatidou, Richter, Schöner, CogSci 2018]


p
reset
i
attention field
(color, space)
scene representation field
(color, space)
color
attention
color mismatch
*
0
0
ture attention fields
(color)
instance creation from boost driven detection
spatial scene
repr. field (space)
target CoS
field (space)
reference field (space)
target fie d (space)
relational field (relati space)


"blue right of red"
reference color memory
reference color
reference processes
spatial relation memory
spatial relation
red
blue
attention (space)
spatial scene representation
cyan
red
blue
green
right
target color memory
target color
reference
target
target centered on reference
red
blue
cyan
green
orange
left
right
above
below
TI
TC
TFI
TFC
TCI
TCC
TMI
TMC
SI
SC
SFI
SFC
SCI
SCC
SMI
SMC



"where is blue relative to green?"
reference color
spatial relation
right
target color
blue
red
reference color memory
reference processes
spatial relation memory
color attention
attention (space)
spatial scene representation
target color memory
cyan
red
blue
green
reference
target
target centered on reference

Conclusion

- … higher dimensional fields
- arranged in architectures…
- deliver higher cognitive functions
- such as perceptual grounding, describing scenes, mental imaging, and inference