

An Experimental Assessment of Embodied Spatial Language Grounding

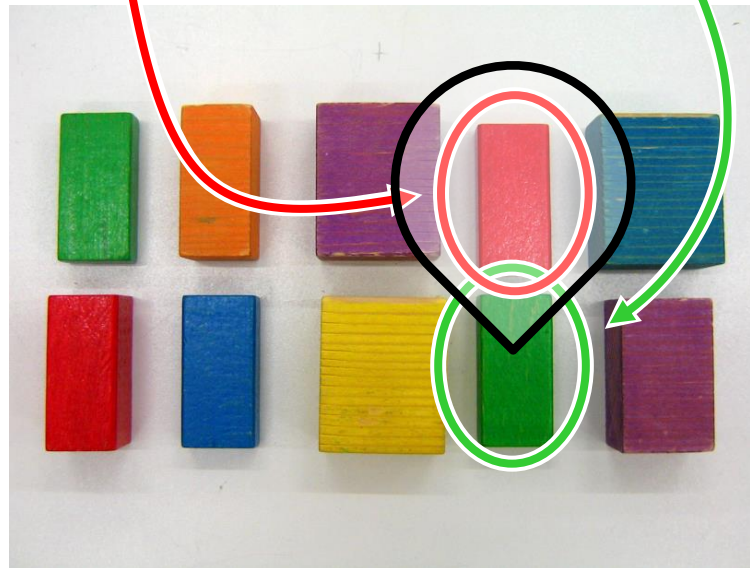
JONAS LINS

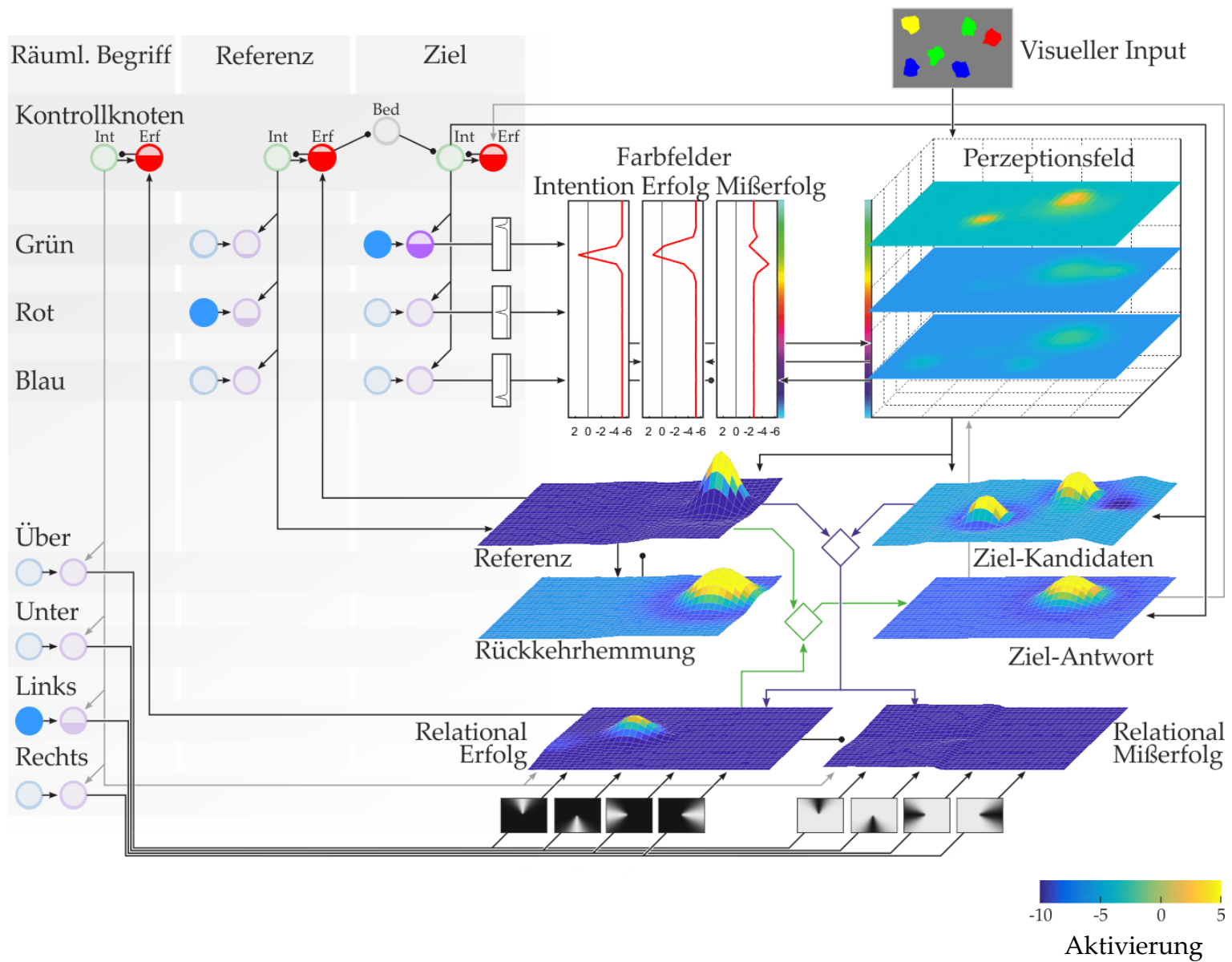
Target

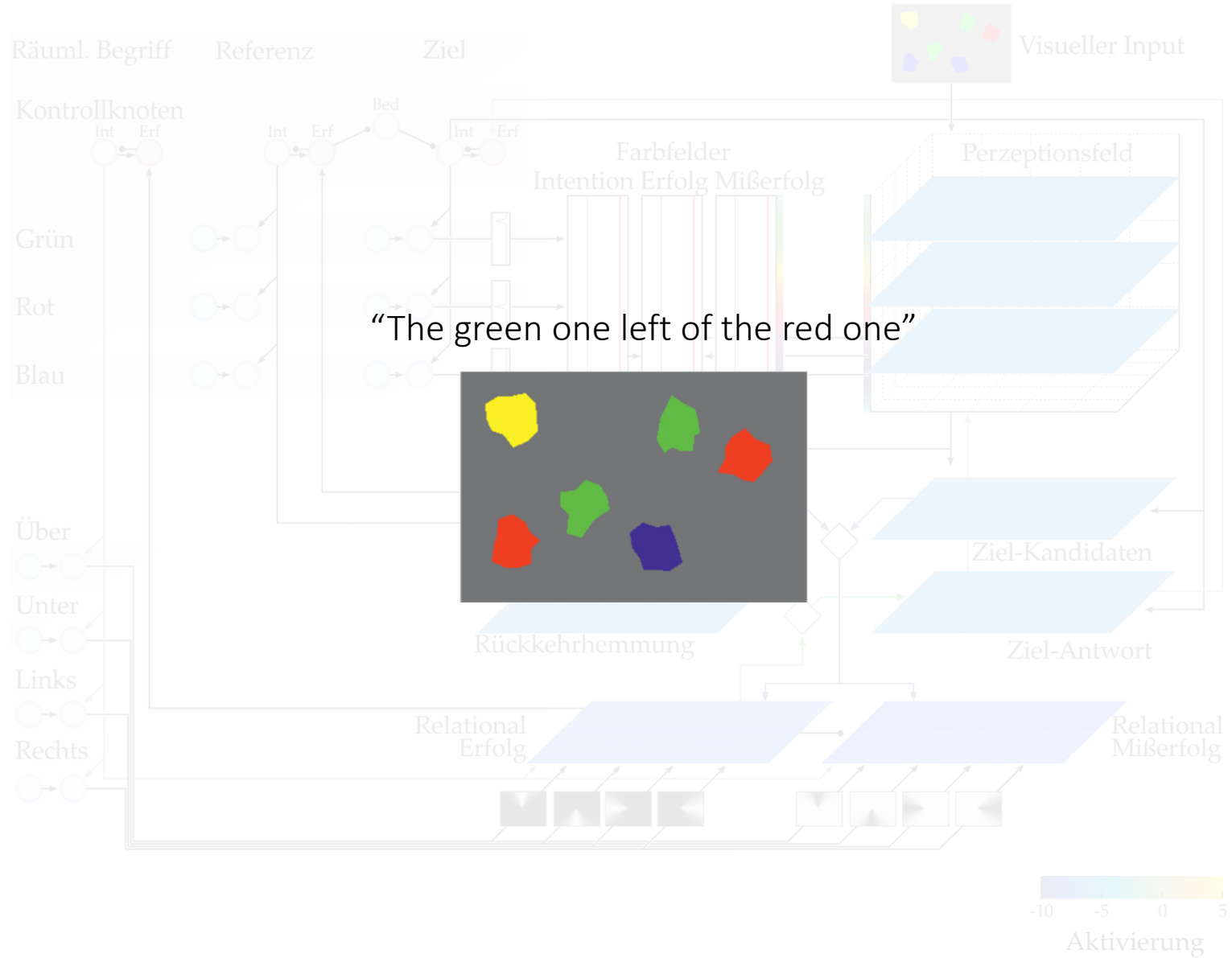
Spatial term

Reference

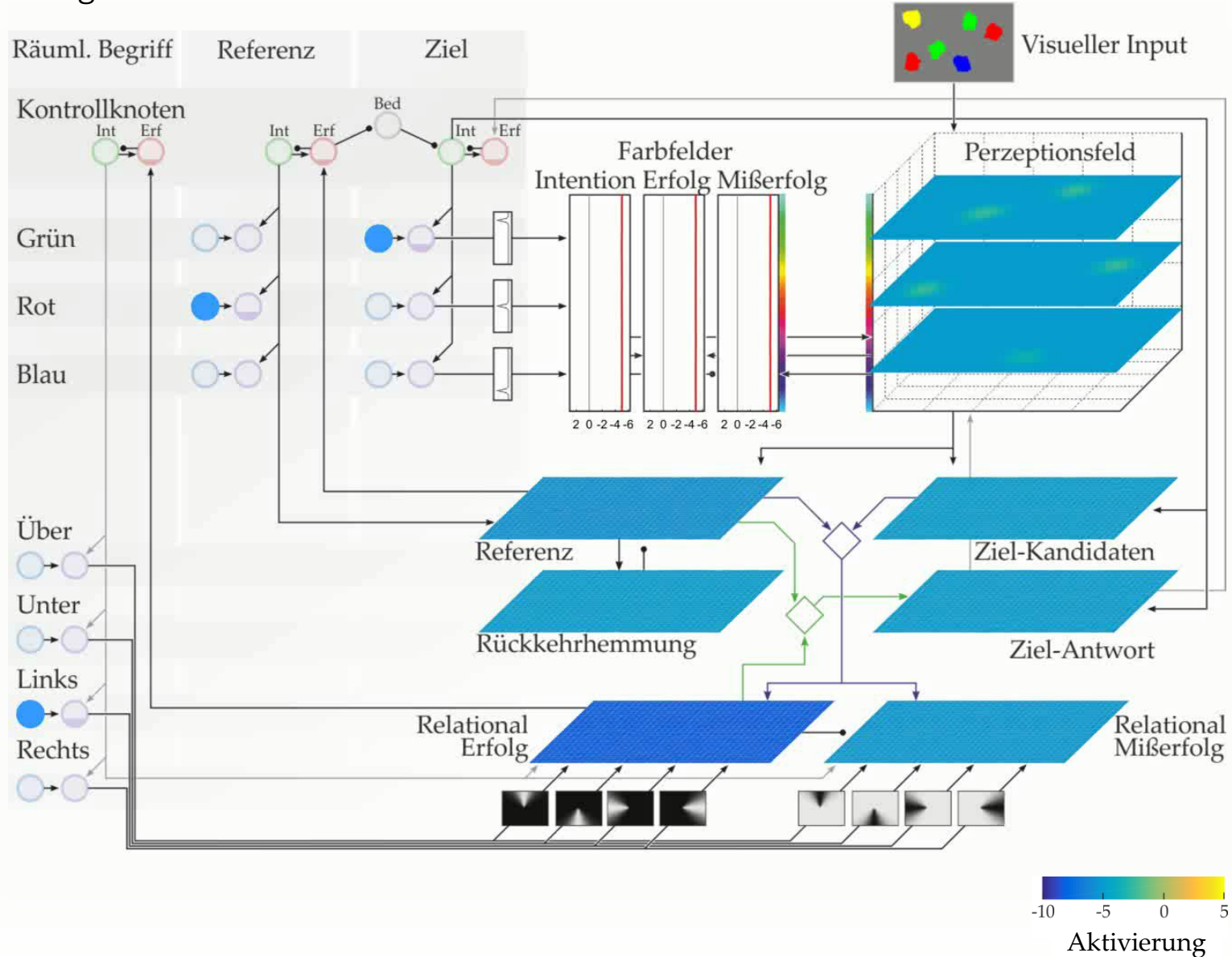
„The red block above the green block.“



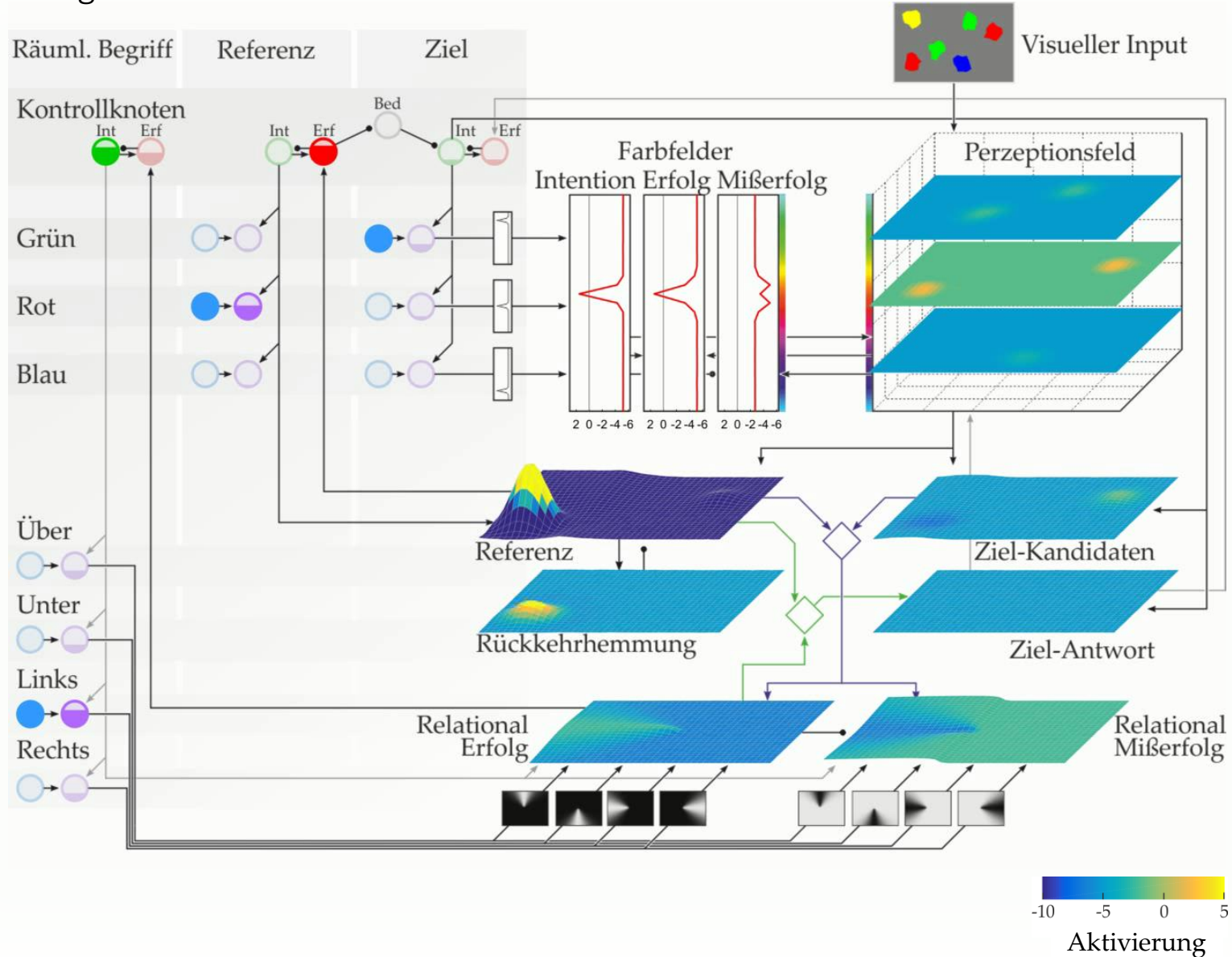




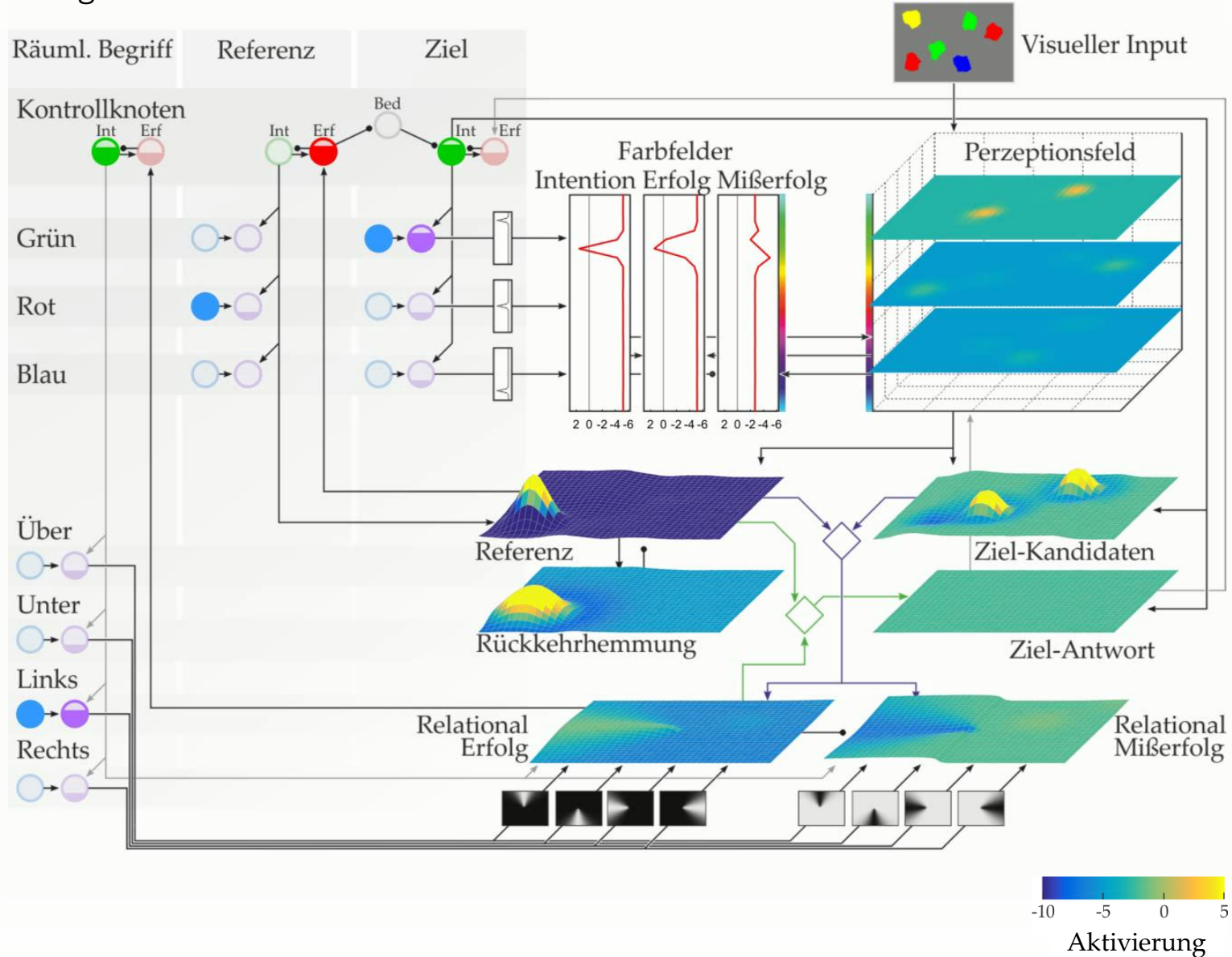
“The green one left of the red one”



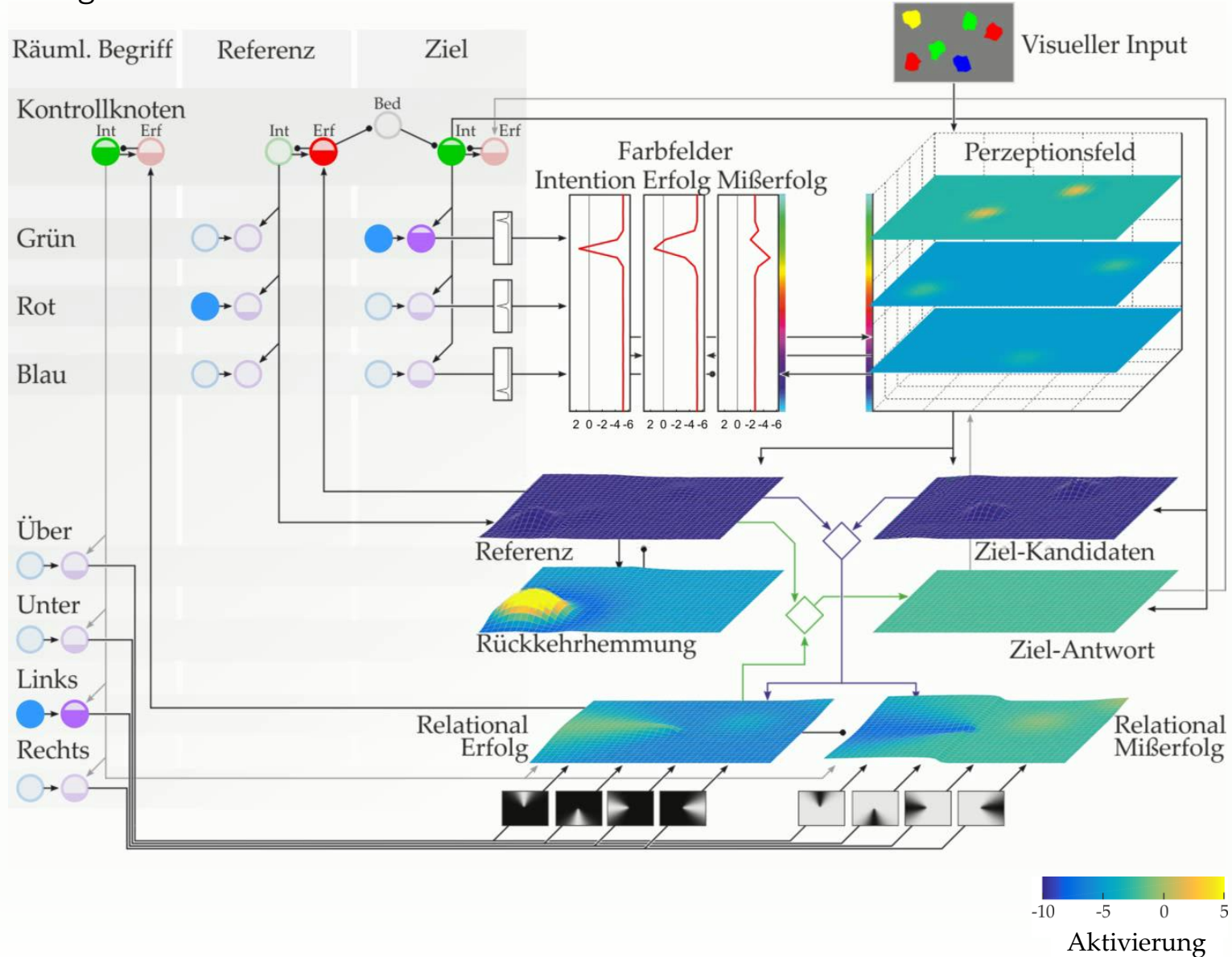
“The green one left of the red one”



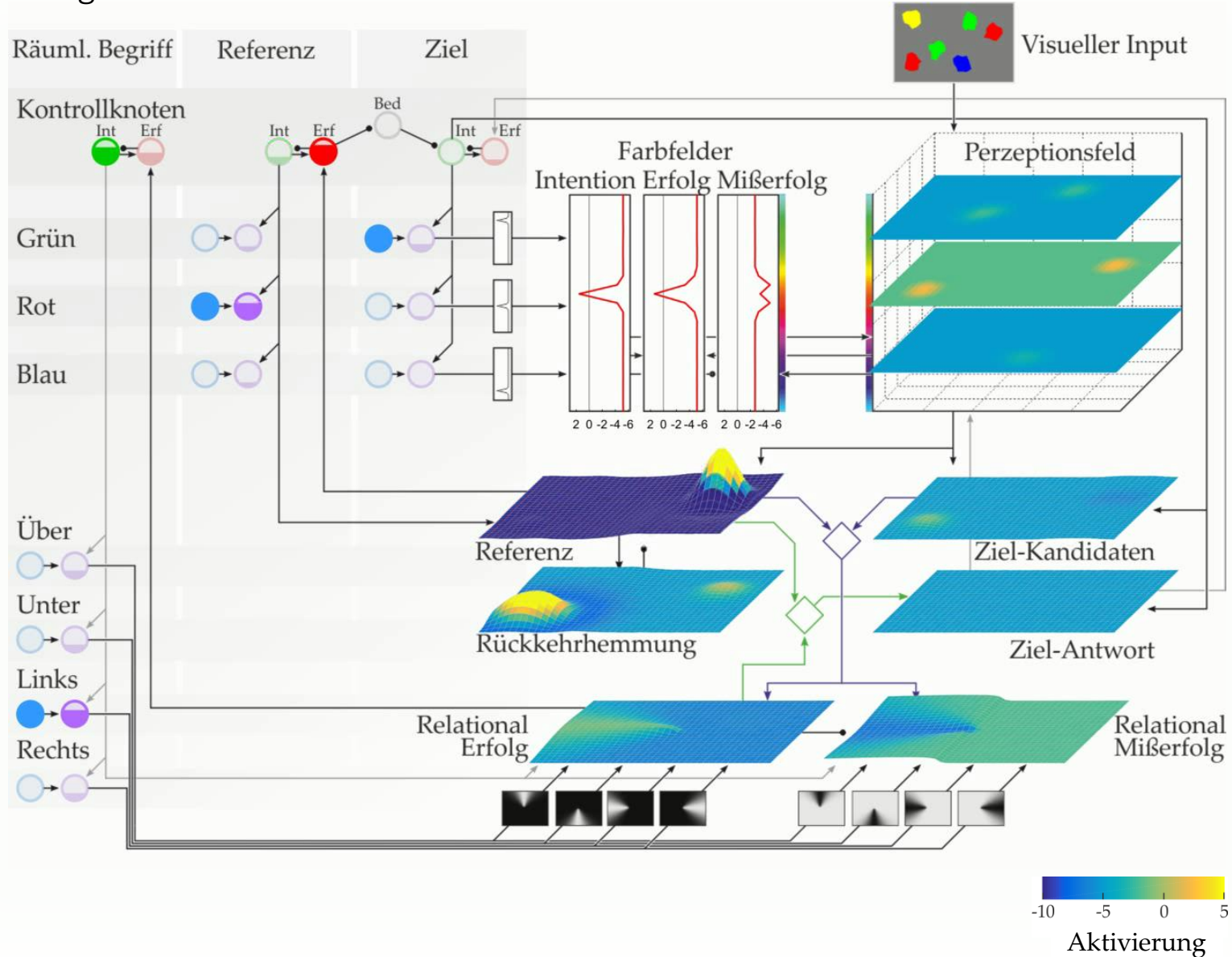
“The green one left of the red one”



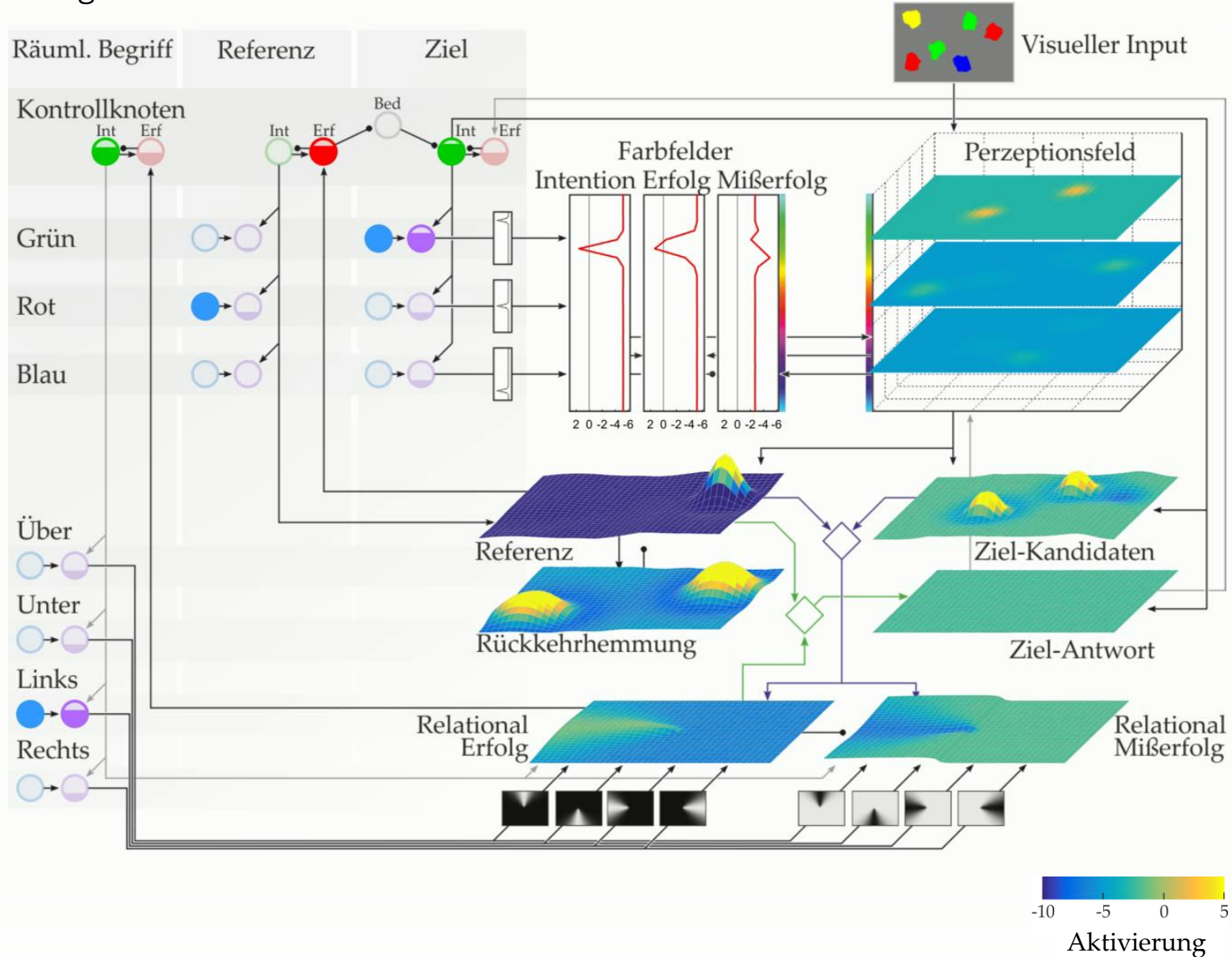
“The green one left of the red one”



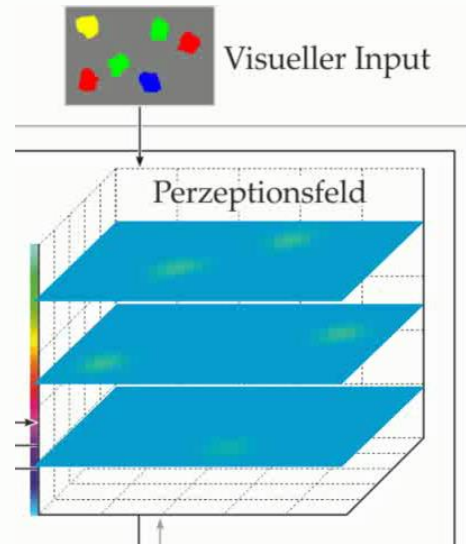
“The green one left of the red one”

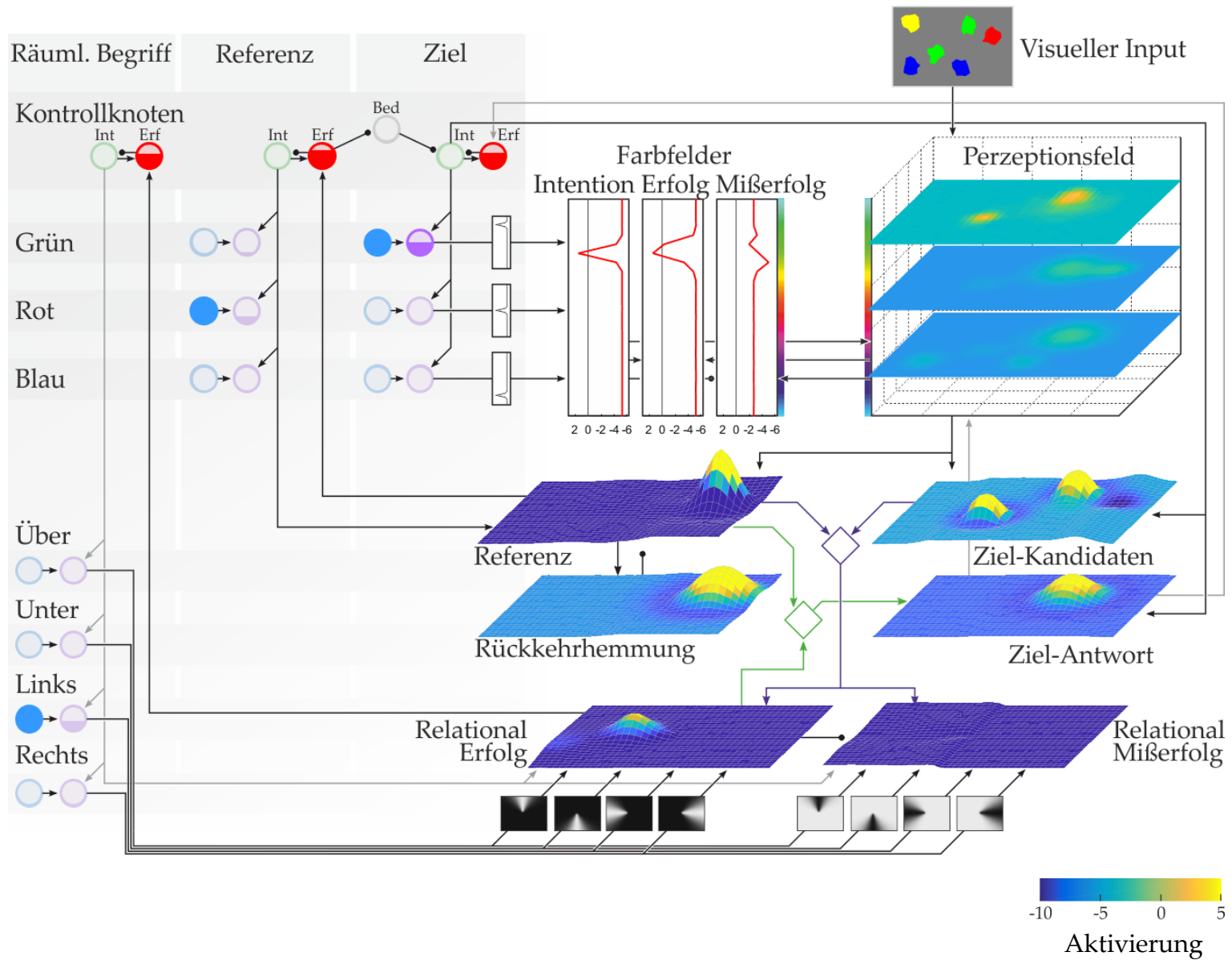


“The green one left of the red one”

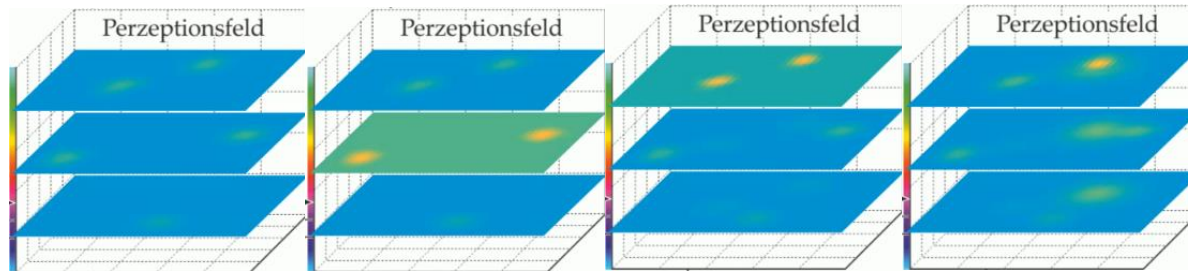


- Activation shifts in visual space
 - Selection of all *potential* reference objects
 - Selection of all *potential* target objects



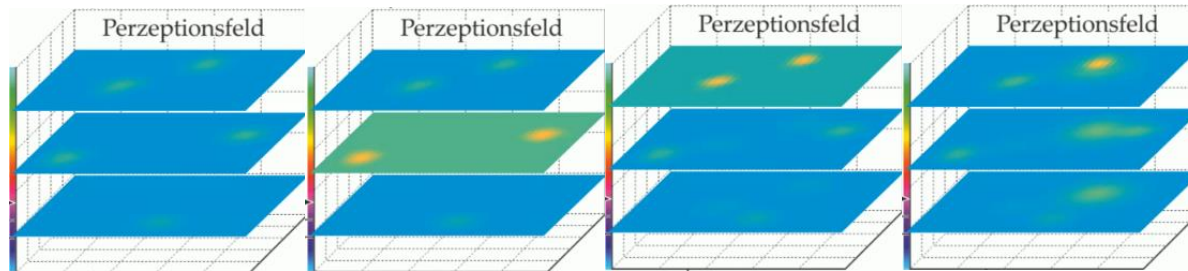


- Activation shifts in visual space
 - Selection of all *potential* reference objects
 - Selection of all *potential* target objects

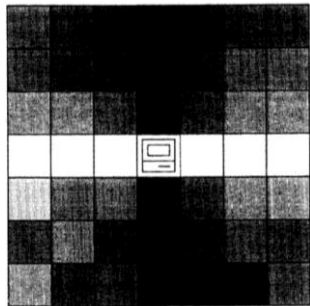


How could such activation shifts be measured?

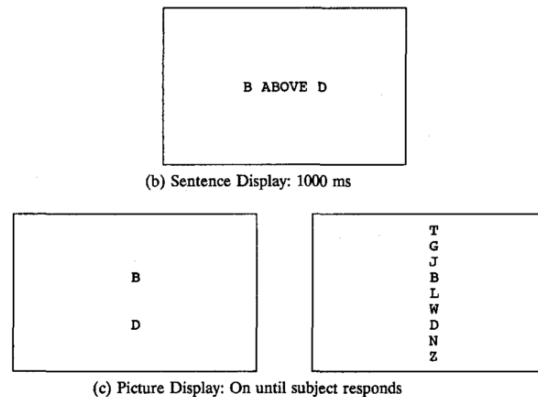
- Activation shifts in visual space
 - Selection of all *potential* reference objects
 - Selection of all *potential* target objects



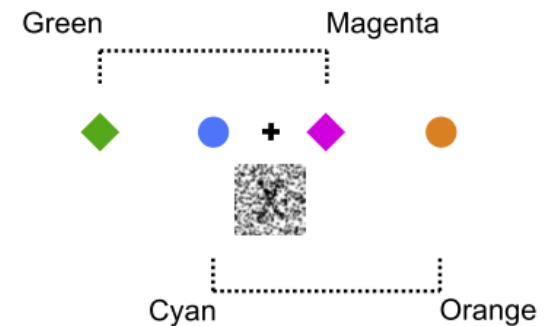
How could such activation shifts be measured?



[Hayward & Tarr, 1995]

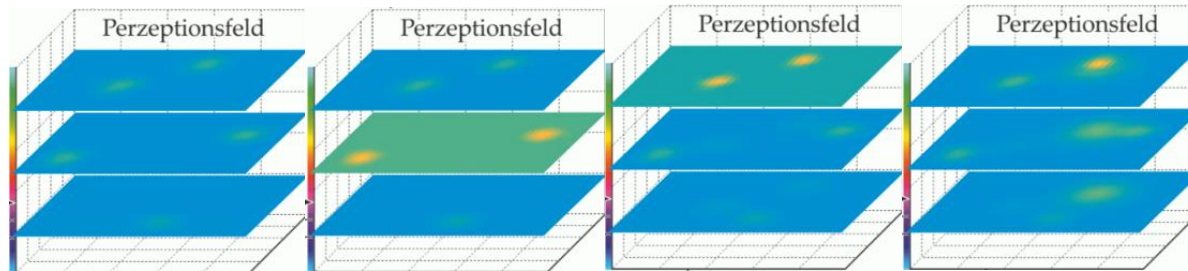


[Logan 1996]

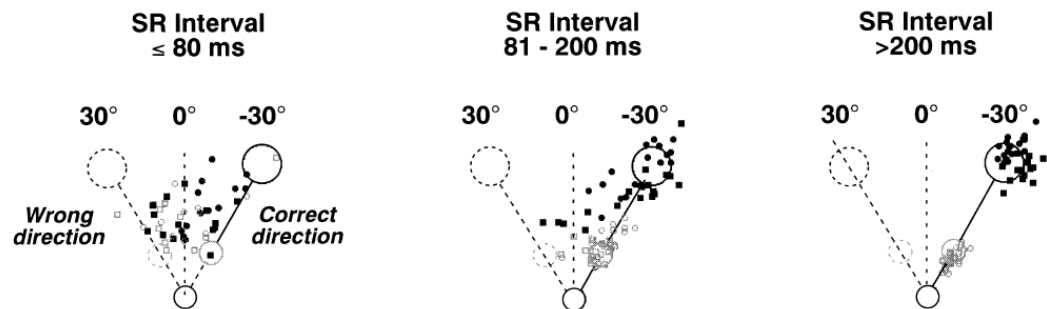


[Franconeri et al. 2012]

- Activation shifts in visual space
 - Selection of all *potential* reference objects
 - Selection of all *potential* target objects

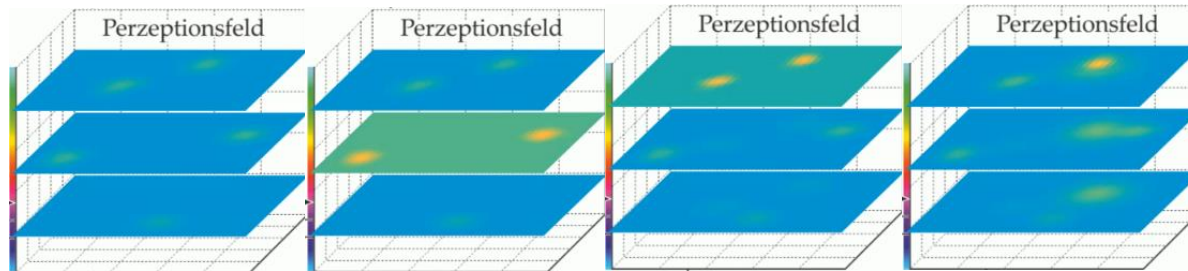


How could such activation shifts be measured?

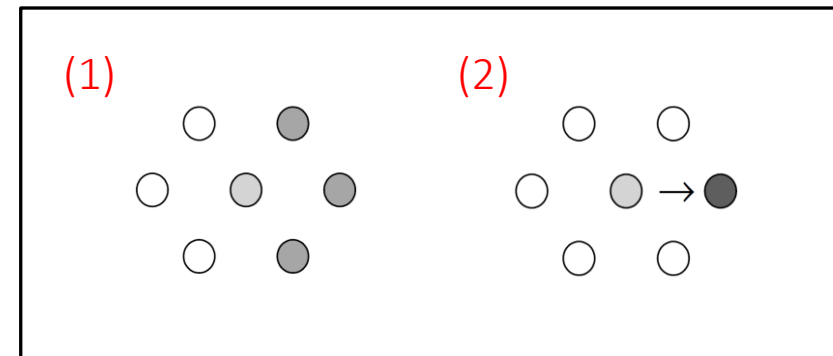
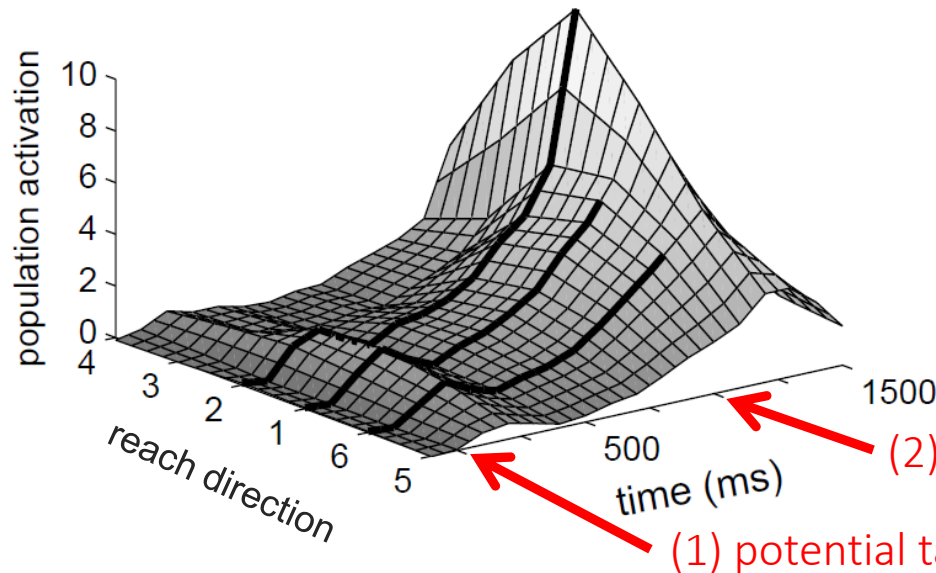


[Ghez et al., 1997]

- Aktivierung an wechselnden Orten im visuellen Raum.
 - alle *potenziellen* Referenzobjekte
 - alle *potenziellen* Zielobjekte



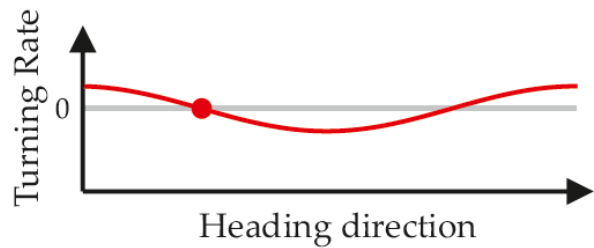
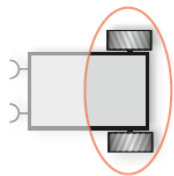
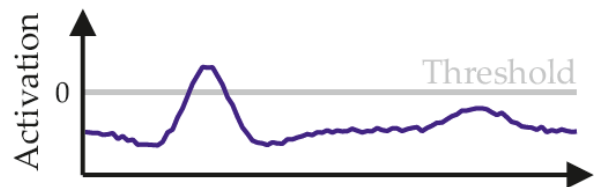
How could such activation shifts be measured?



(2) Definitive Target marked / GO

(1) potential targets shown

[Bastian et al., 2003]



Sensory surface

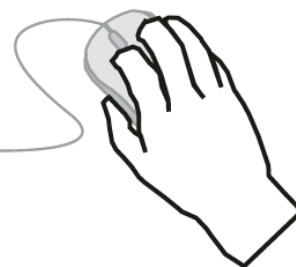
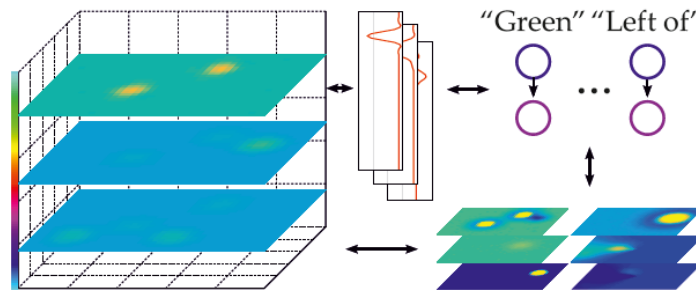
Sensorimotor fields

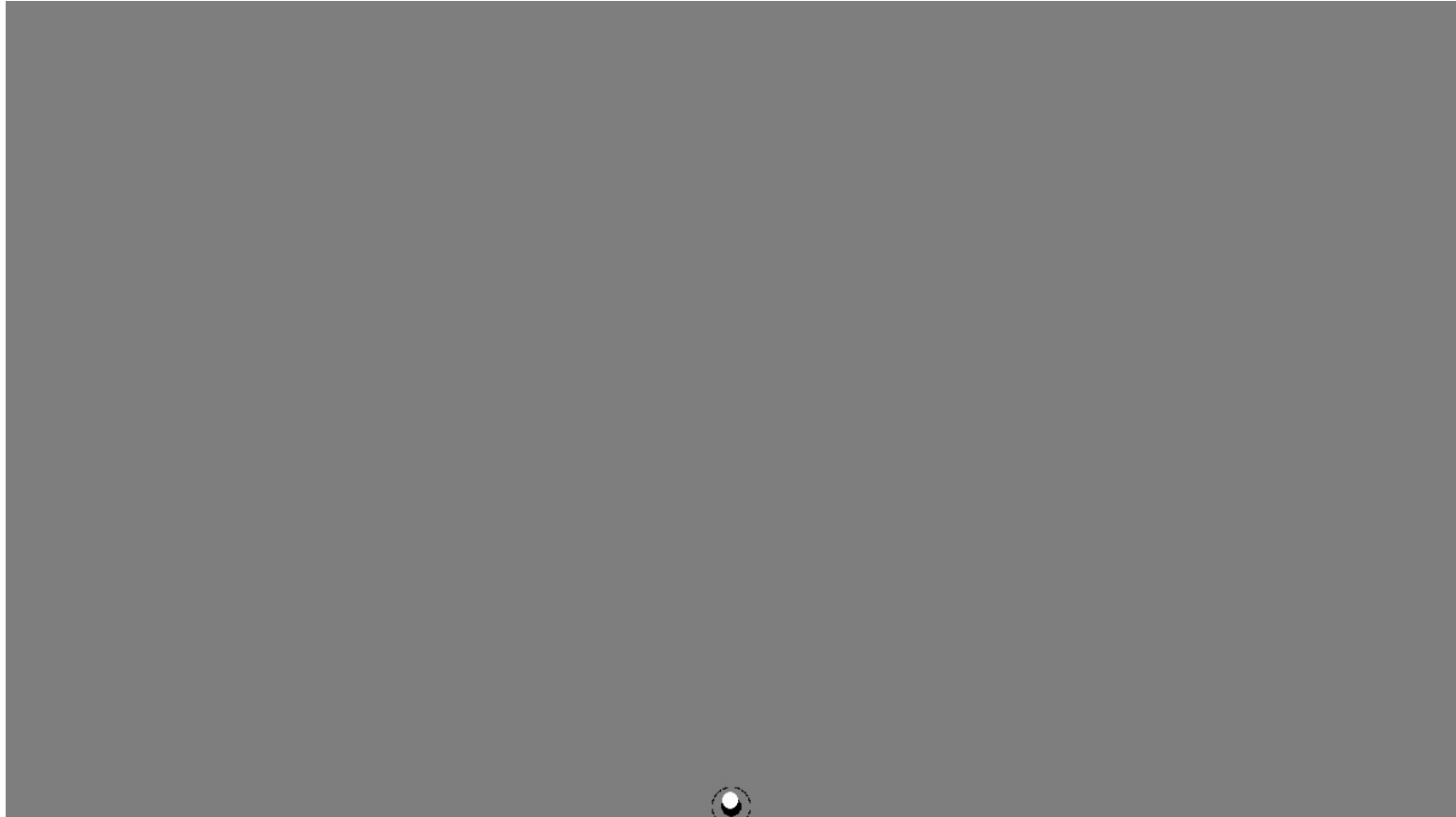
Motor surface

Behavior



"The green item to the left of the red one."





1-2 s	Spatial phrase
1 s	Movement start
Upon onset	Visual scene
2 s	Select target object

Das Grüne unter dem Weißen.



1-2 s

Spatial phrase

1 s

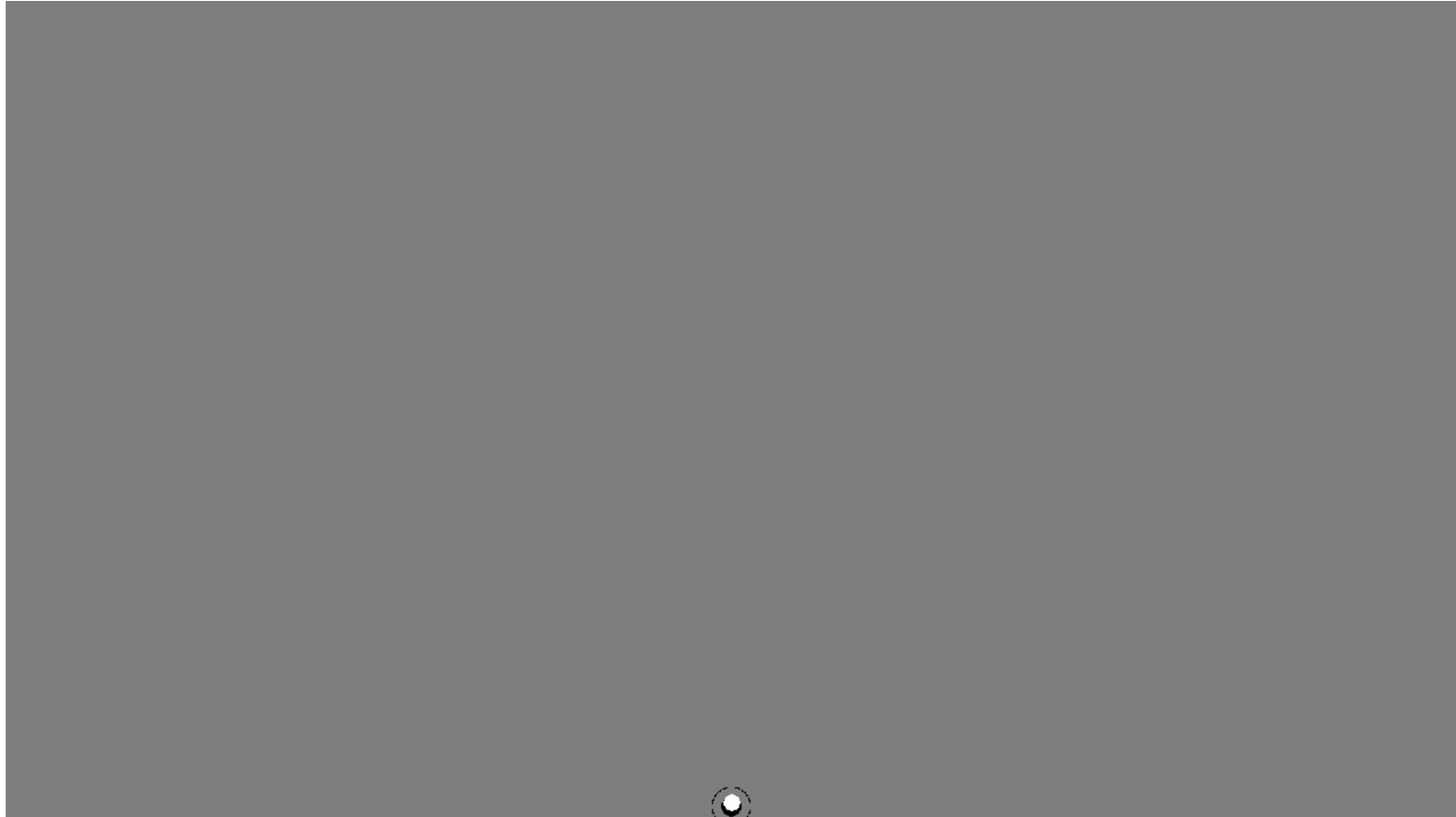
Movement start

Upon onset

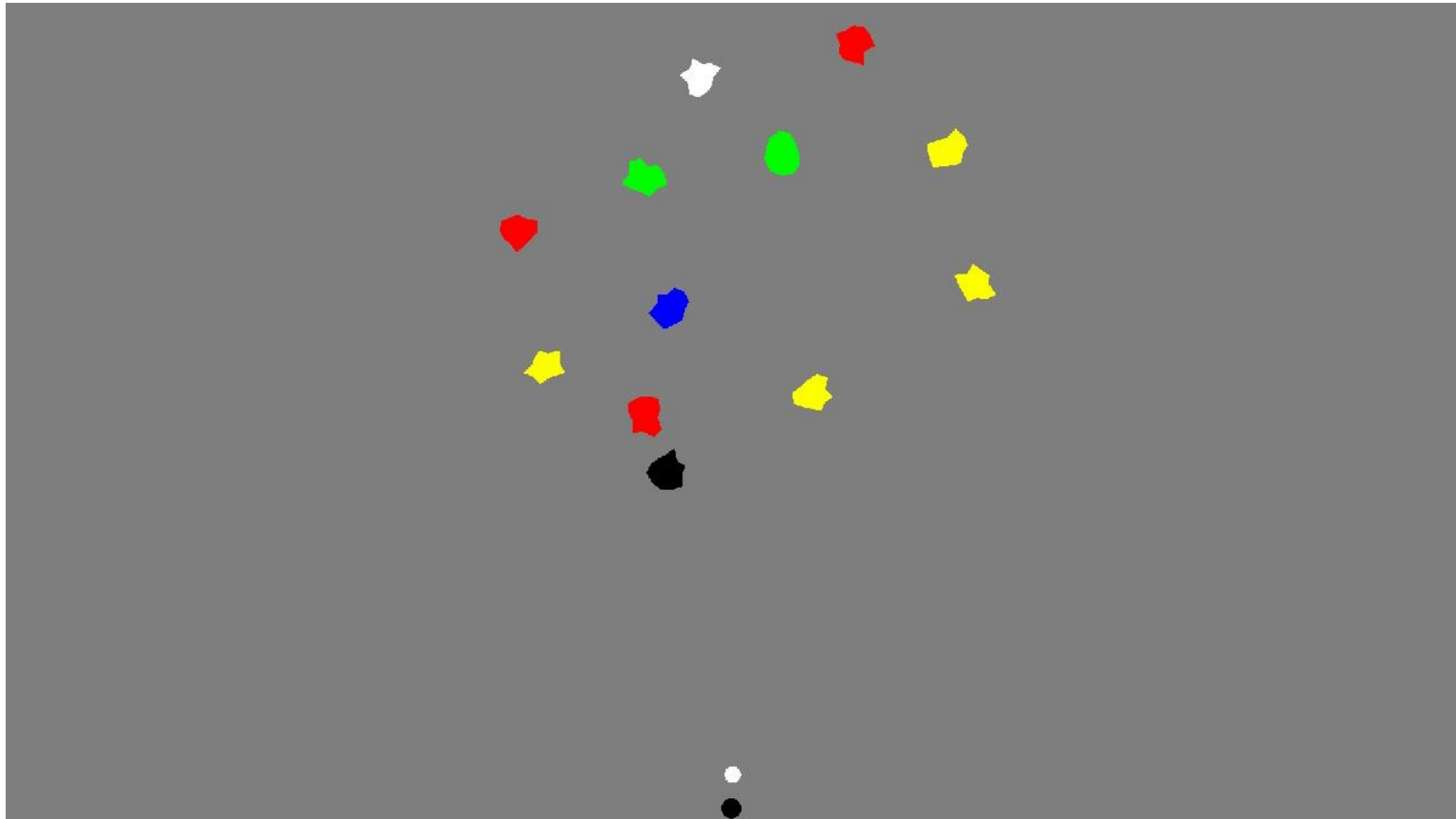
Visual scene

2 s

Select target object

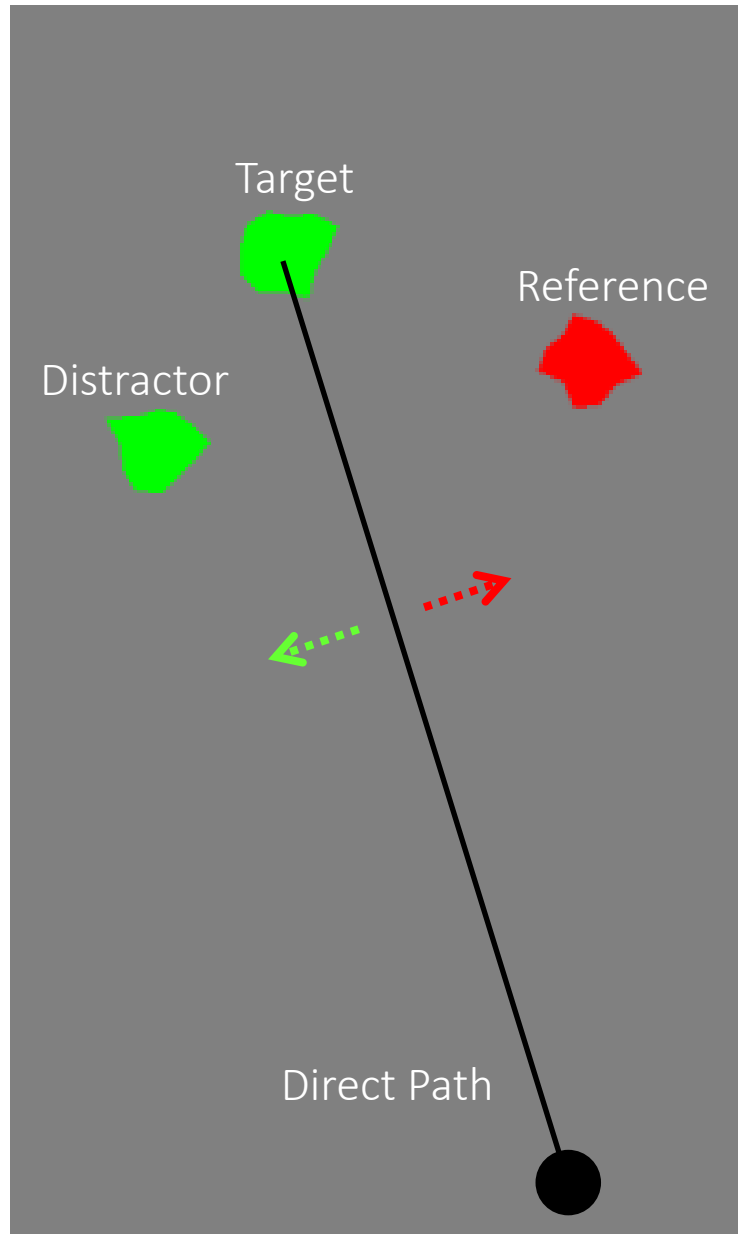


1-2 s	Spatial phrase
1 s	Movement start
Upon onset	Visual scene
2 s	Select target object



1-2 s	Spatial phrase
1 s	Movement start
Upon onset	Visual scene
2 s	Select target object

“The green one left of the red one”

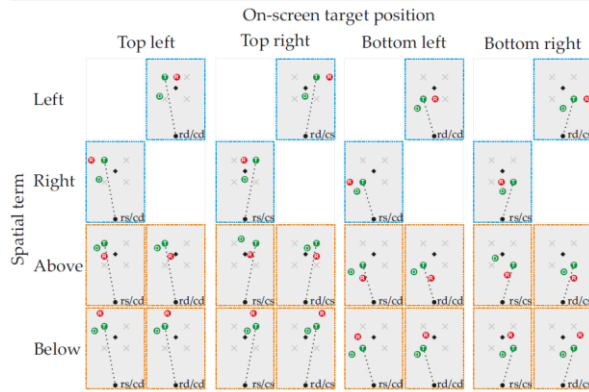


Hypotheses

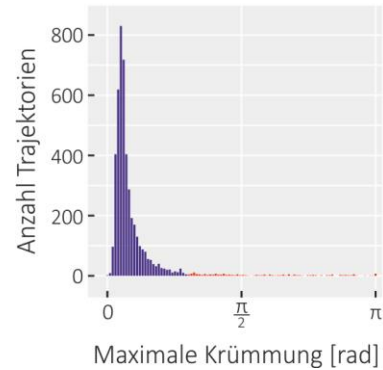
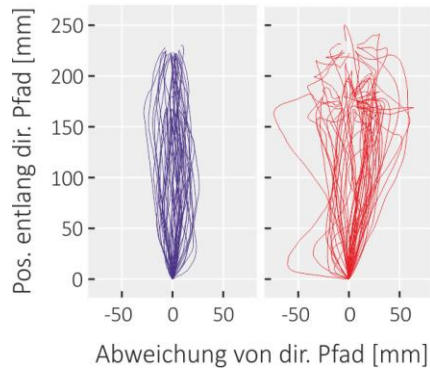
.....➤ Attraction to distractor object

.....➤ Attraction to reference object

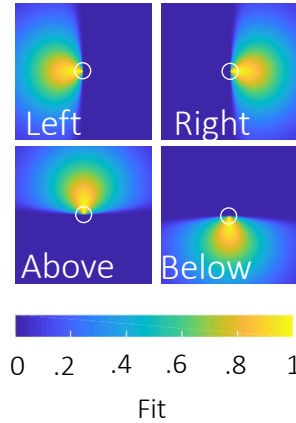
Distractor left



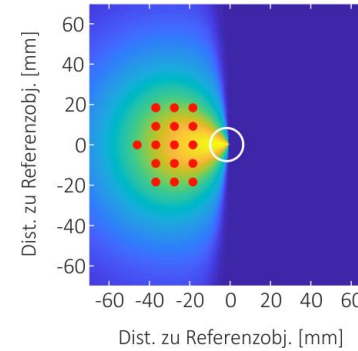
Distractor right



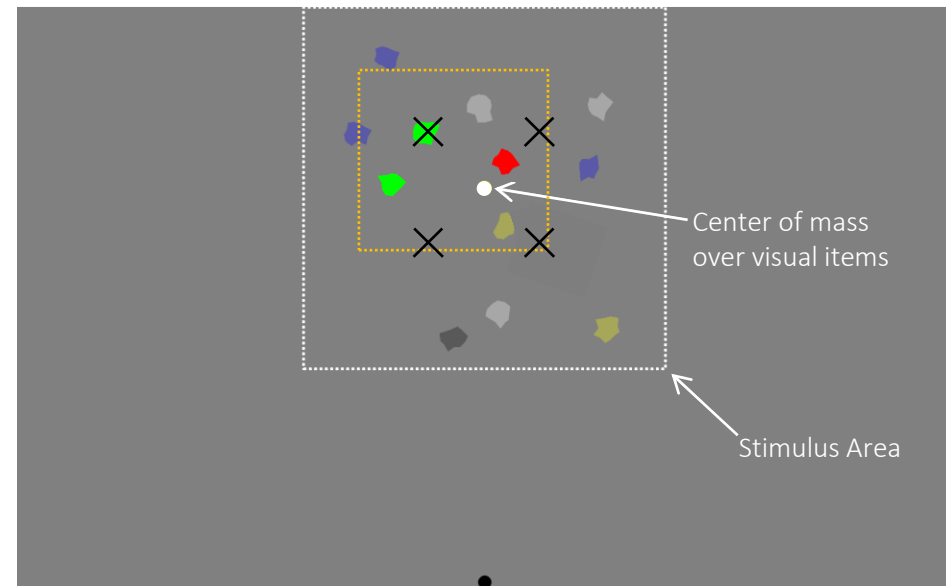
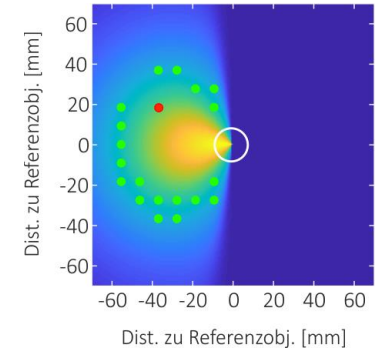
Spatial templates

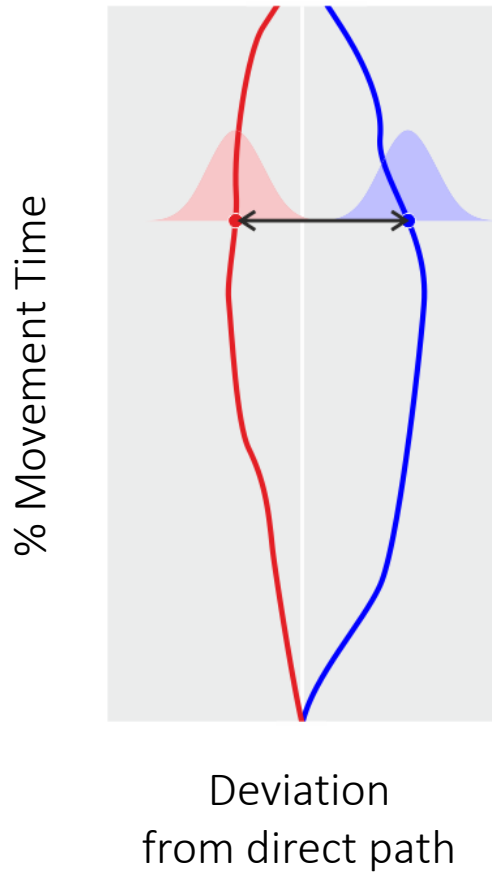


Target Positions



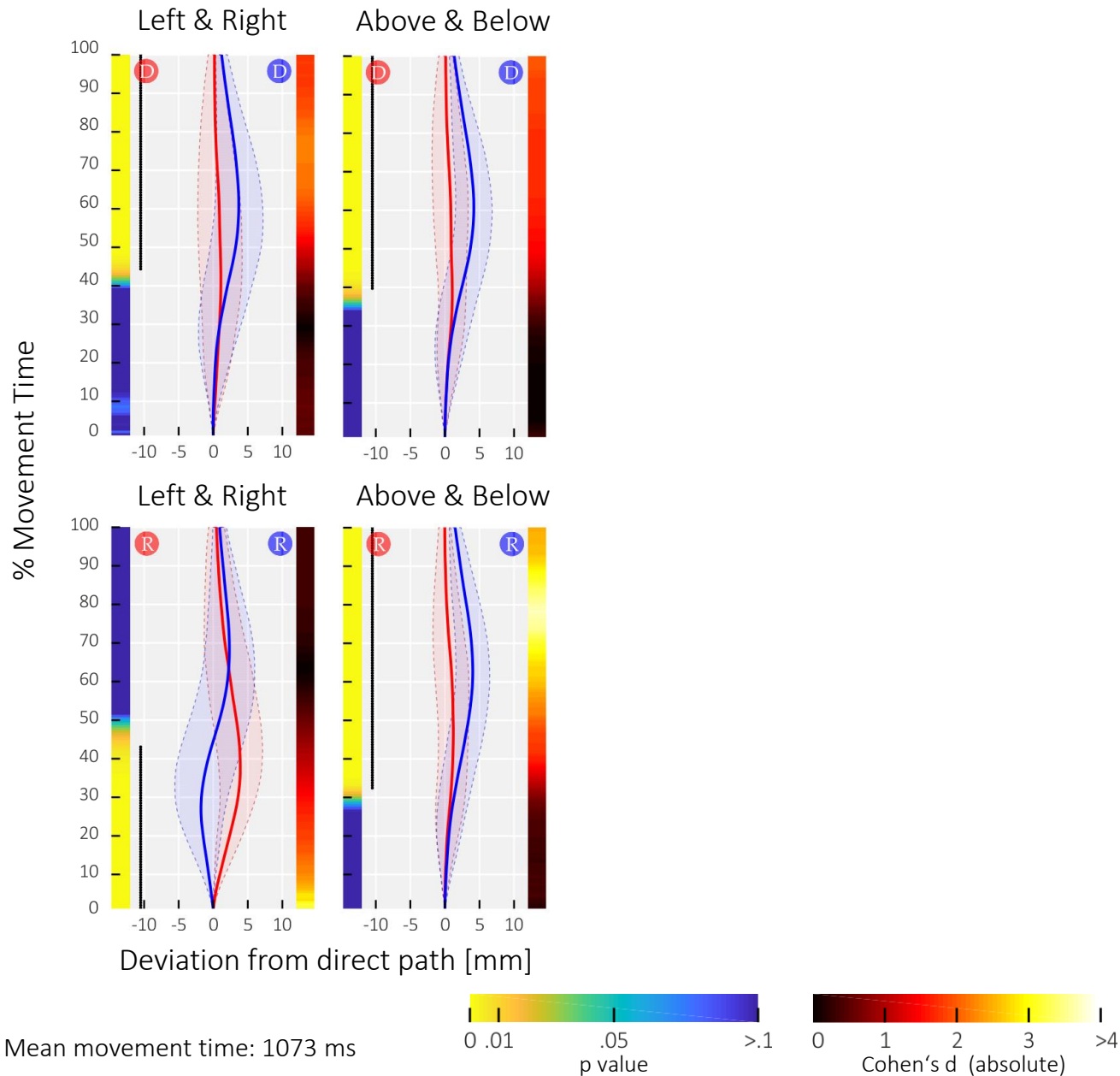
Distractor Positions



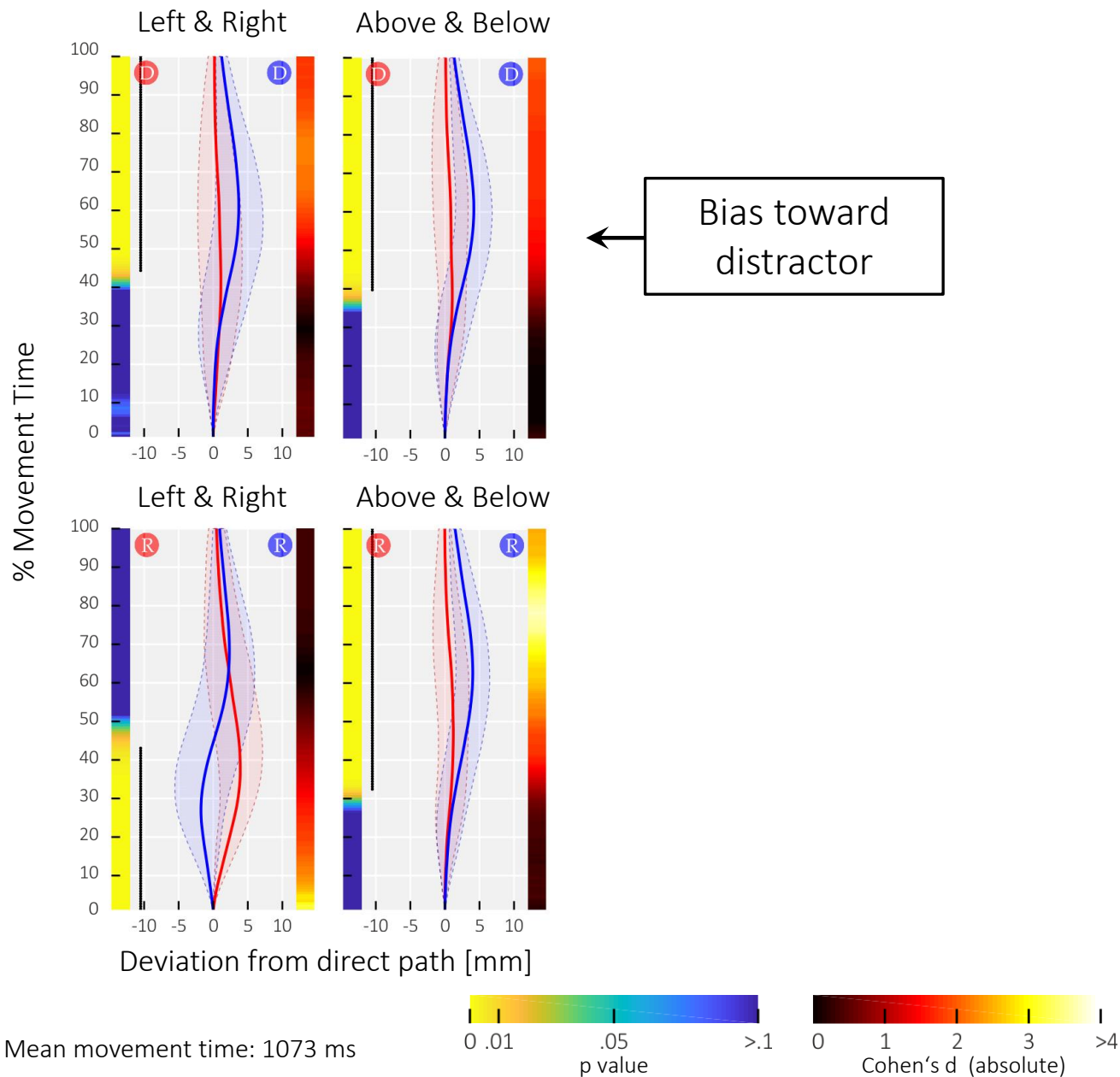


- 151 interpolated time steps
- Paired sample t-tests ($p < .01$)
- Overall significance: Bootstrap [Dale, 2007]

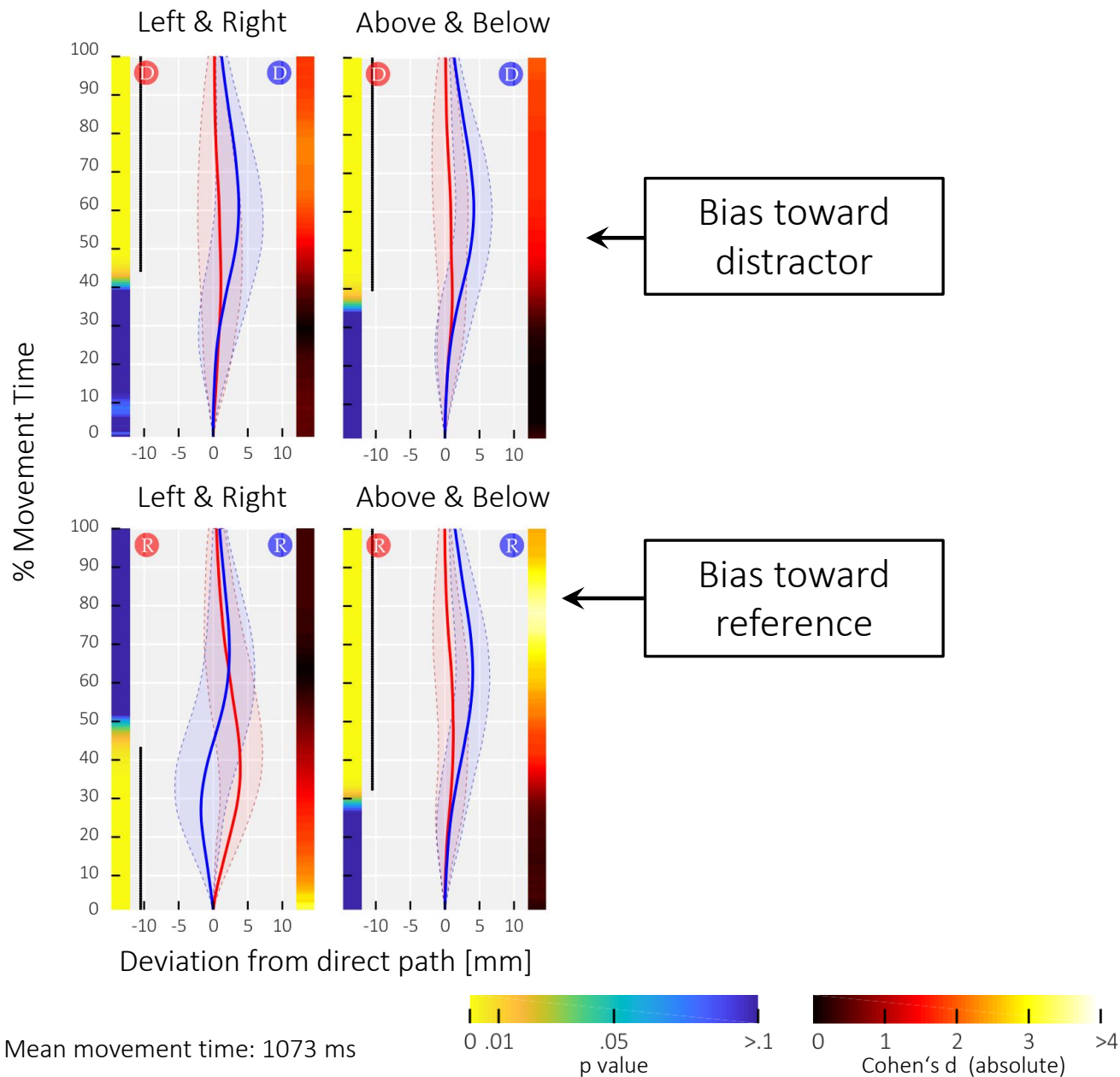
Experiment 1 : Movement bottom → top



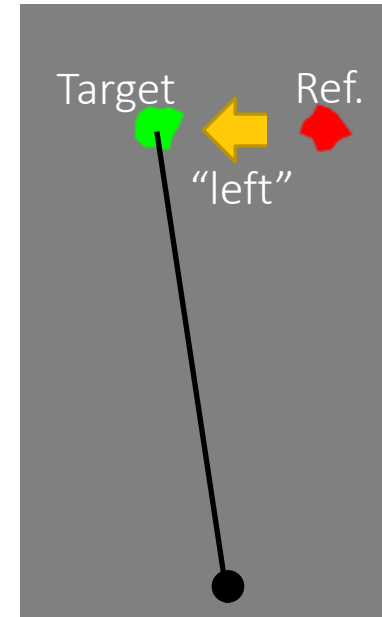
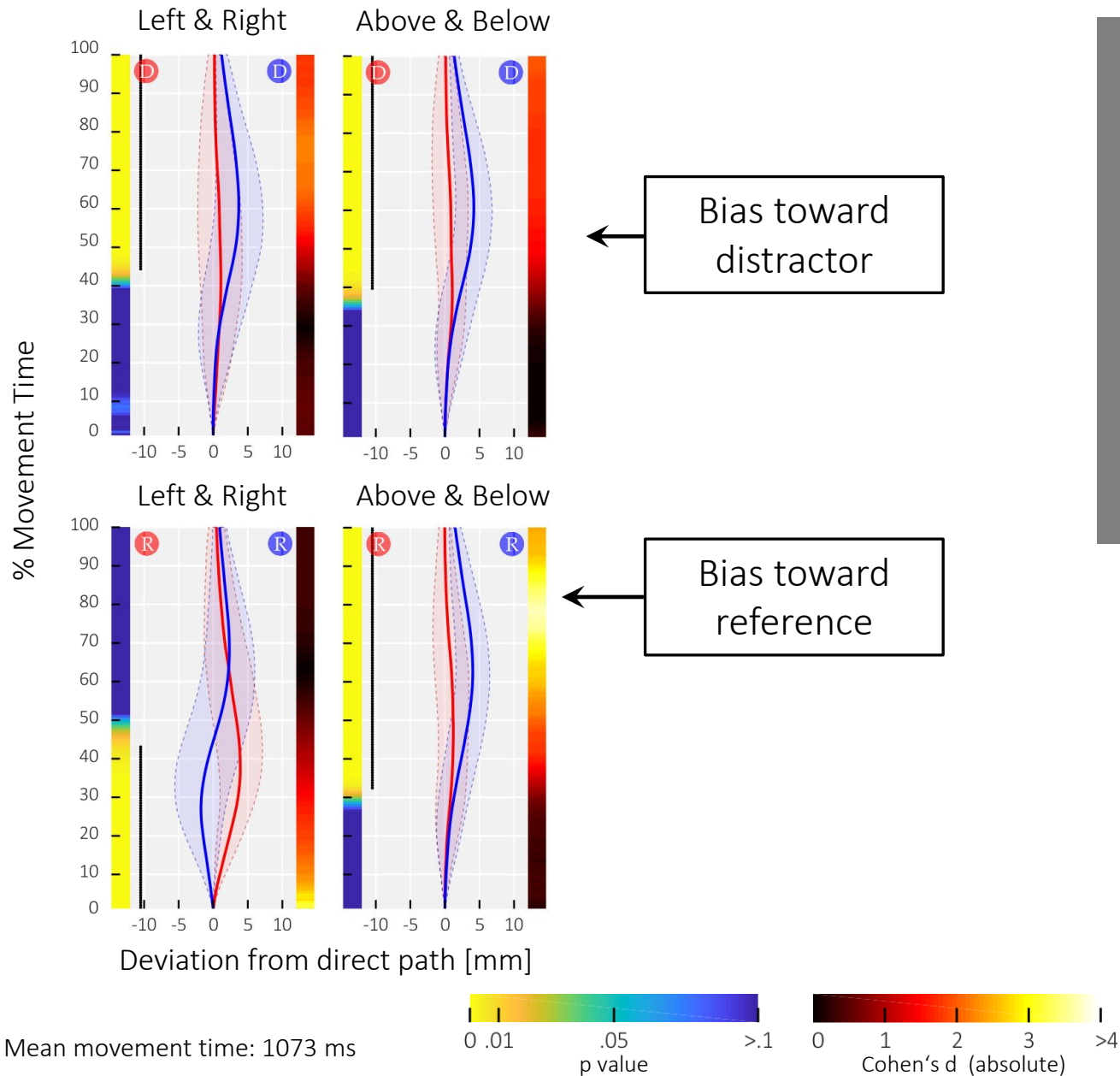
Experiment 1 : Movement bottom → top



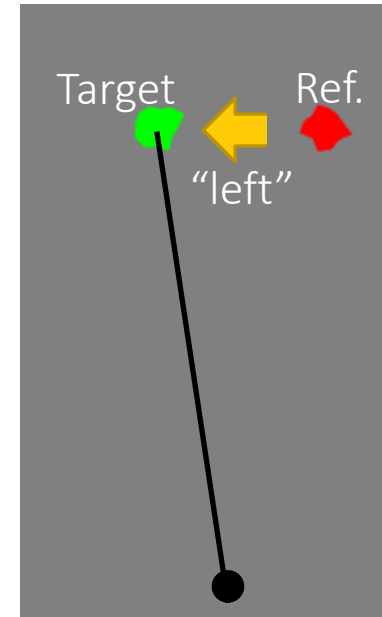
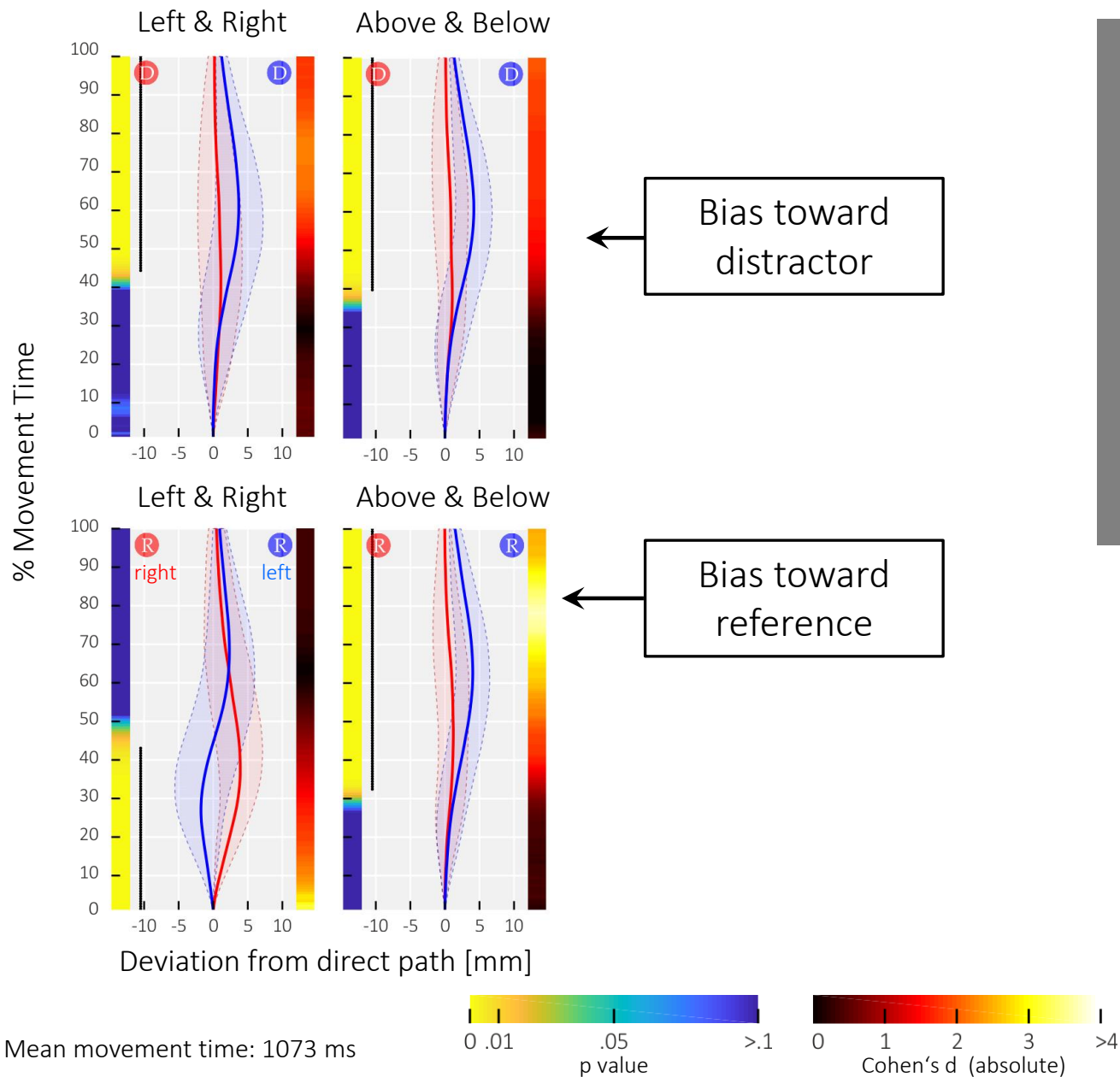
Experiment 1 : Movement bottom → top



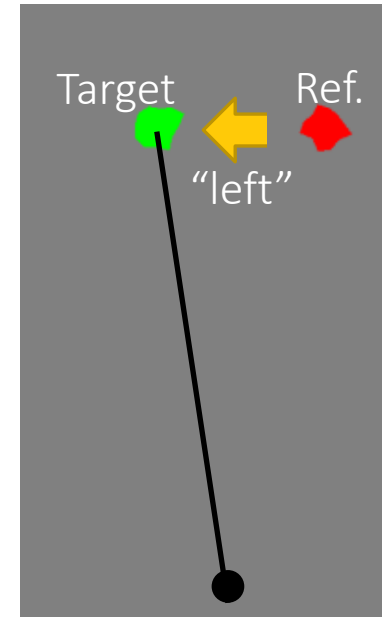
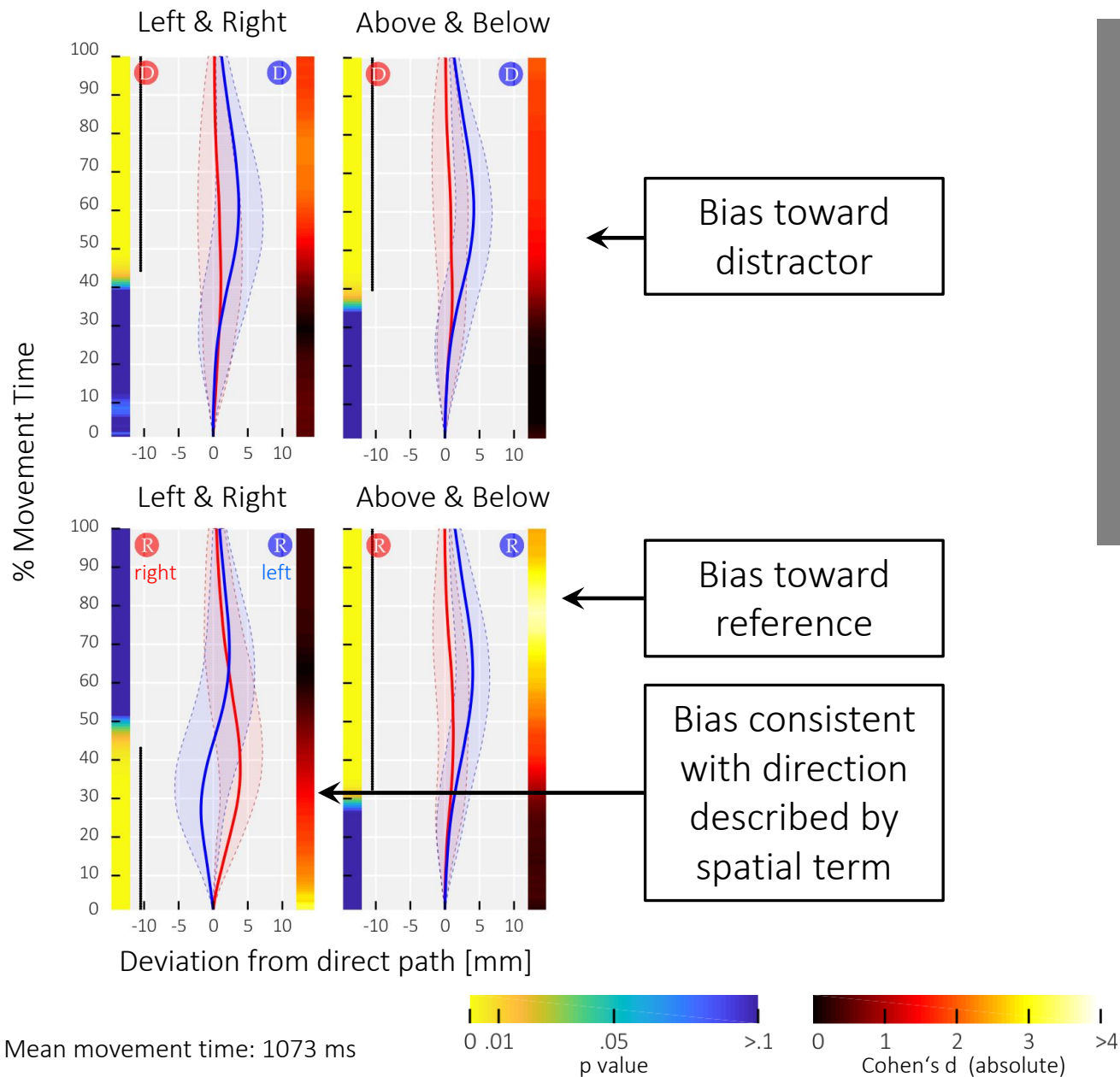
Experiment 1 : Movement bottom → top



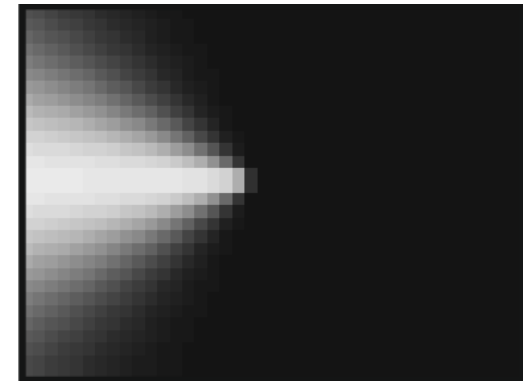
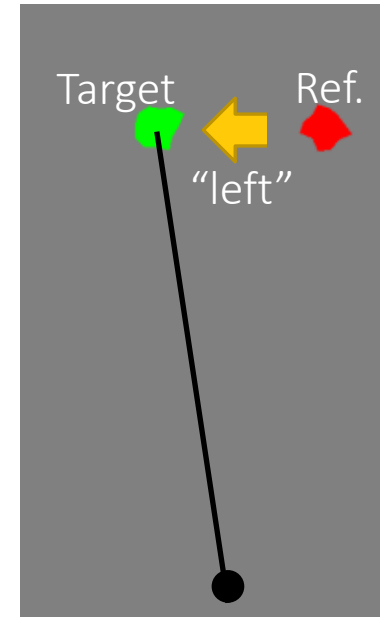
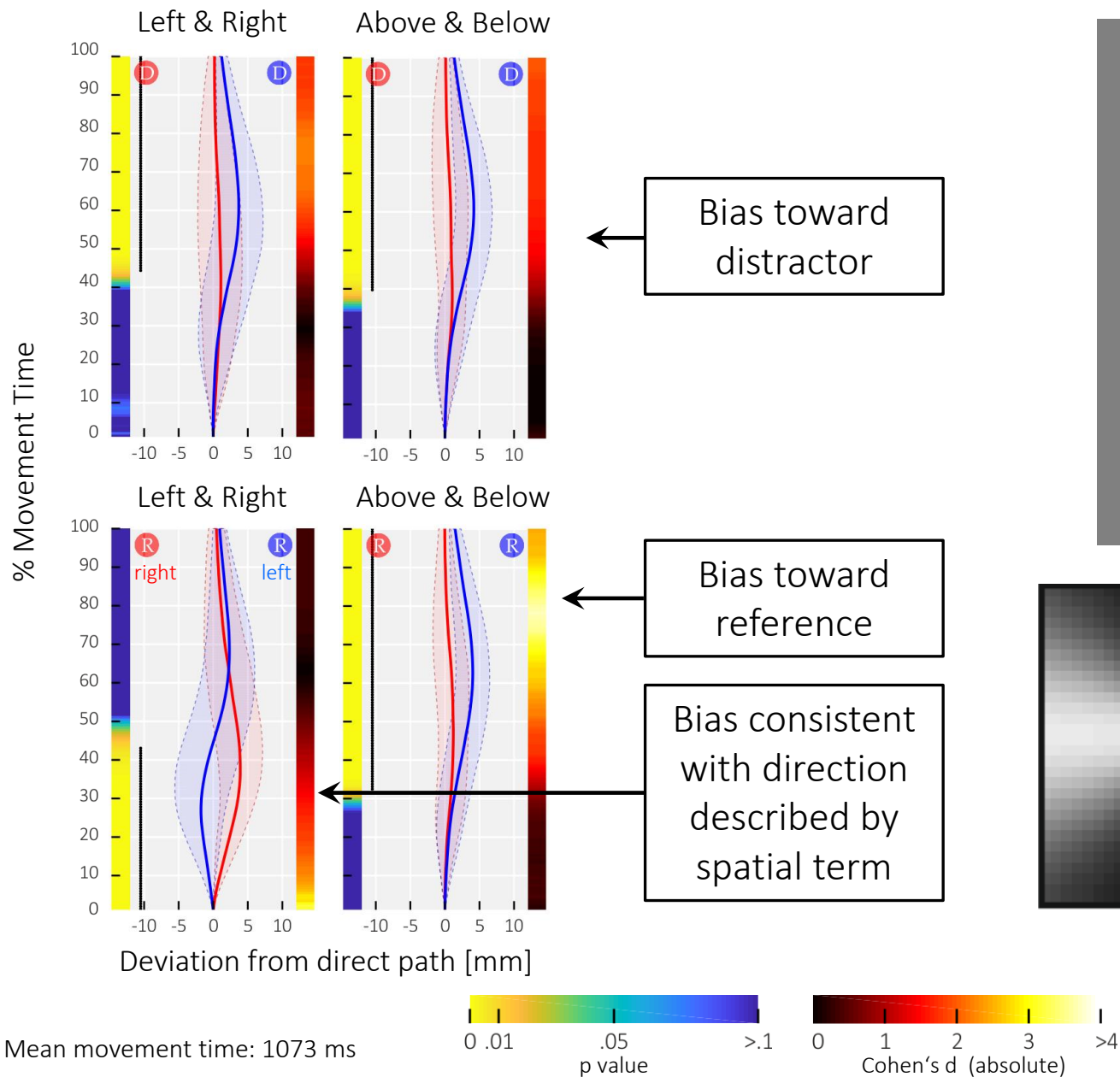
Experiment 1 : Movement bottom → top



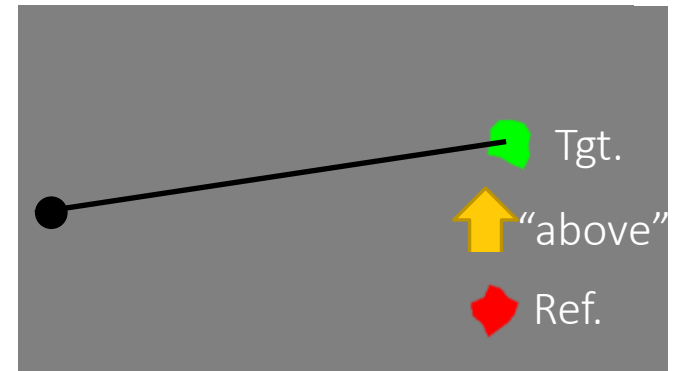
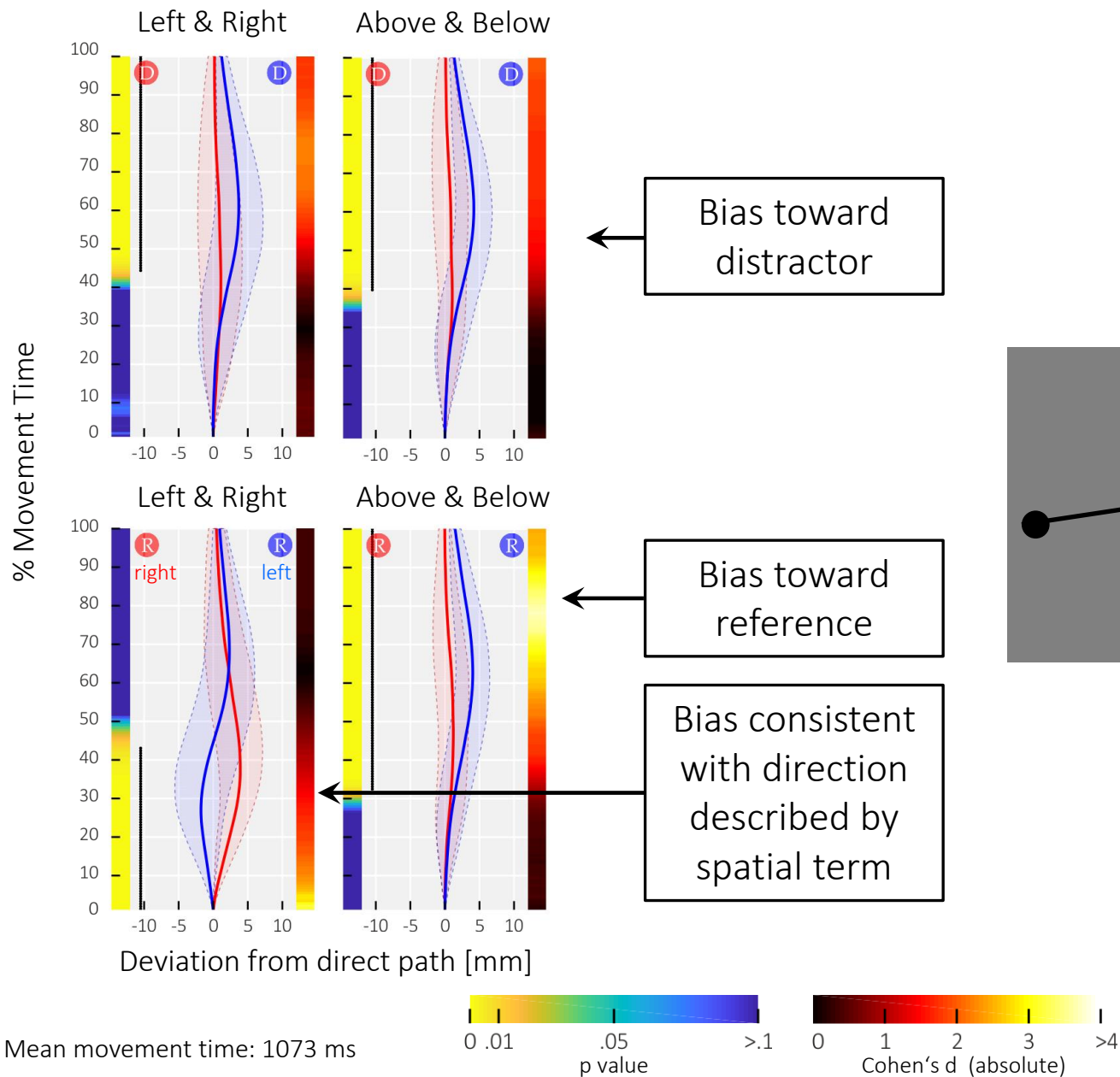
Experiment 1 : Movement bottom → top



Experiment 1 : Movement bottom → top

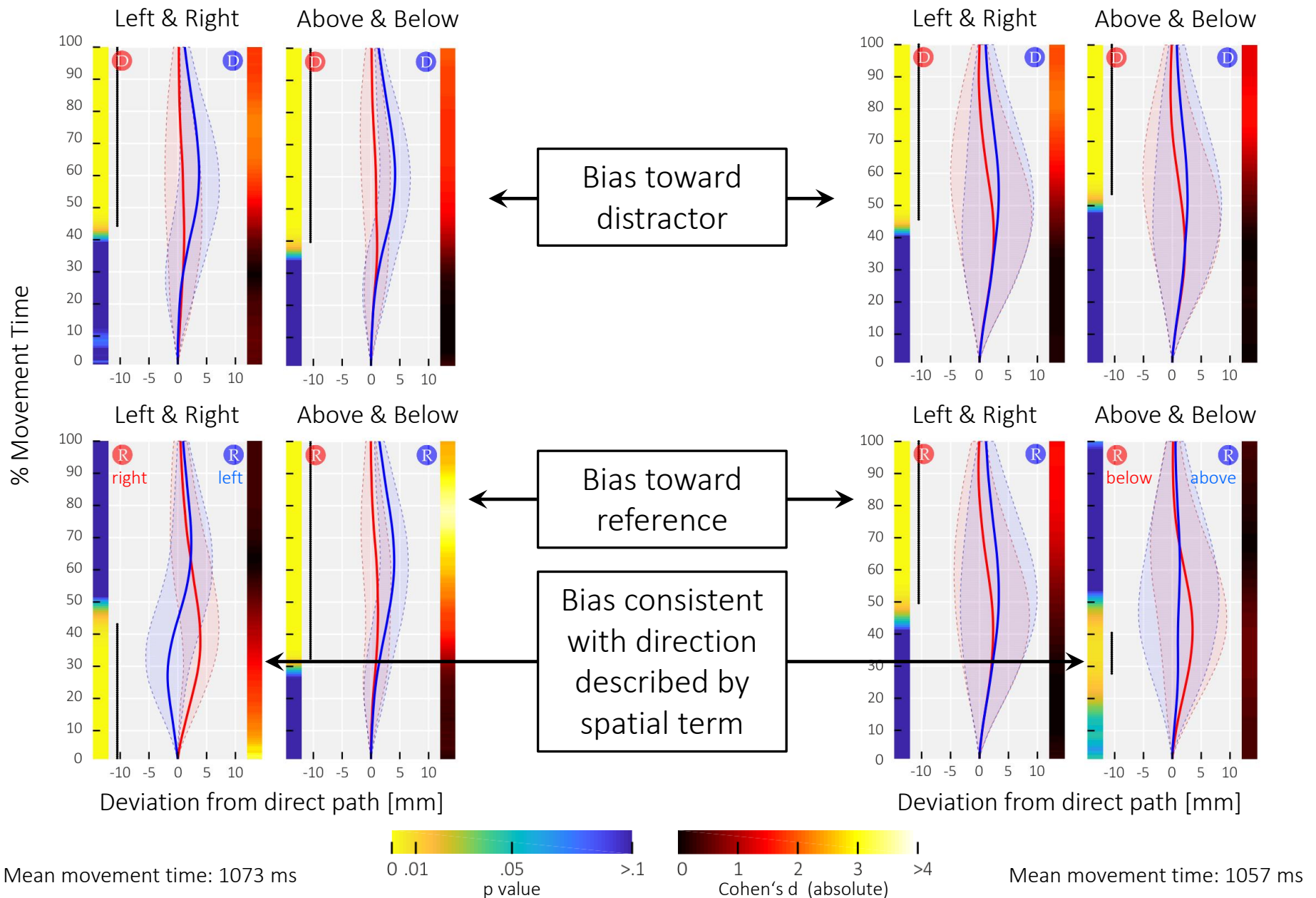


Experiment 1 : Movement bottom → top



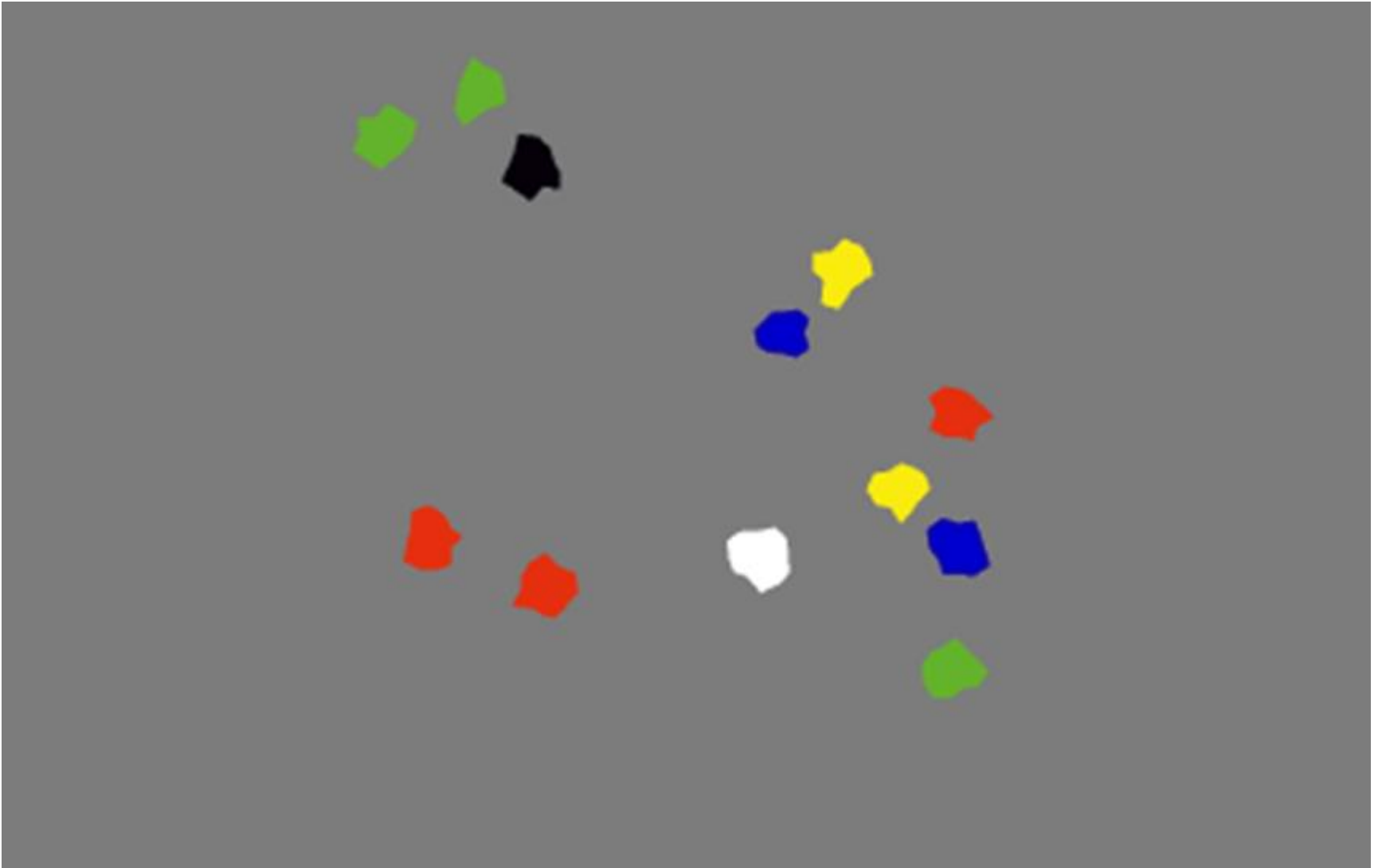
Experiment 1 : Movement bottom → top

Experiment 2 : Movement left → right



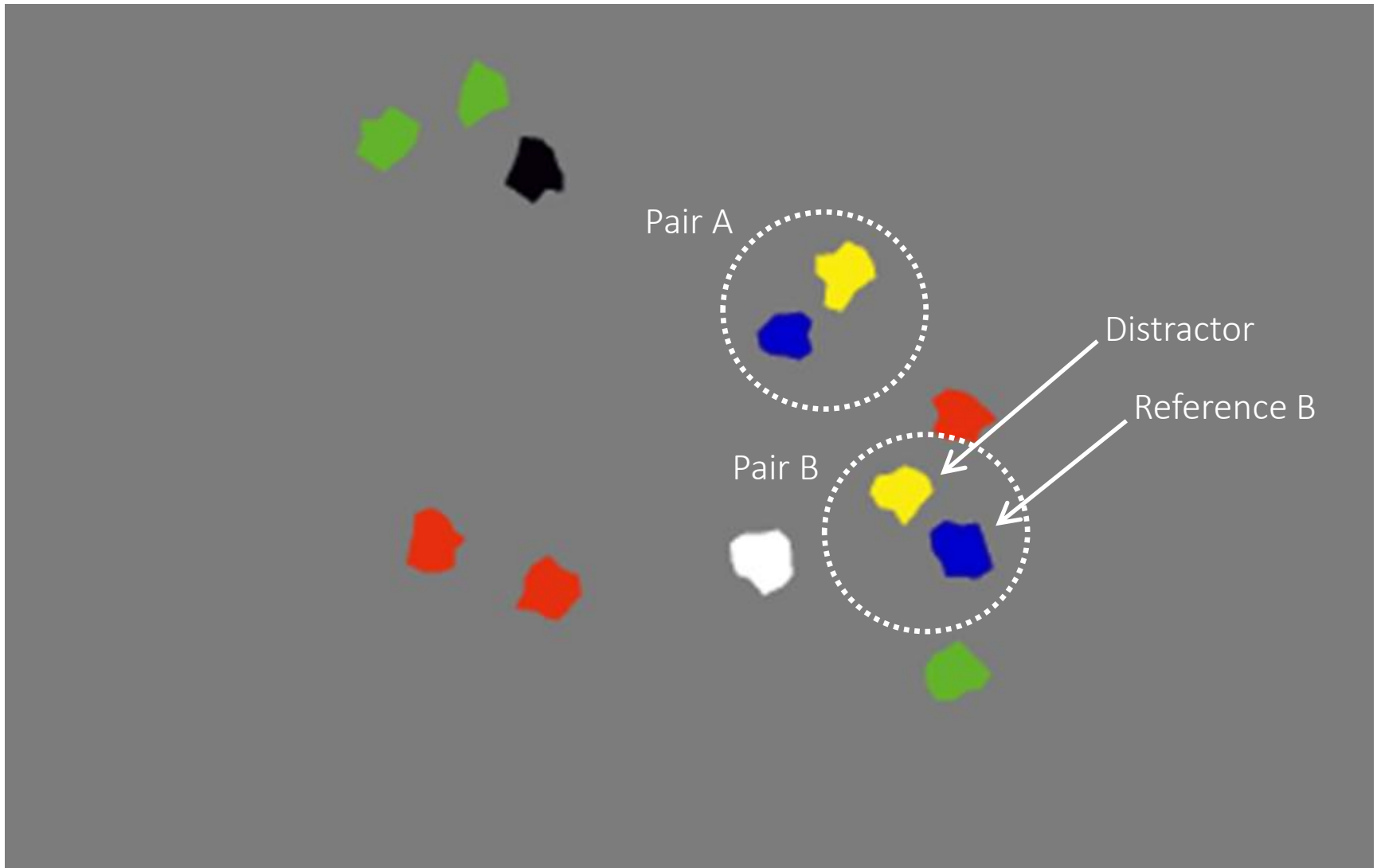
Do the reference and distractor effects result from *relational grounding*?

“The yellow one right of the blue one”



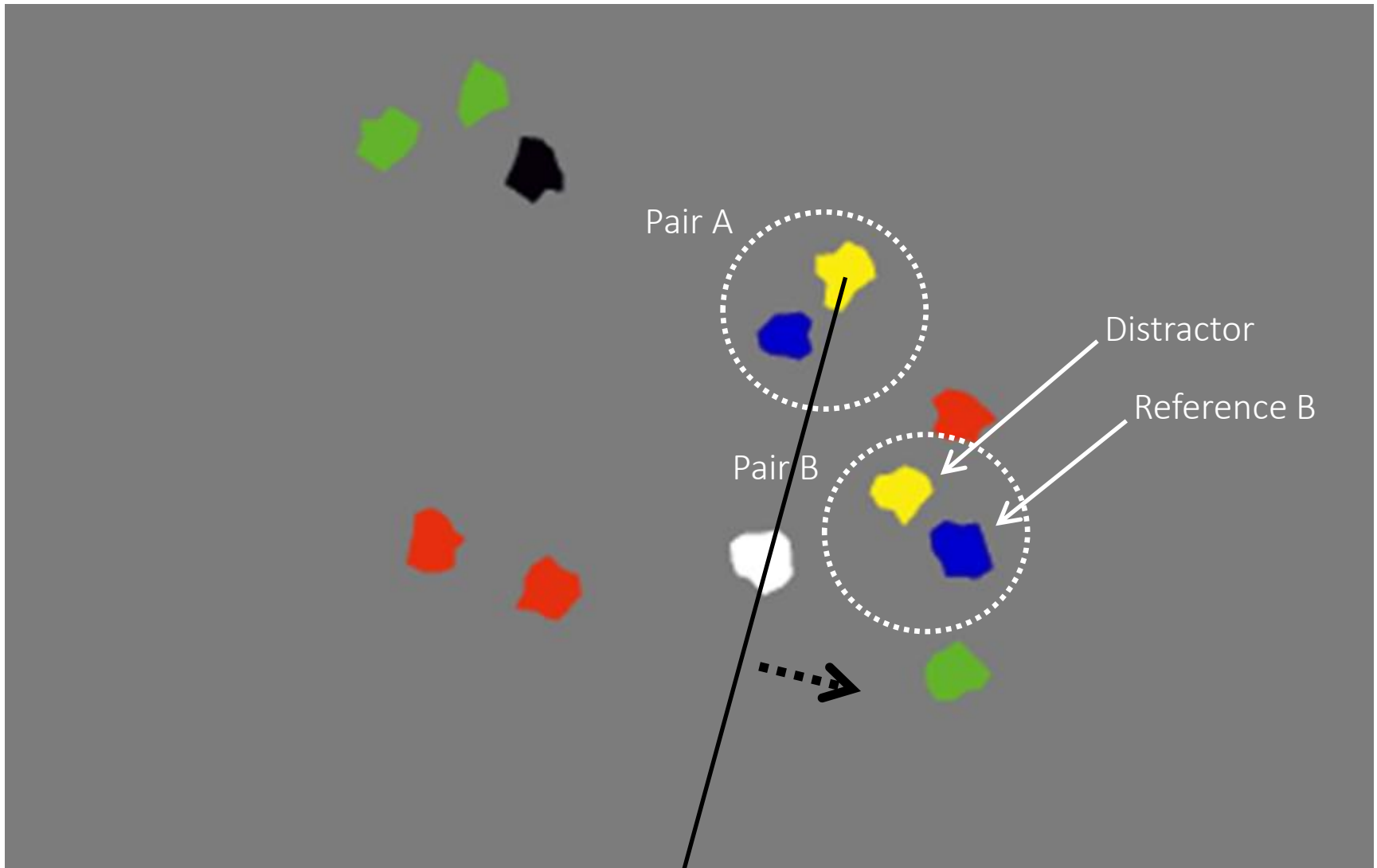
Do the reference and distractor effects result from *relational grounding*?

“The yellow one right of the blue one”



Do the reference and distractor effects result from *relational grounding*?

“The yellow one right of the blue one”



Reference B presence

✓

✗

Pair B

Distractor only

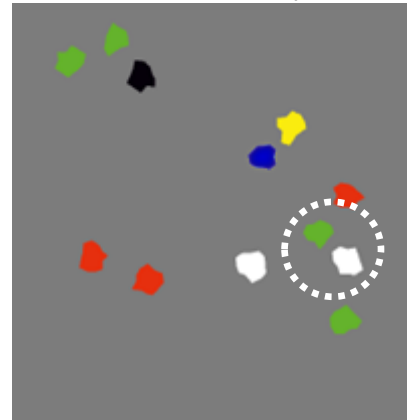
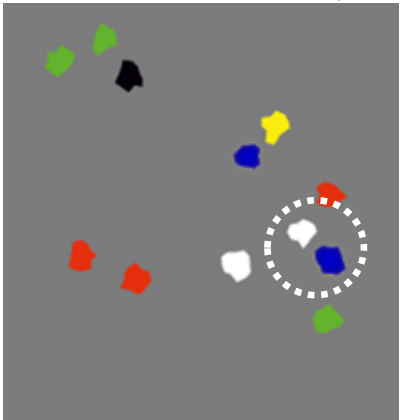


Distractor
presence

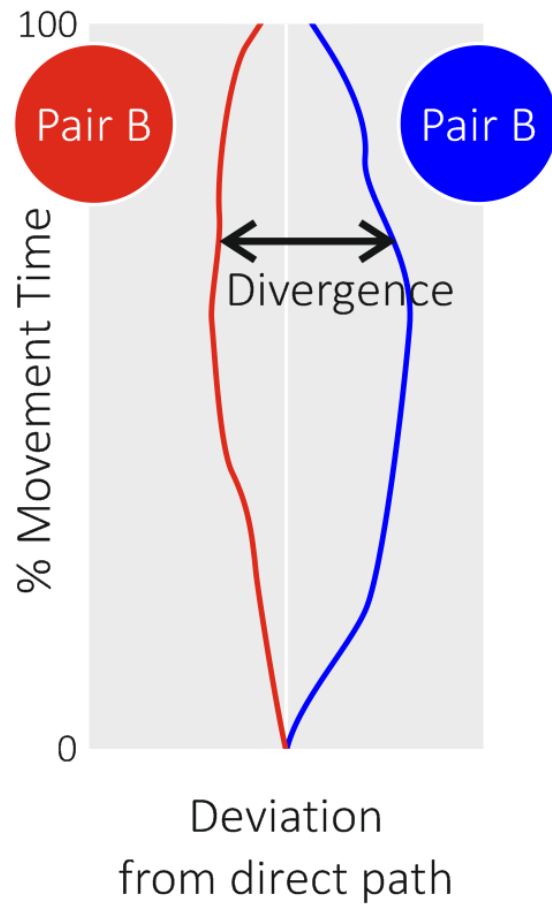
✓

Reference B only

Pair A only

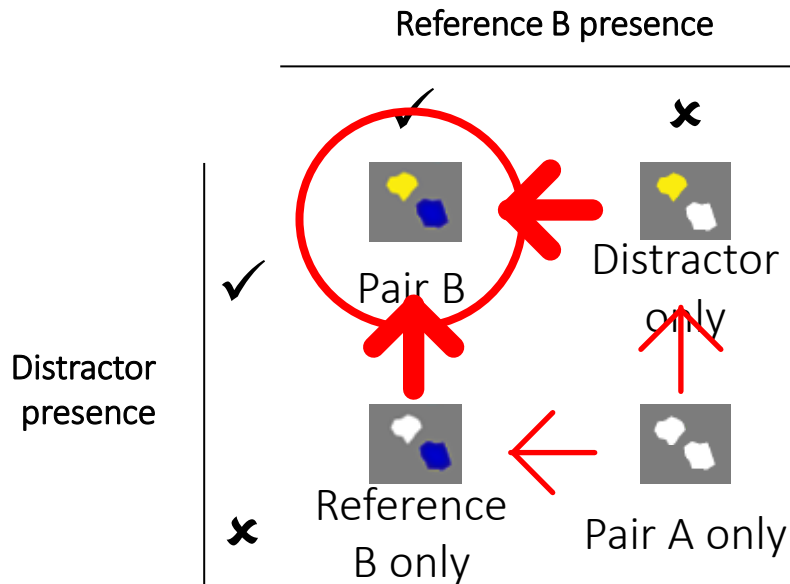
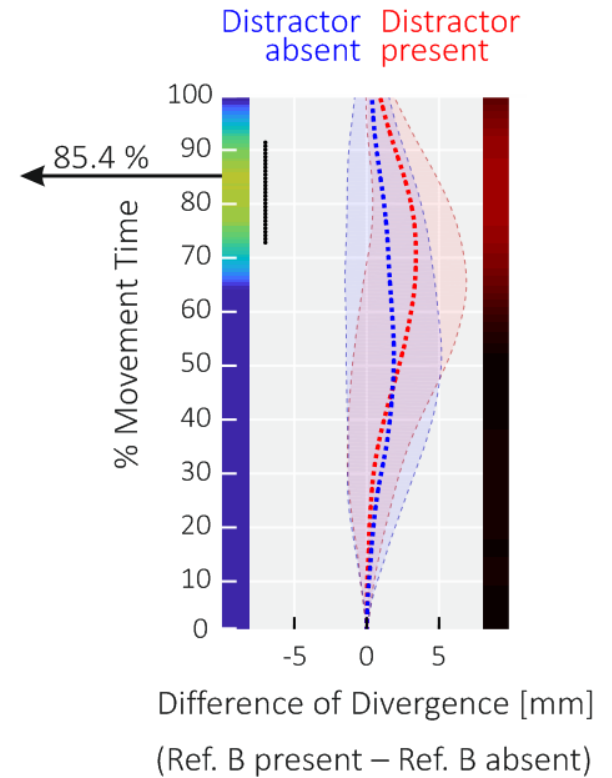
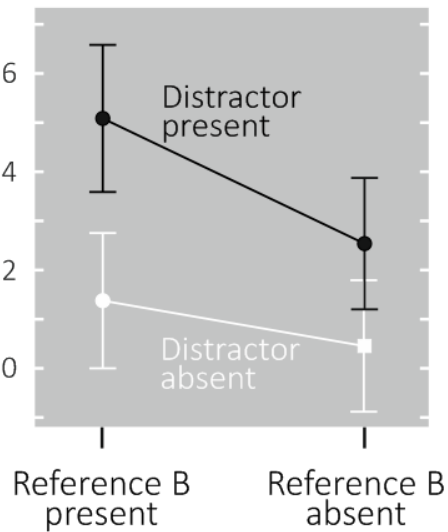
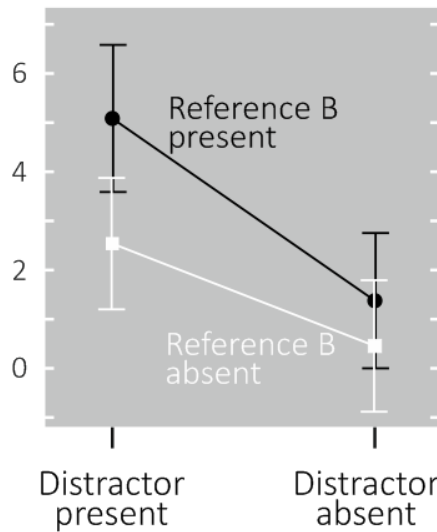


“The yellow one right of the blue one”

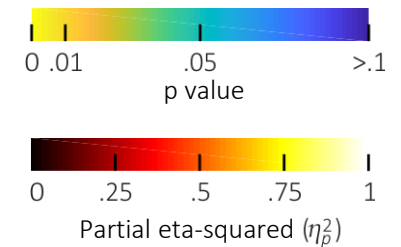


- 2x2 repeated measures ANOVA ($p < .04$)

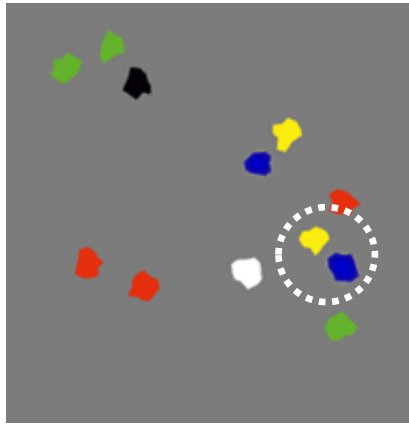
Trajectory divergence [mm]



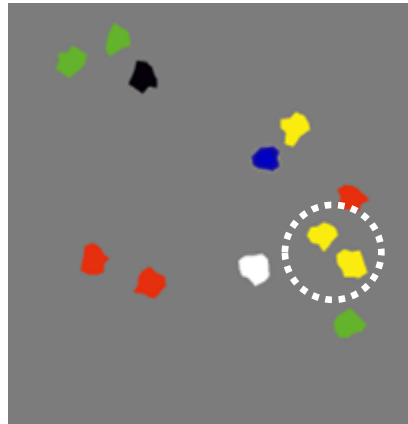
Mean Movement Time: 1101 ms



Is the interaction specific for a potential *relational* pair?



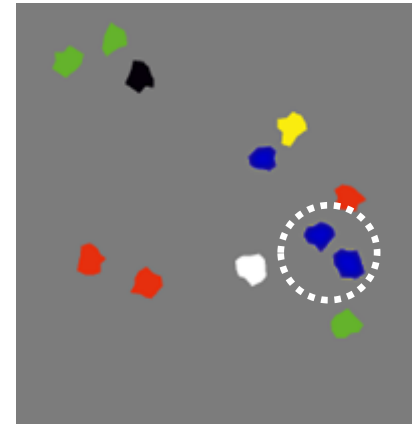
Pair B



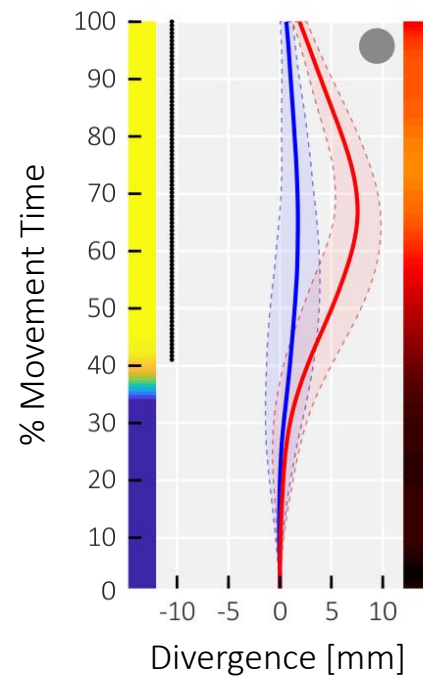
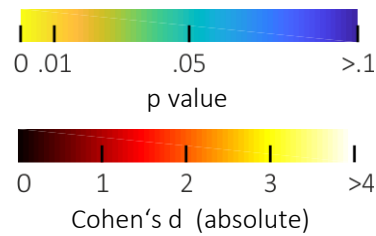
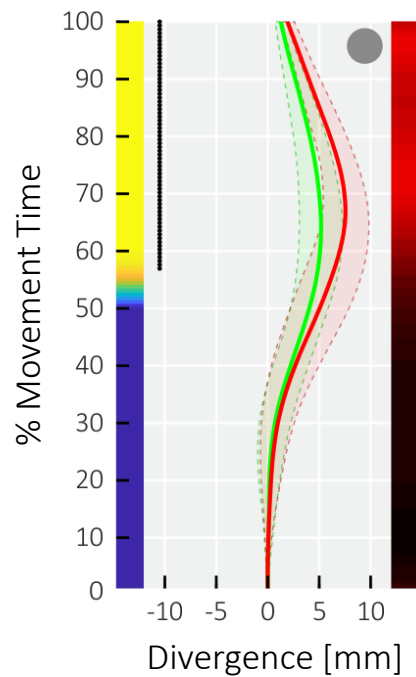
2 x Distractor



Pair B



2 x Reference B



Summary & Outlook

- In the neural process model of spatial language grounding **all potential reference and target items must receive attentional allocation**
- Four motor signatures of grounding processes
 - Attraction to **distractor**
 - Attraction to **reference** (not a potential target!)
 - Strong attraction to **potential relational pair**
 - But not toward same-colored pair
- Example for how DFT models can serve as heuristics for experimental research questions.
- Next step: Biologically plausible movement generation → trajectories?