Active Bayesian Associative Learning

John K. Kruschke
Indiana University

April, 2008

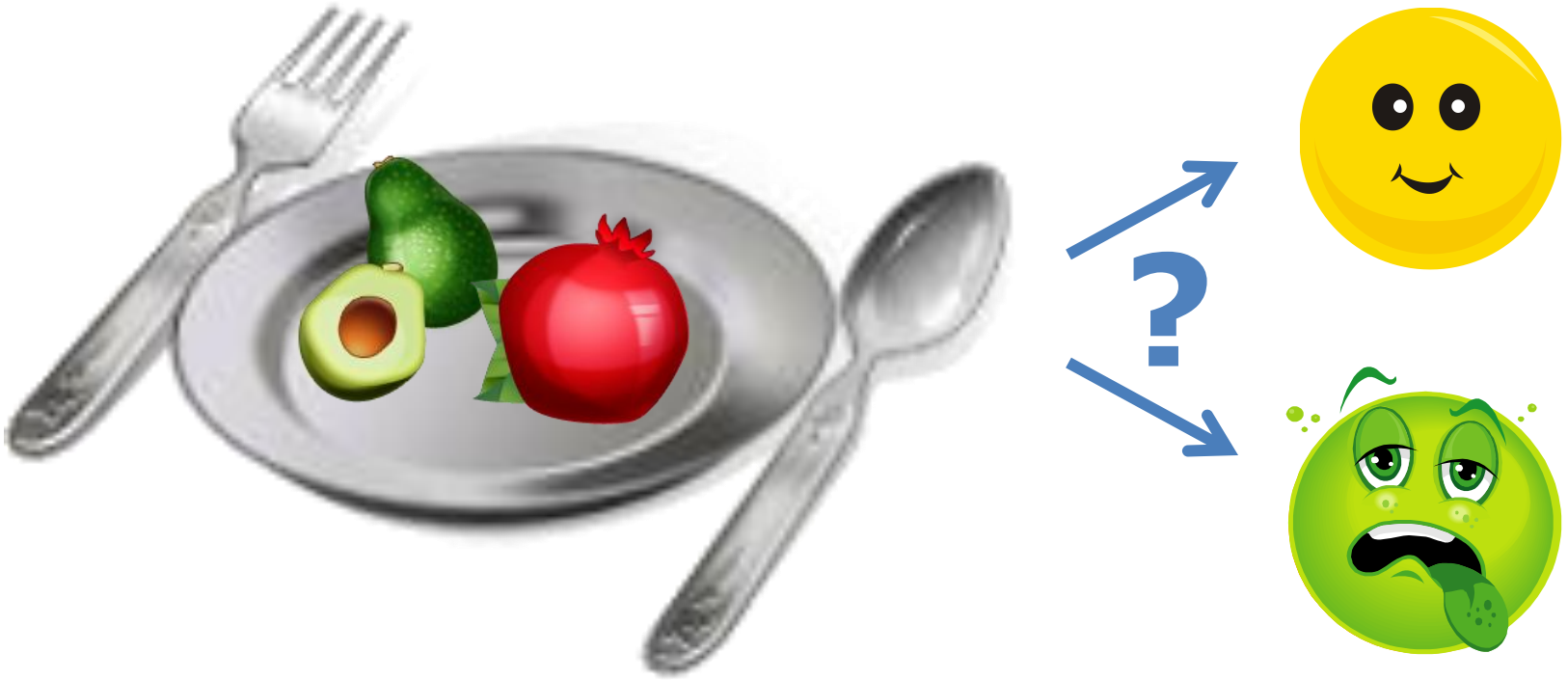
Background:

- **A typical associative learning task, and a phenomenon from associative learning: "Backward Blocking".**

- **A traditional (non-Bayesian) model: The Rescorla-Wagner model.**

Typical Learning Task

Learn which foods produce *or prevent* nausea:



Backward Blocking

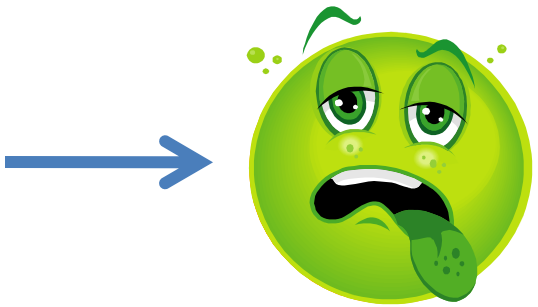
Phase	Frequency	Cue 1	Cue 2	Outcome
I	10	1 	1 	1 
II	10	1 	0	1 

Note: Cell with a 1 indicates presence of cue or outcome. Cell with a 0 indicates absence of cue or outcome.

Please predict nausea or no nausea:

?

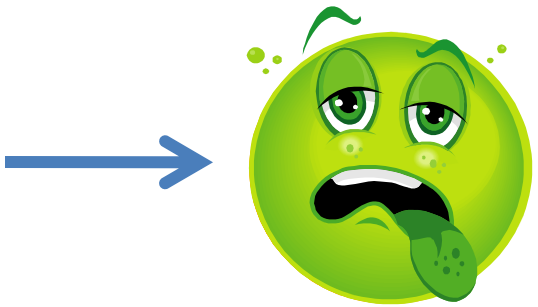
Actual result:



Please predict nausea or no nausea:

?

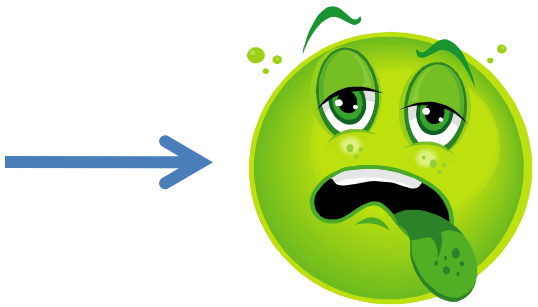
Actual result:



Please predict nausea or no nausea:

?

Actual result:



Interim Test (no corrective feedback):

Please predict nausea or no nausea:

?

Interim Test (no corrective feedback):

Please predict nausea or no nausea:

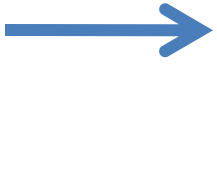
Usual response is *middling* prediction of nausea.

→ ?

Please predict nausea or no nausea:

?

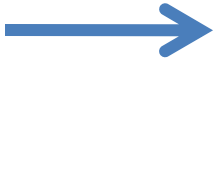
Actual result:



Please predict nausea or no nausea:

?

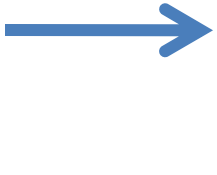
Actual result:



Please predict nausea or no nausea:

?

Actual result:



Final Test (no corrective feedback):

Please predict nausea or no nausea:

?

Final Test (no corrective feedback):

Please predict nausea or no nausea:

Usual response is *reduced* prediction of nausea.

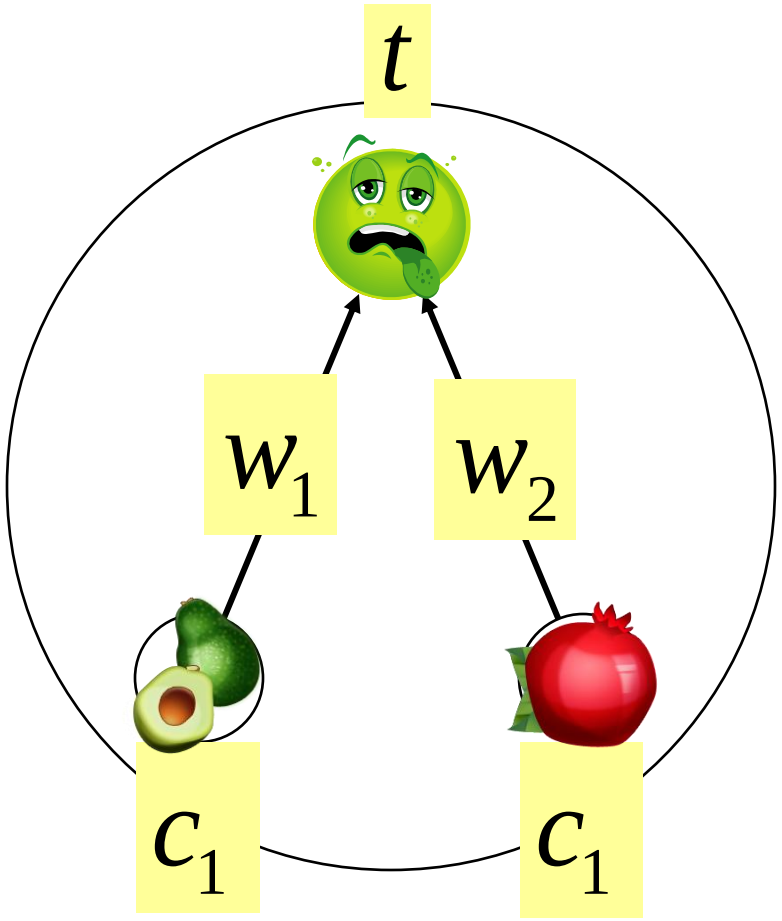
This result (with its training structure) is called "backward blocking".

Active Learning after Backward Blocking

Phase	Frequency	Cue 1	Cue 2	Outcome
I	10	1 	1 	1 
II	10	1 	0	1 

After being trained on backward blocking, your goal is to better determine which cues produce or prevent nausea. Which probe below do you prefer to test?

Associative Models:



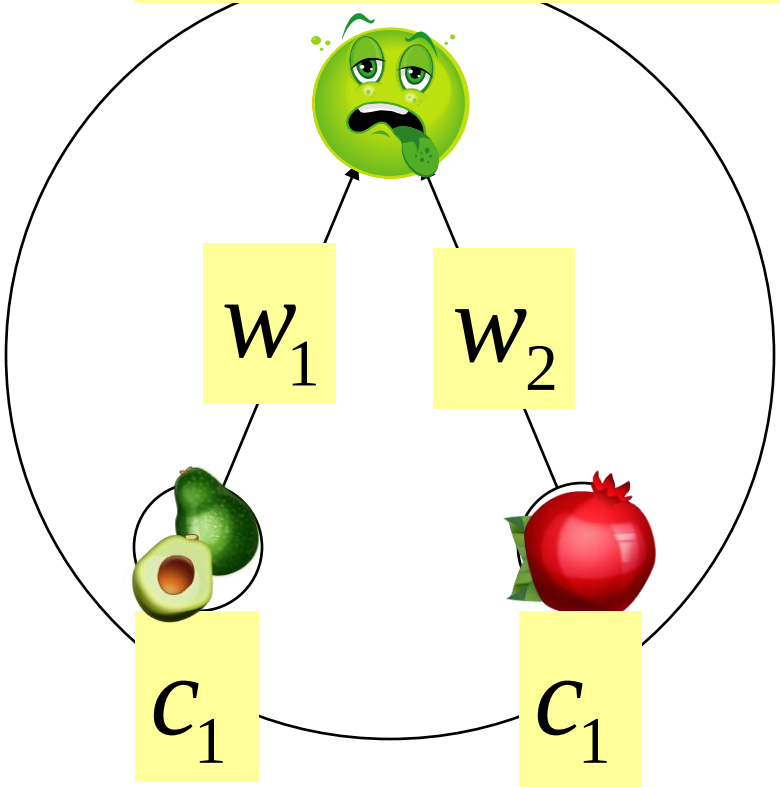
Outcome:
Modeled as a combination of weighted cues.

Associative Weights

Cues

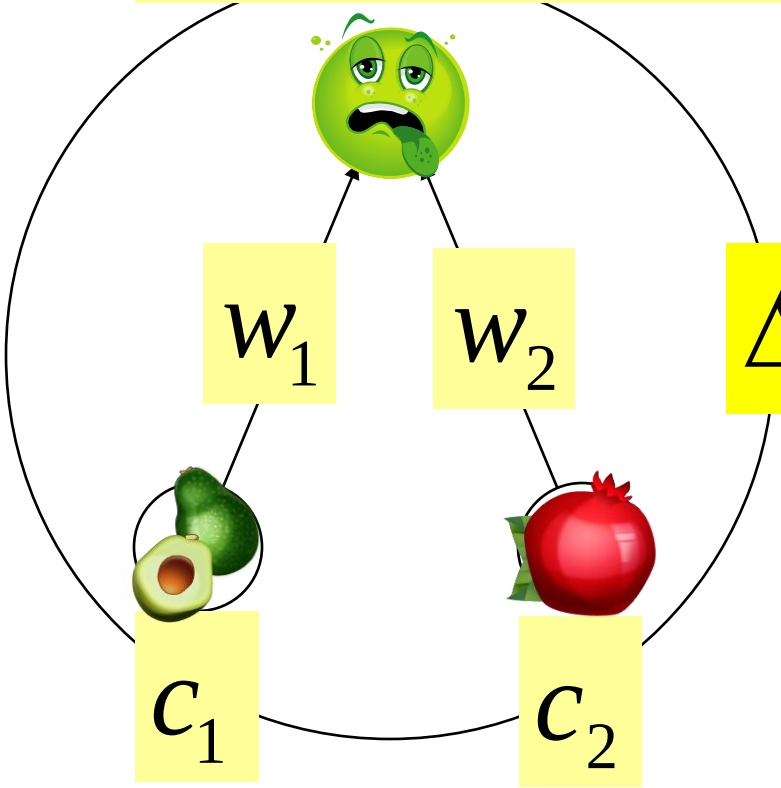
A classic non-Bayesian approach: The Rescorla-Wagner model.

$$a = \sum_i w_i c_i = \vec{w}^T \vec{c}$$



Learning
in the Rescorla-Wagner model.

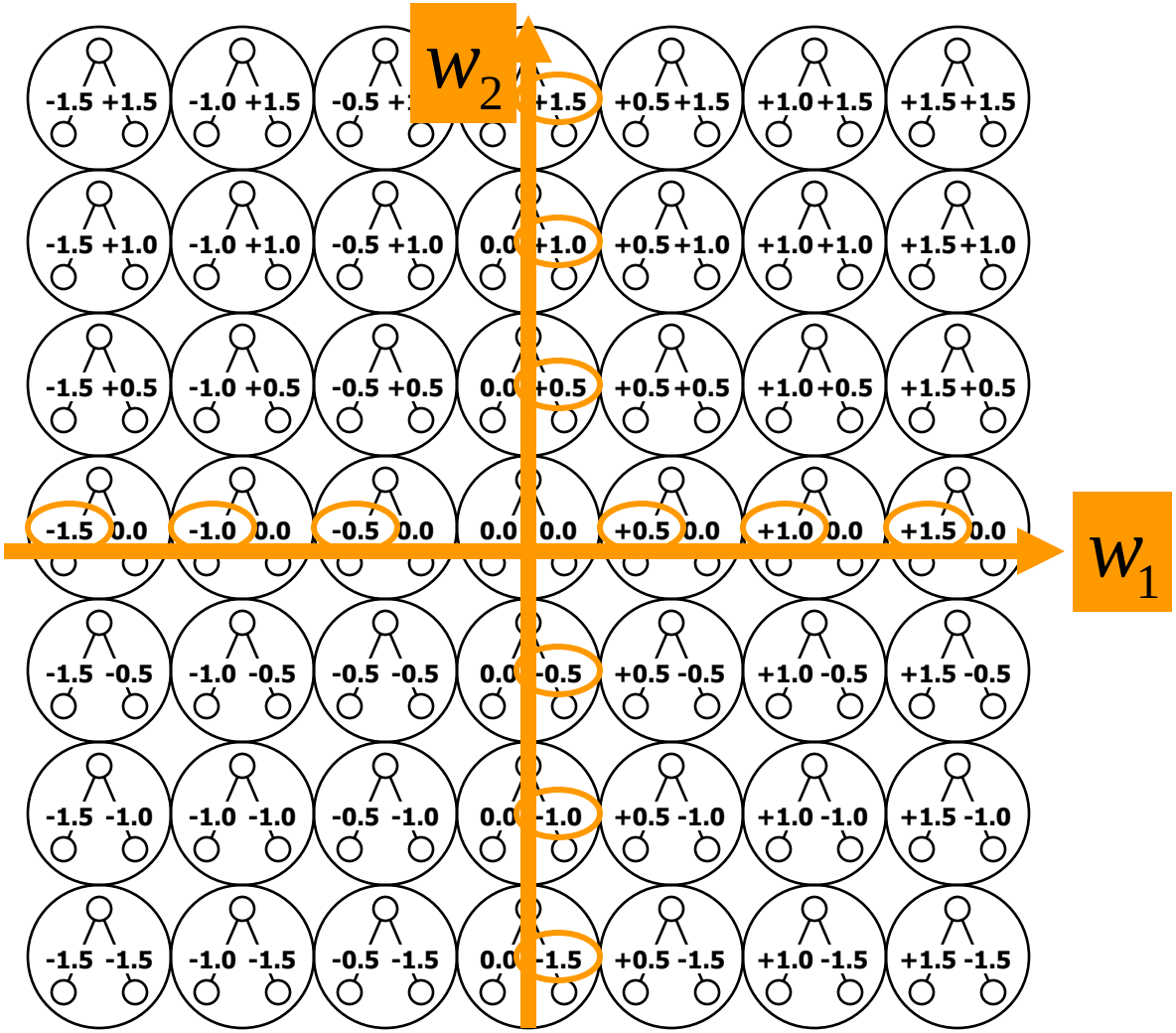
$$a = \sum_i w_i c_i = \vec{w}^T \vec{c}$$



$$\Delta \vec{w} = \lambda \left(t - \vec{w}^T \vec{c} \right) \vec{c}$$

Weight Space:

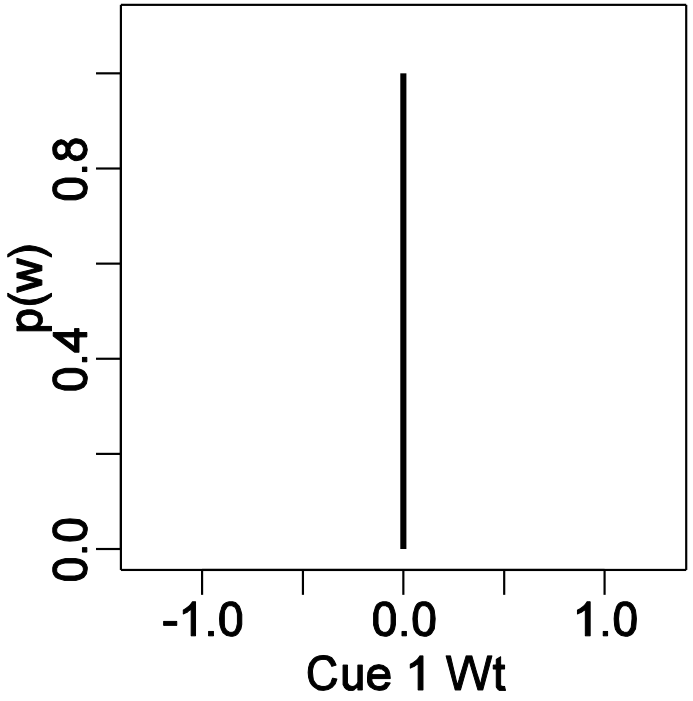
Learning is change from one position to another.



Backward Blocking

Phase	Frequency	Cue 1	Cue 2	Outcome
I	10	1 	1 	1 
II	10	1 	0	1 

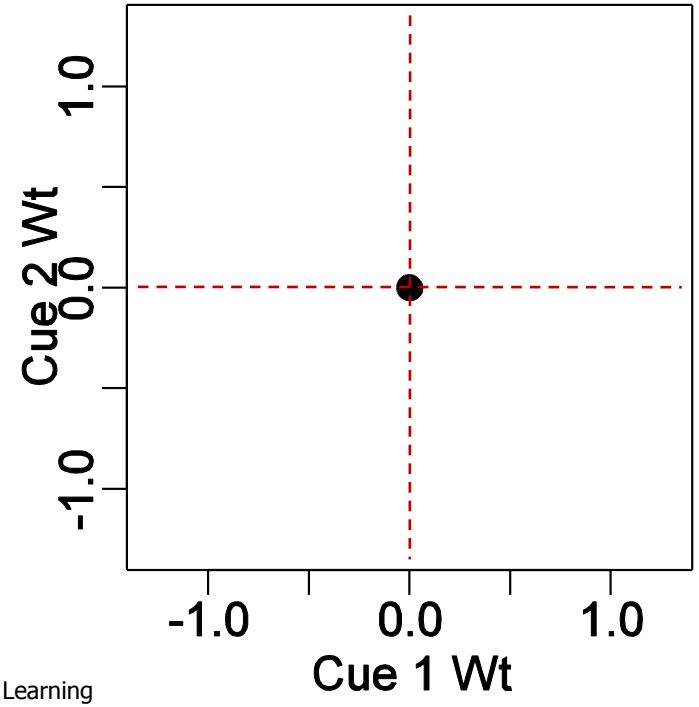
Note: Cell with a 1 indicates presence of cue or outcome. Cell with a 0 indicates absence of cue or outcome.

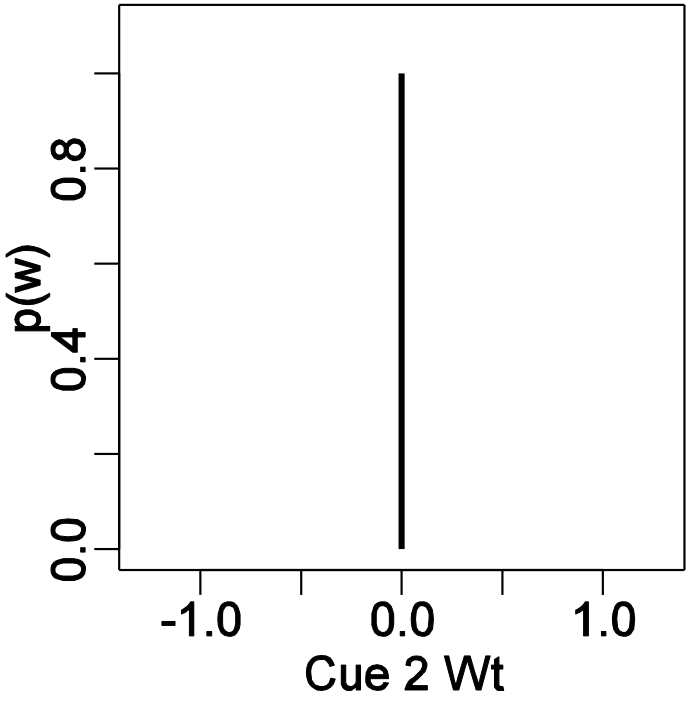


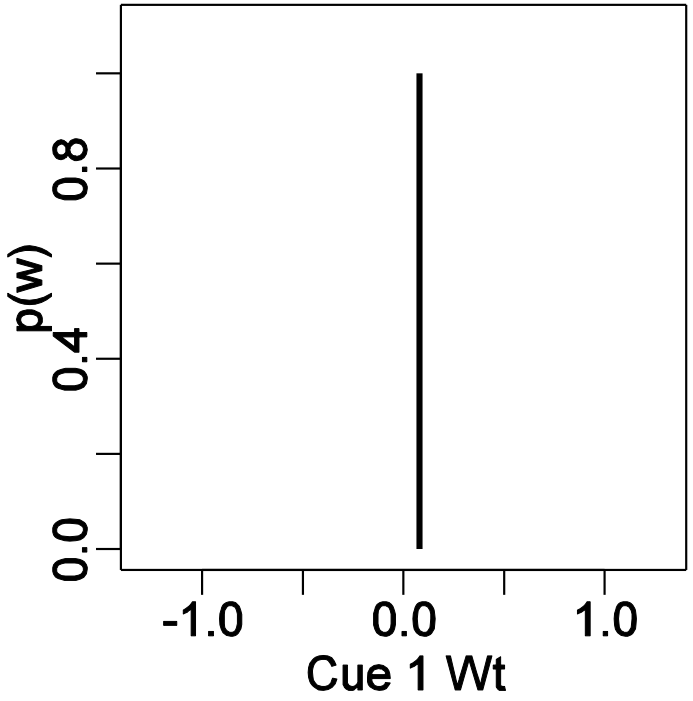
Training: BackwardBlocking

Beginning of Trial 1

Mean: 0 0



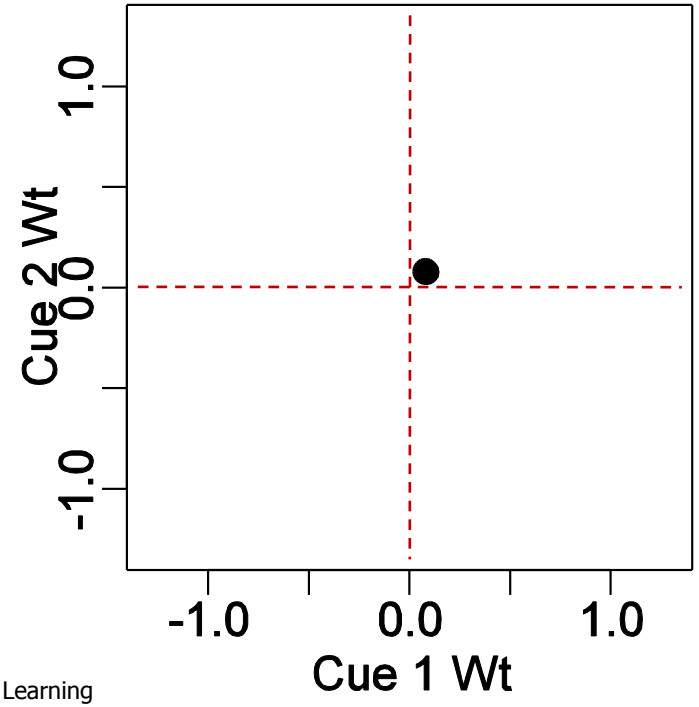


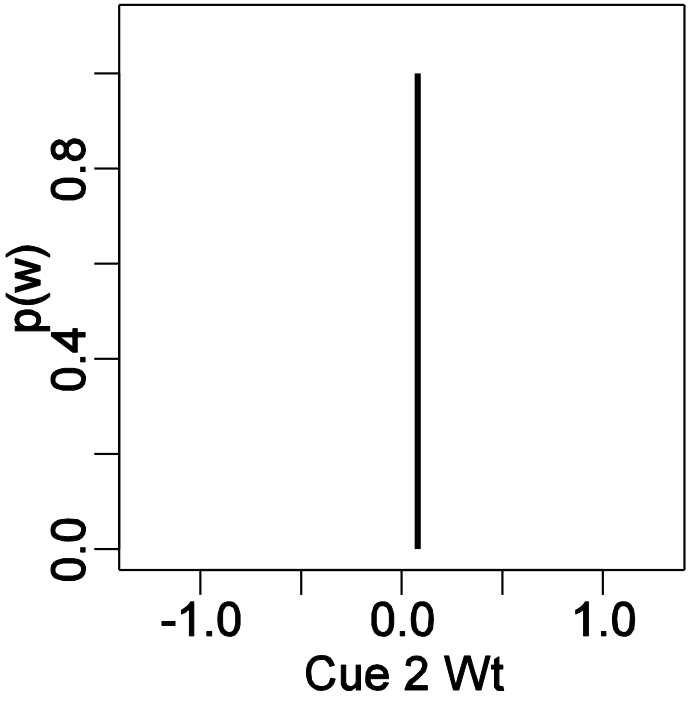


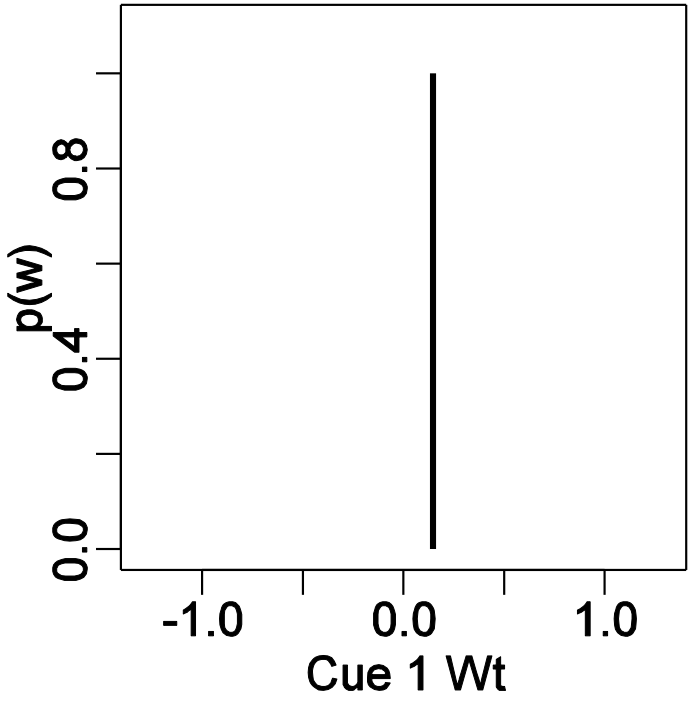
Training: BackwardBlocking

Beginning of Trial 2

Mean: 0.08 0.08



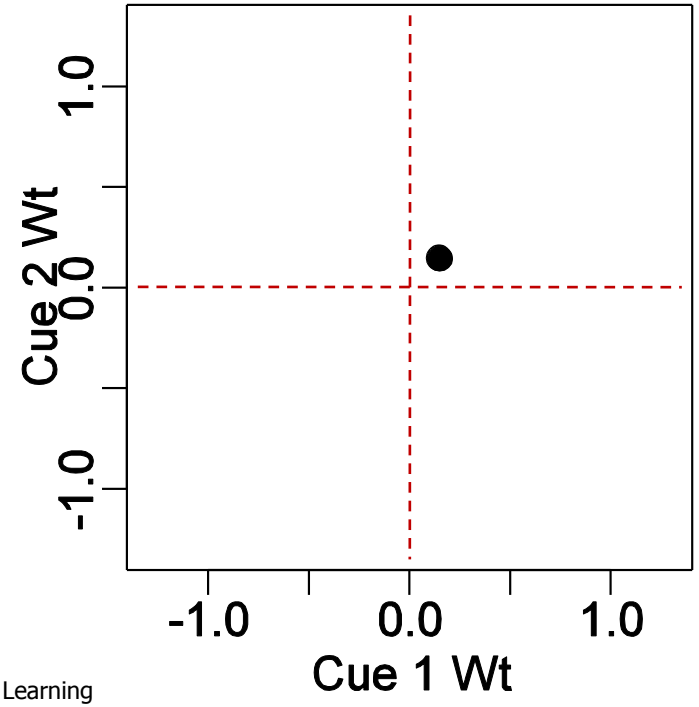


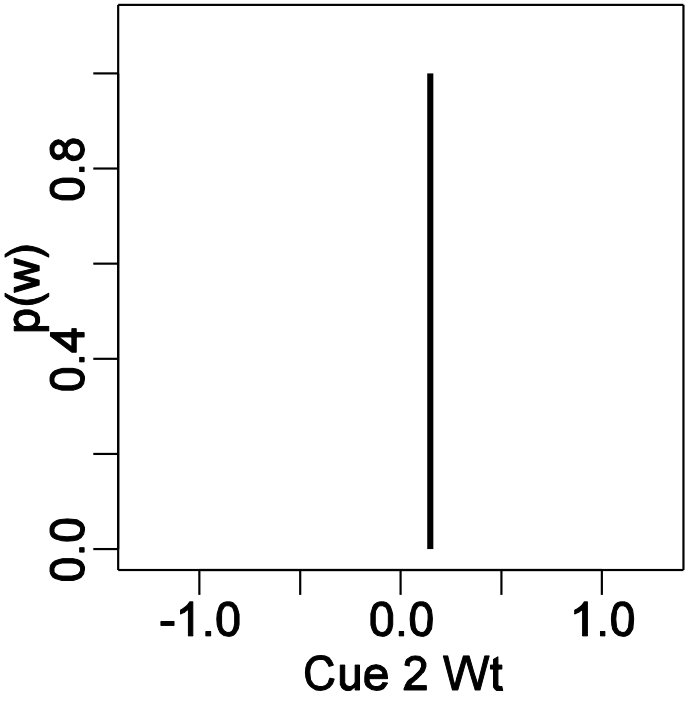


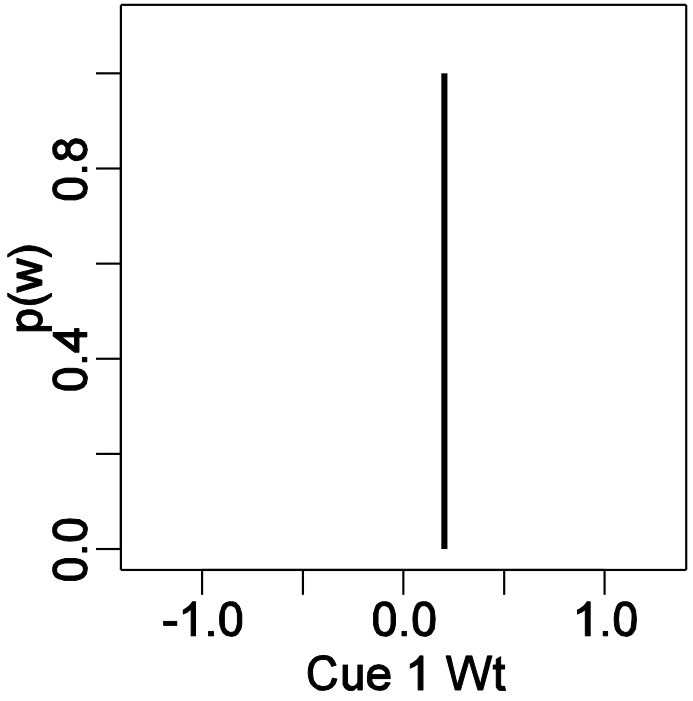
Training: BackwardBlocking

Beginning of Trial 3

Mean: 0.147 0.147



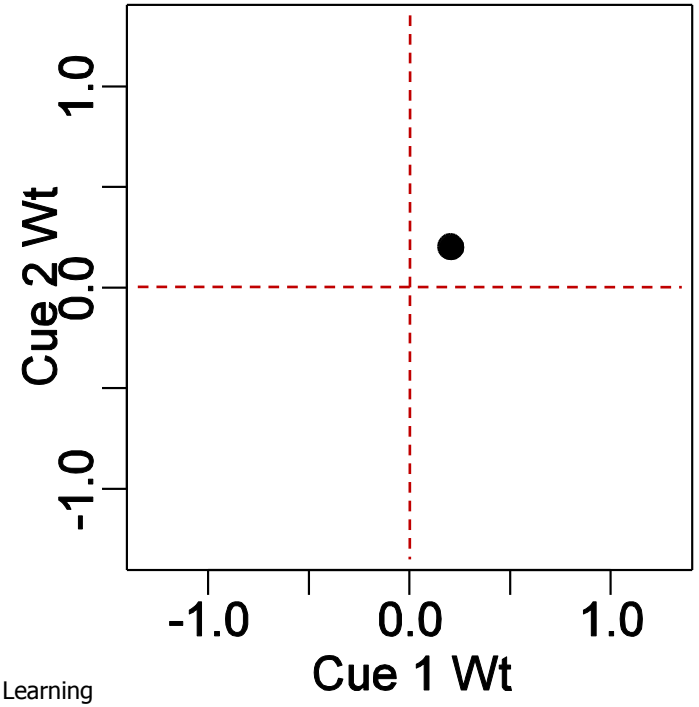


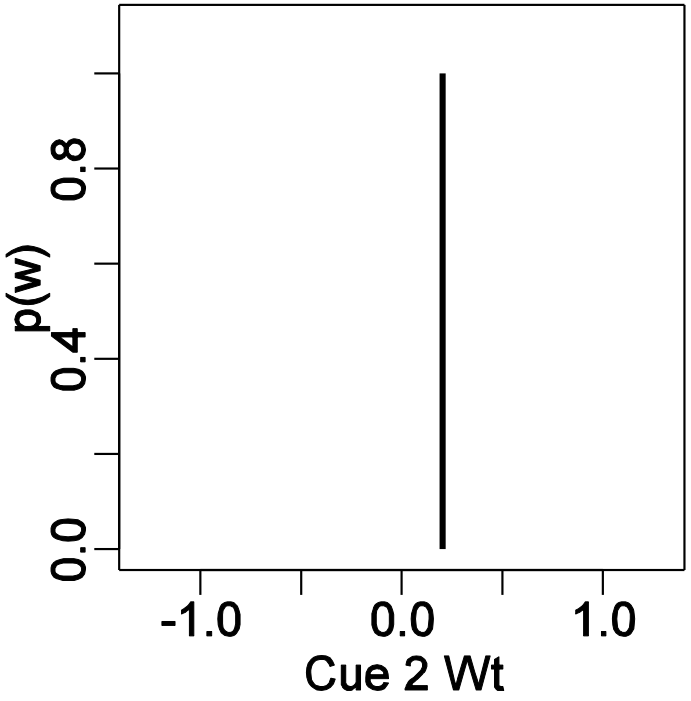


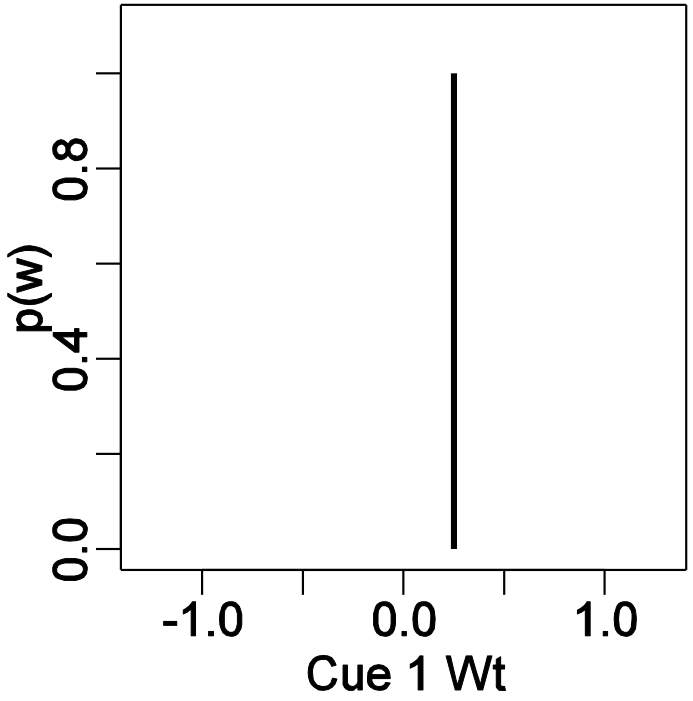
Training: BackwardBlocking

Beginning of Trial 4

Mean: 0.204 0.204



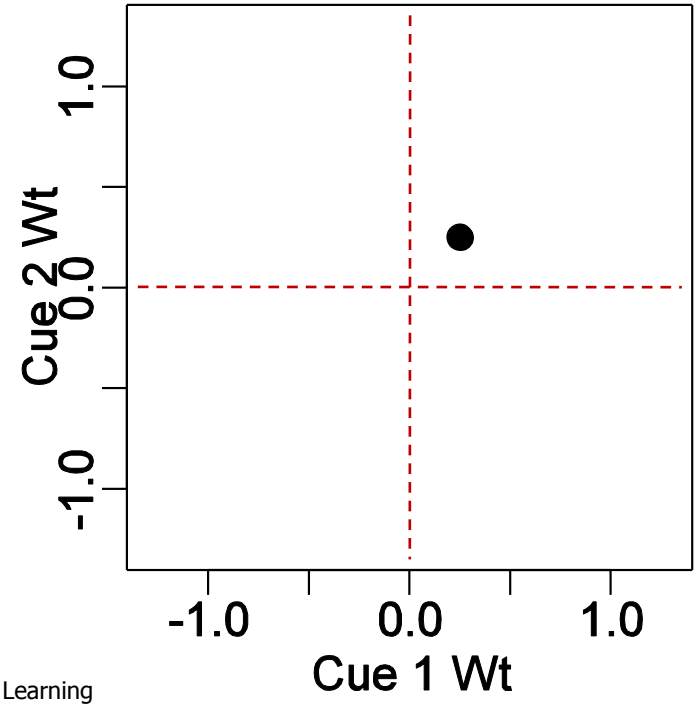


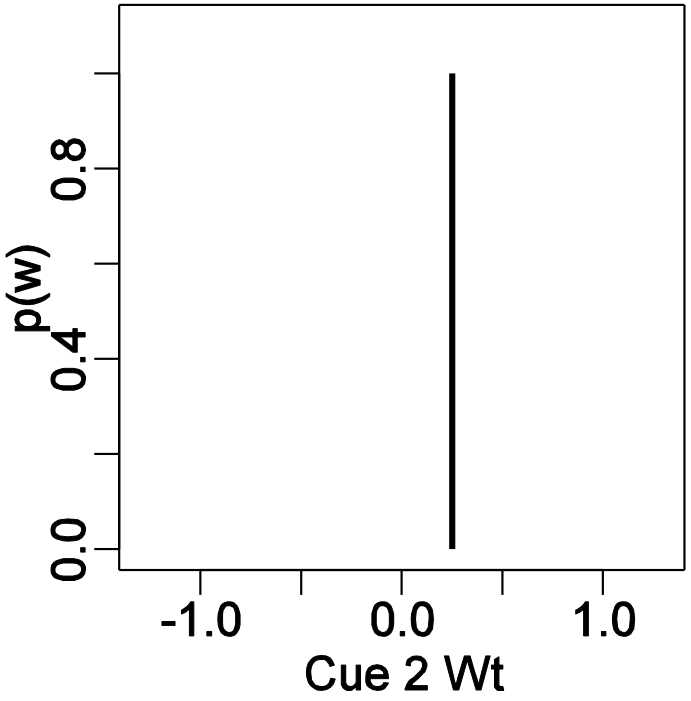


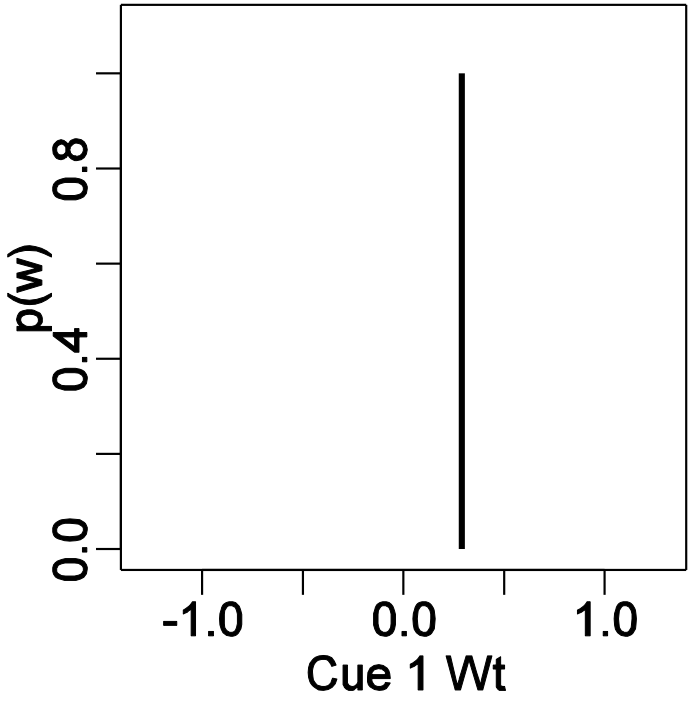
Training: BackwardBlocking

Beginning of Trial 5

Mean: 0.251 0.251



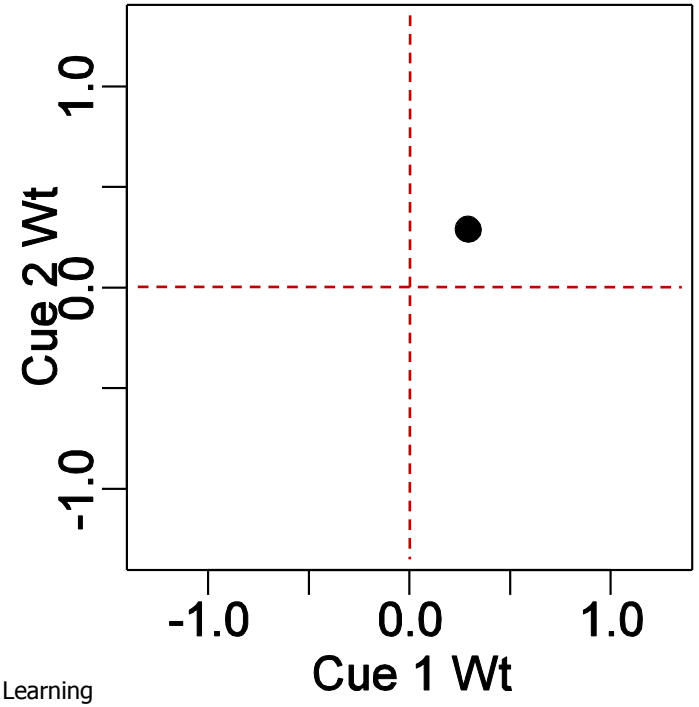


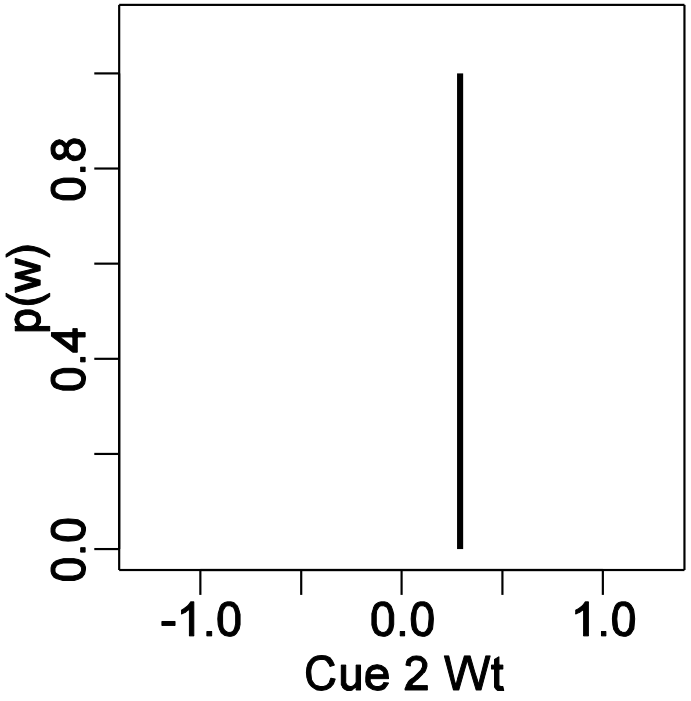


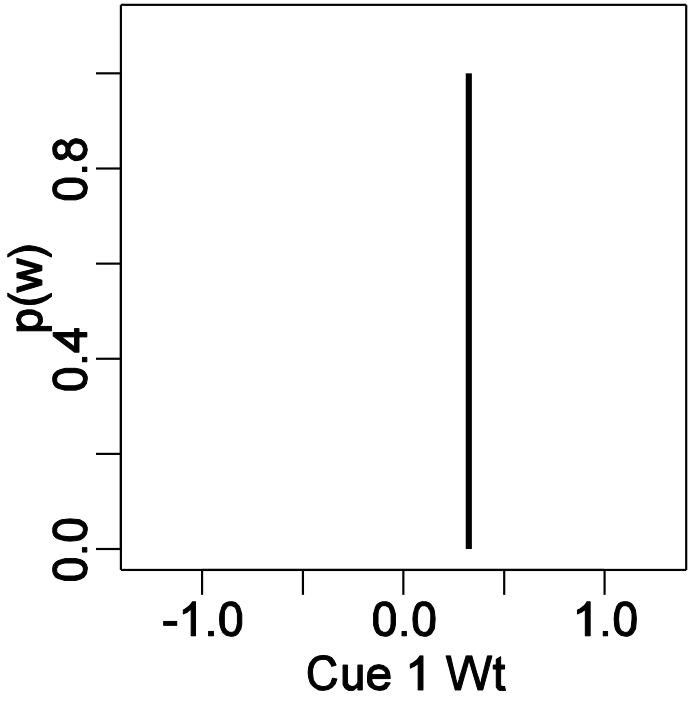
Training: BackwardBlocking

Beginning of Trial 6

Mean: 0.291 0.291



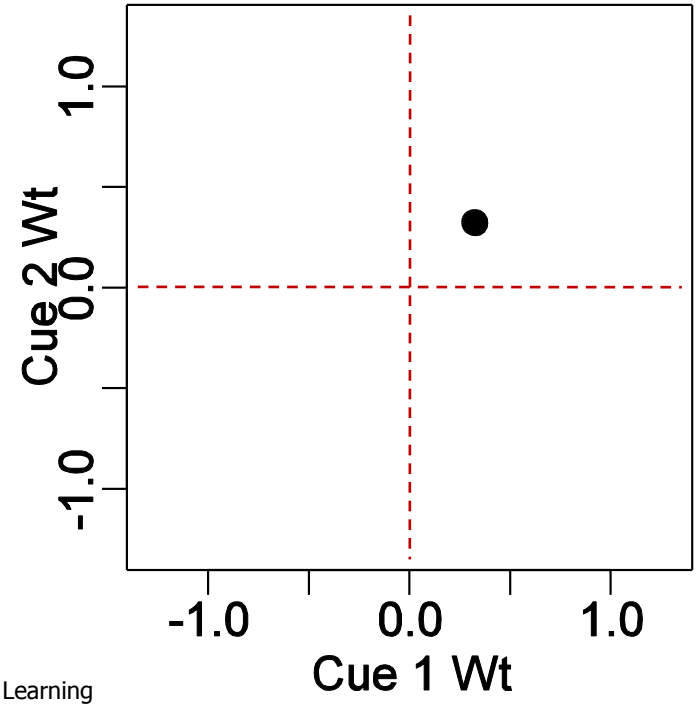


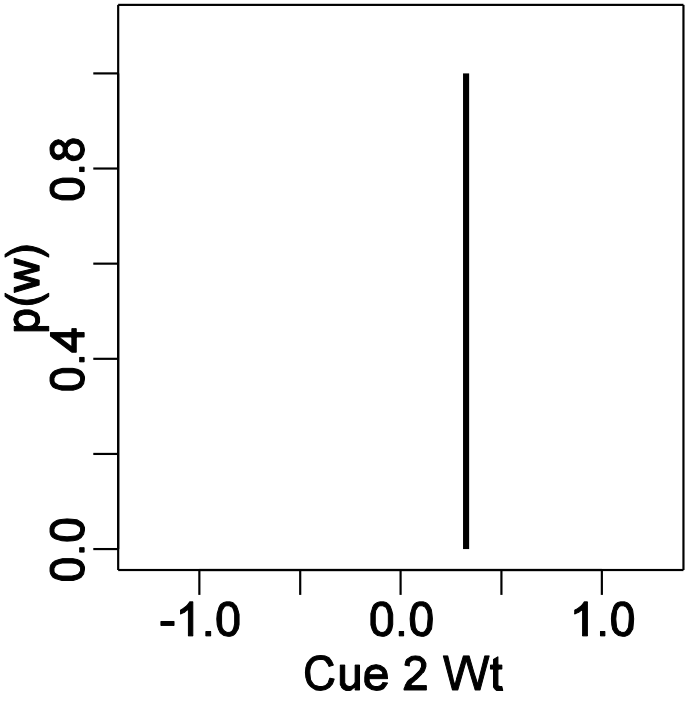


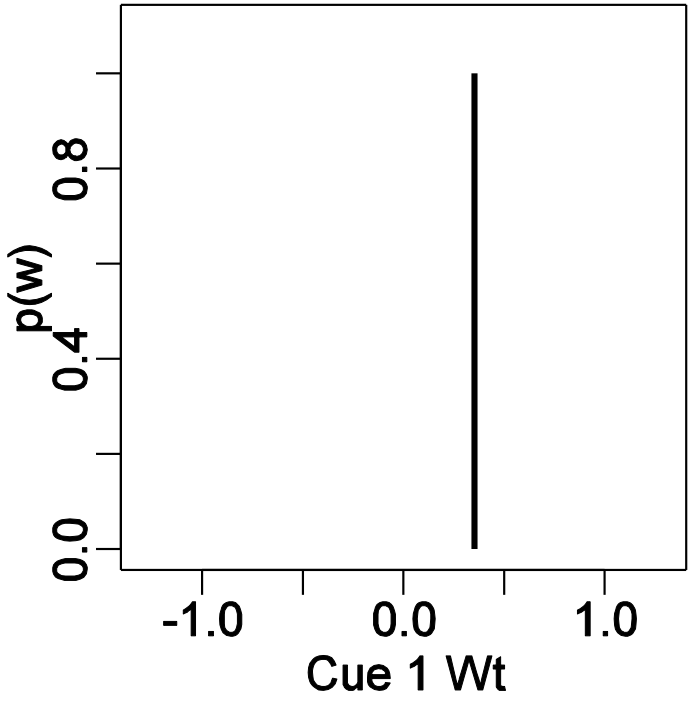
Training: BackwardBlocking

Beginning of Trial 7

Mean: 0.324 0.324



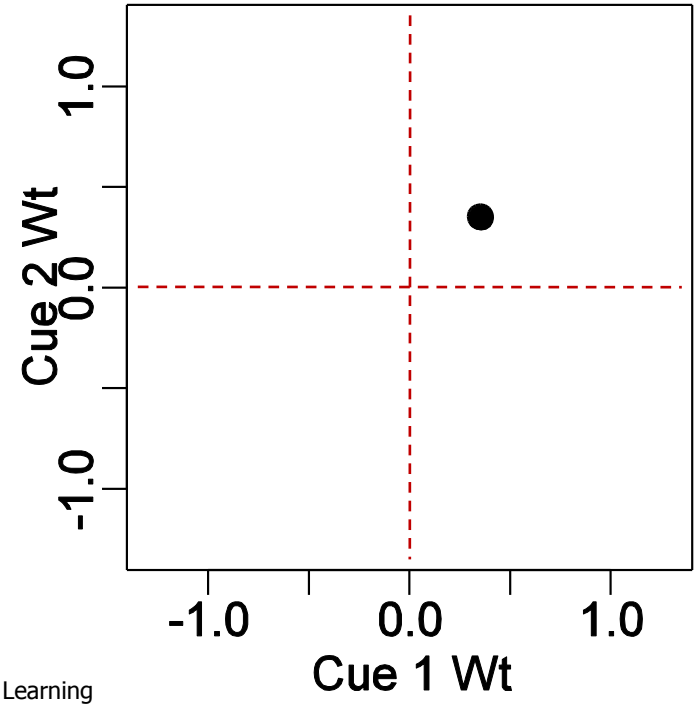


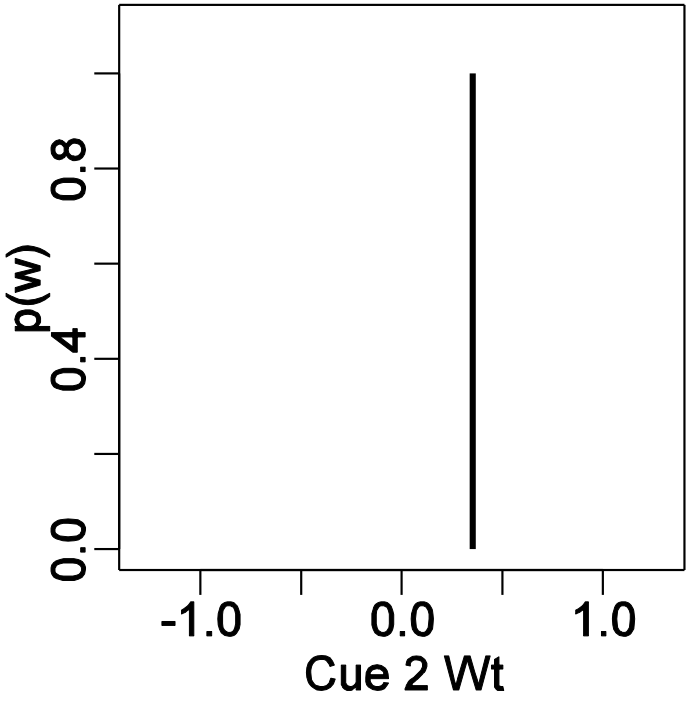


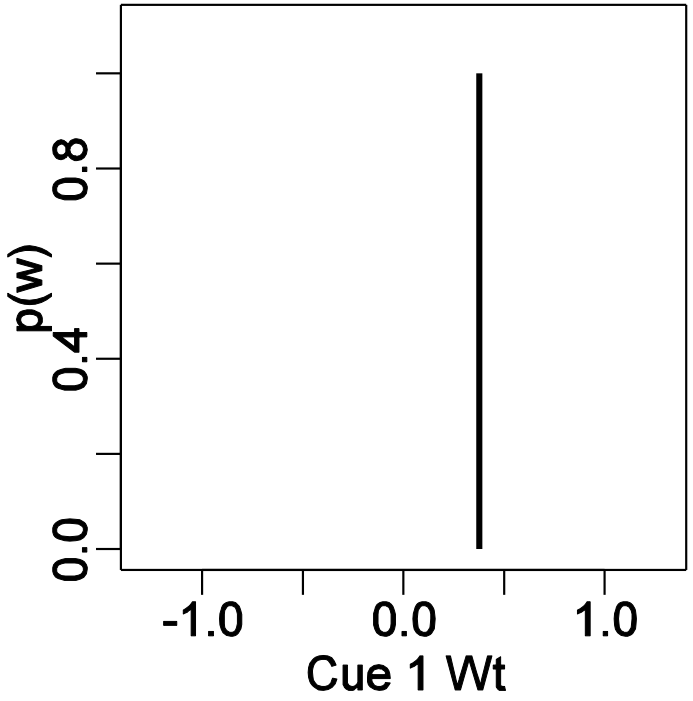
Training: BackwardBlocking

Beginning of Trial 8

Mean: 0.352 0.352



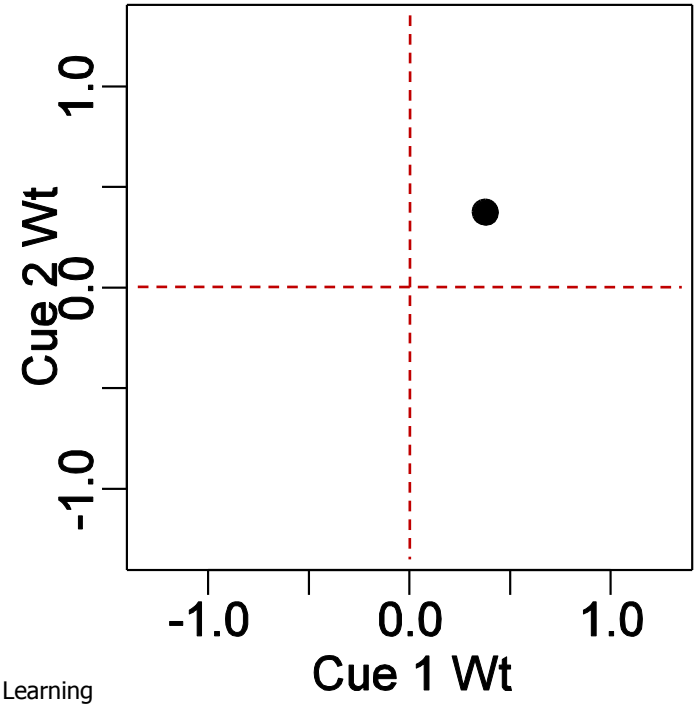


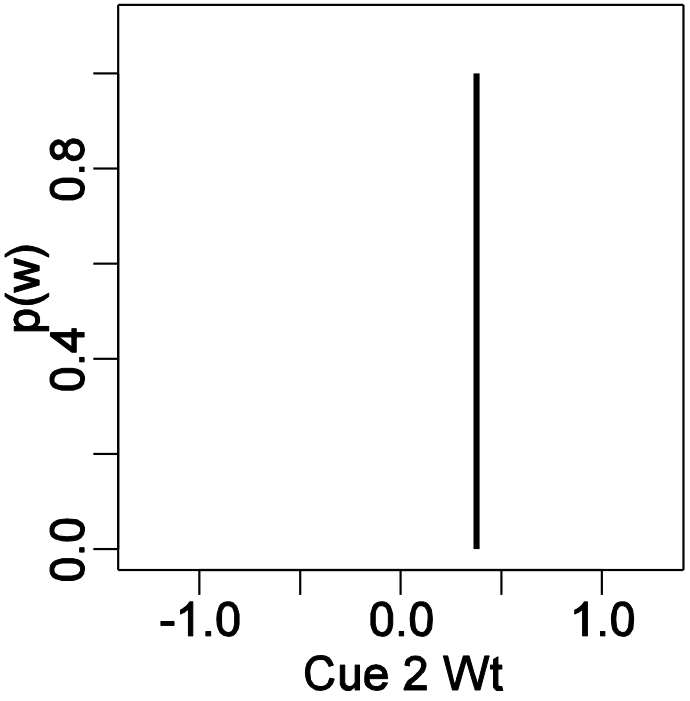


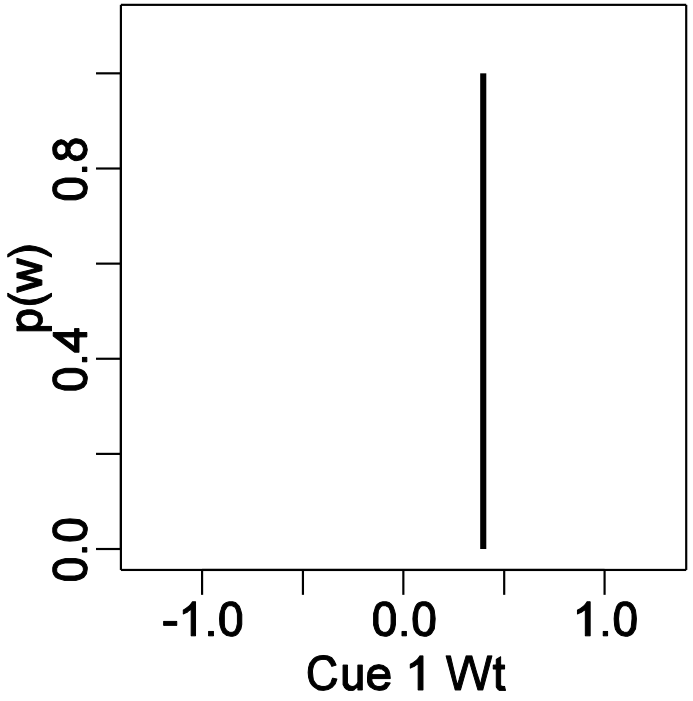
Training: BackwardBlocking

Beginning of Trial 9

Mean: 0.376 0.376



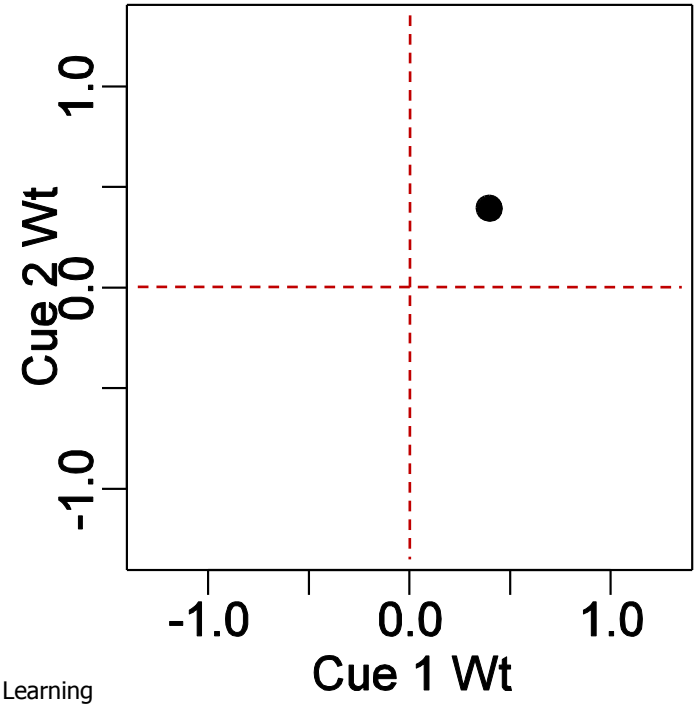


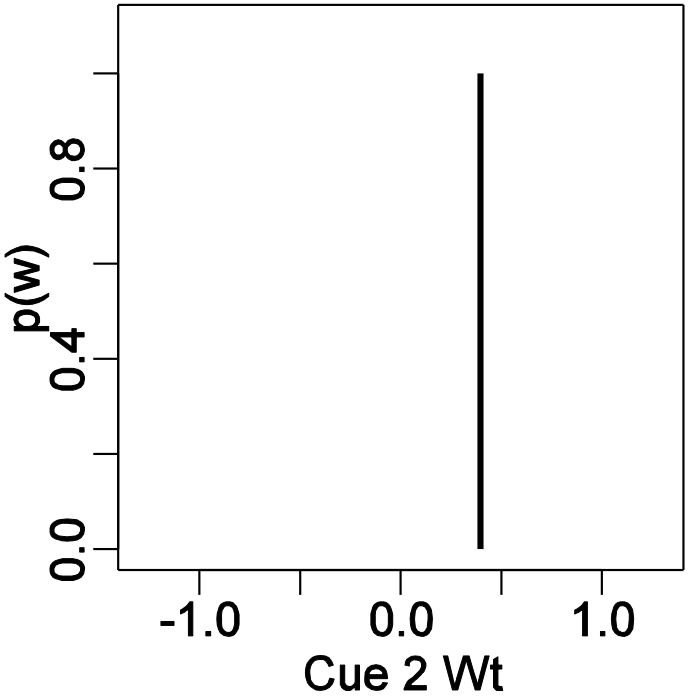


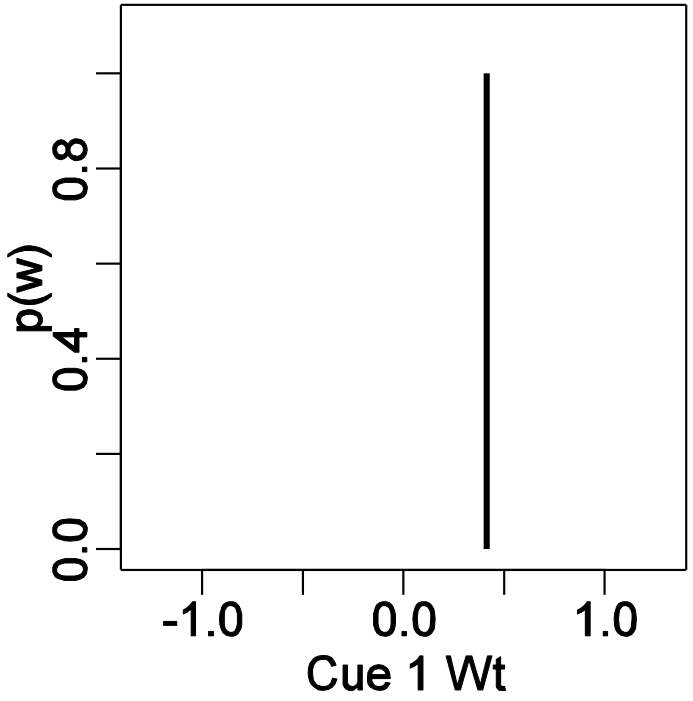
Training: BackwardBlocking

Beginning of Trial 10

Mean: 0.396 0.396



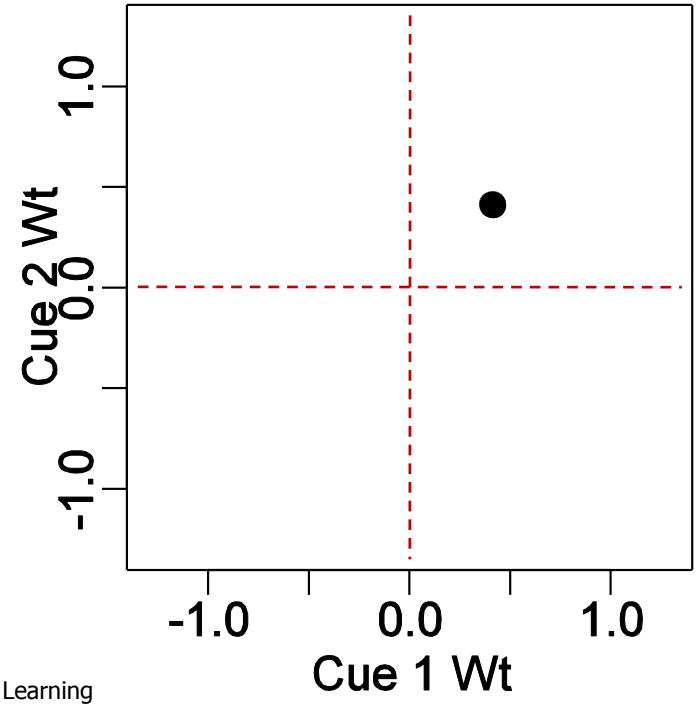


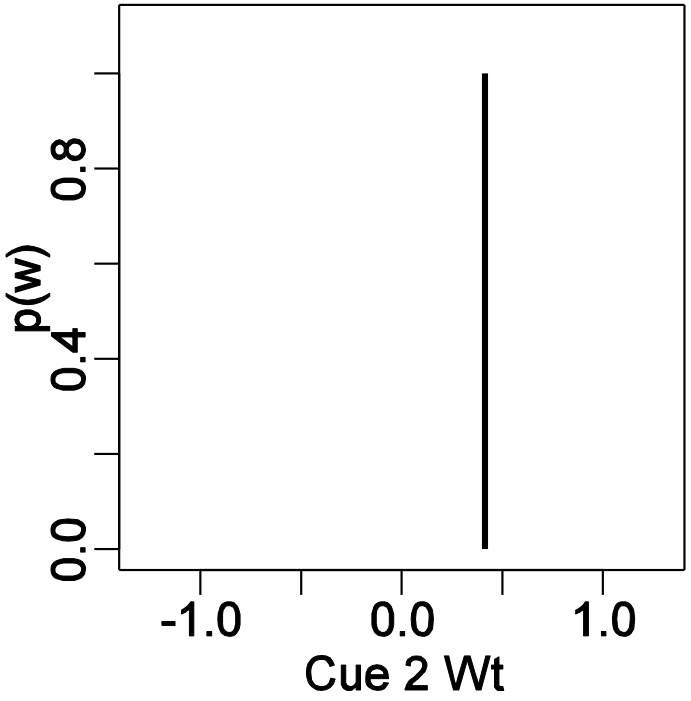


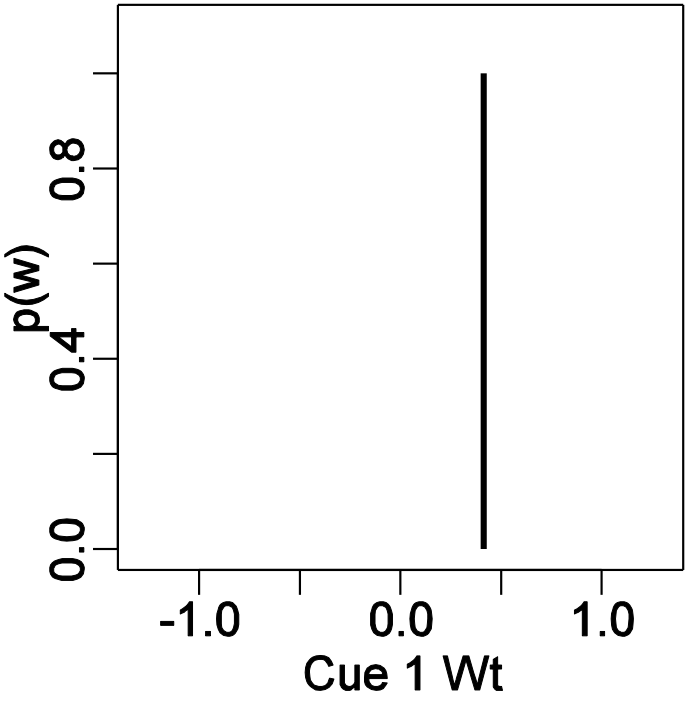
Training: BackwardBlocking

Beginning of Trial 11

Mean: 0.413 0.413





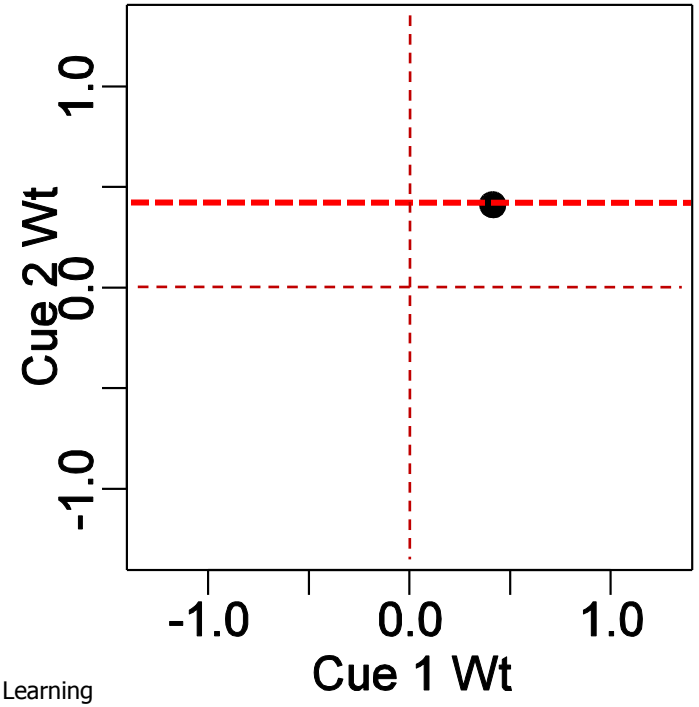


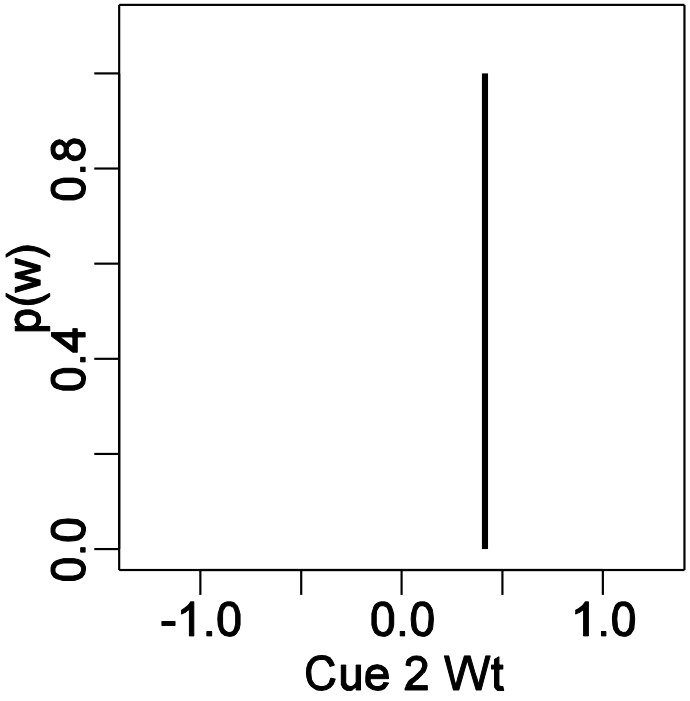
Training: BackwardBlocking

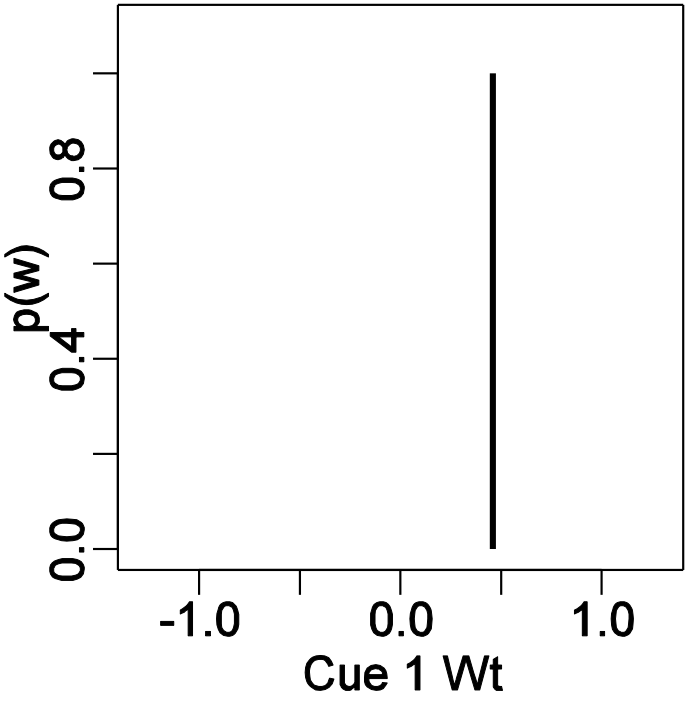
Beginning of Trial 11

Mean: 0.413 0.413

End of Phase I.



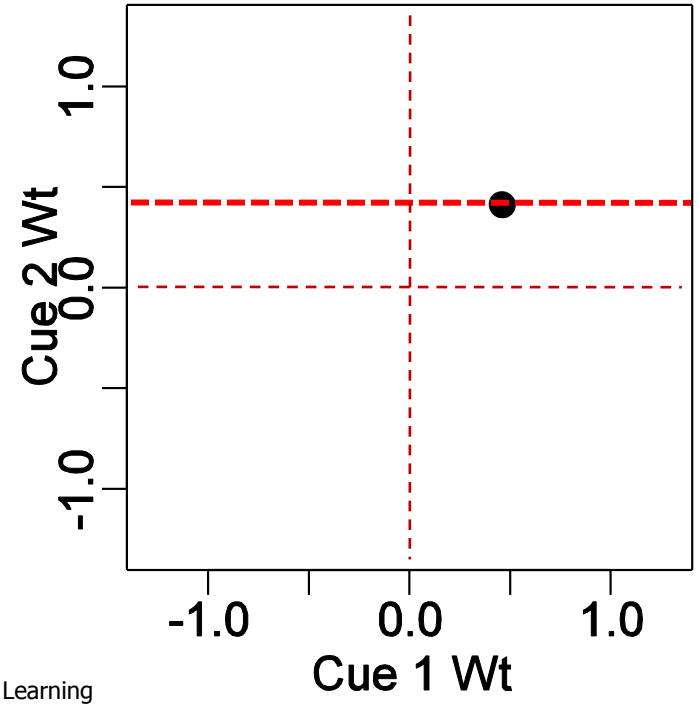


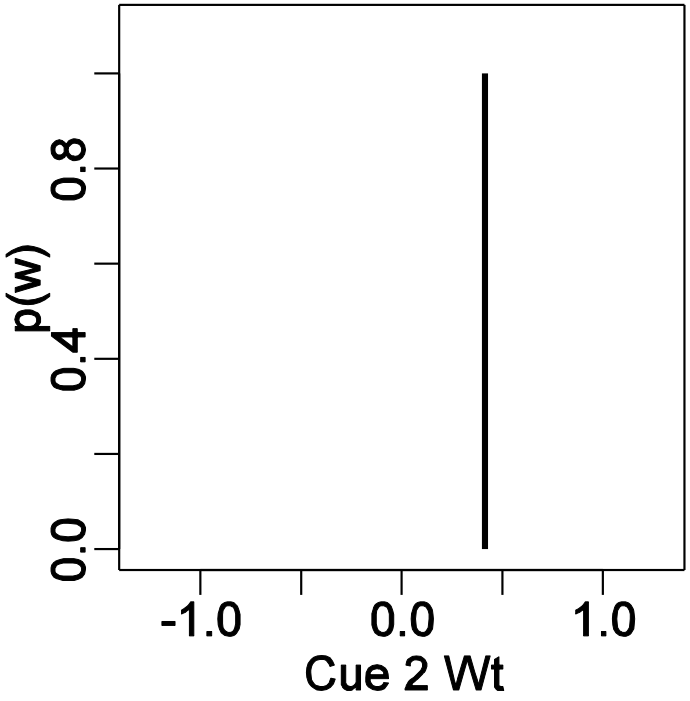


Training: BackwardBlocking

Beginning of Trial 12

Mean: 0.46 0.413

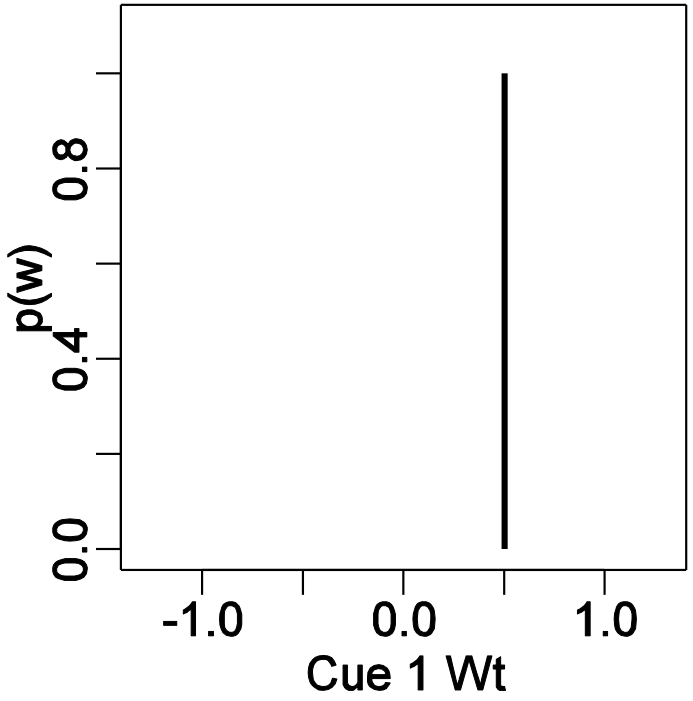


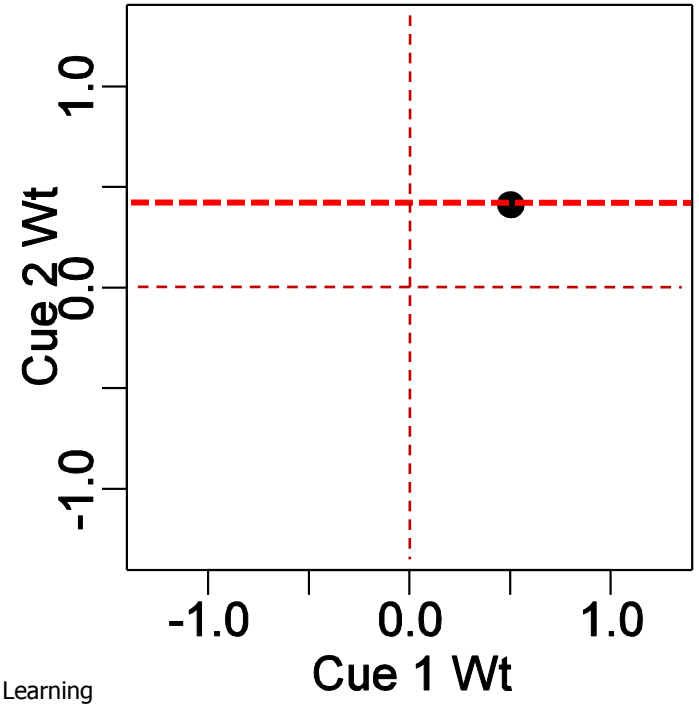


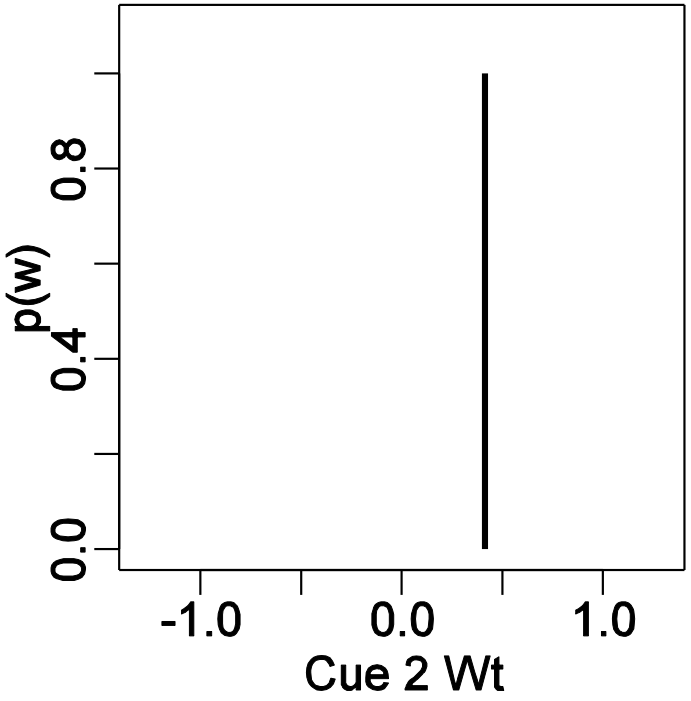
Training: BackwardBlocking

Beginning of Trial 13

Mean: 0.503 0.413



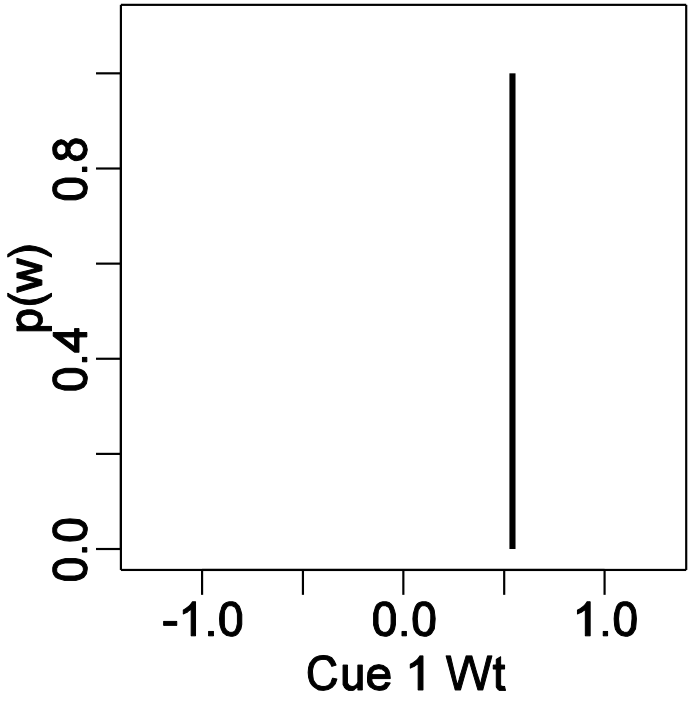


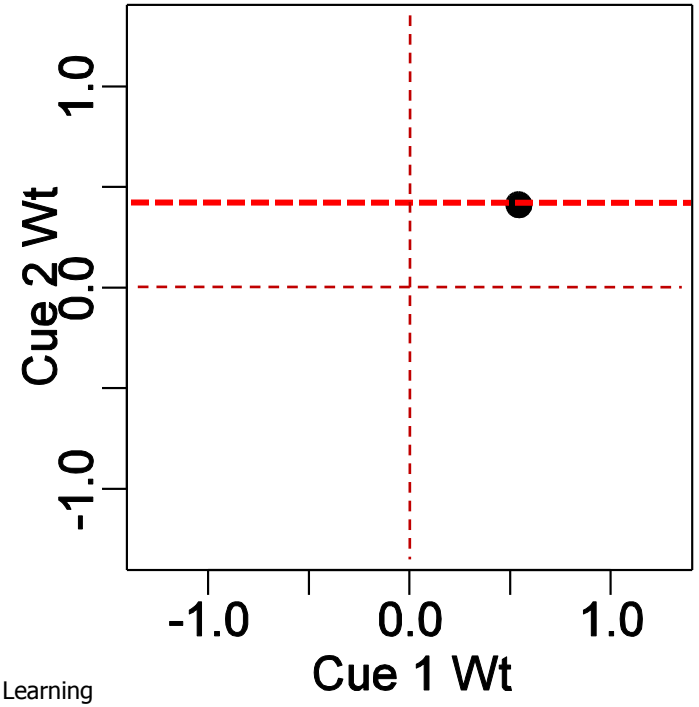


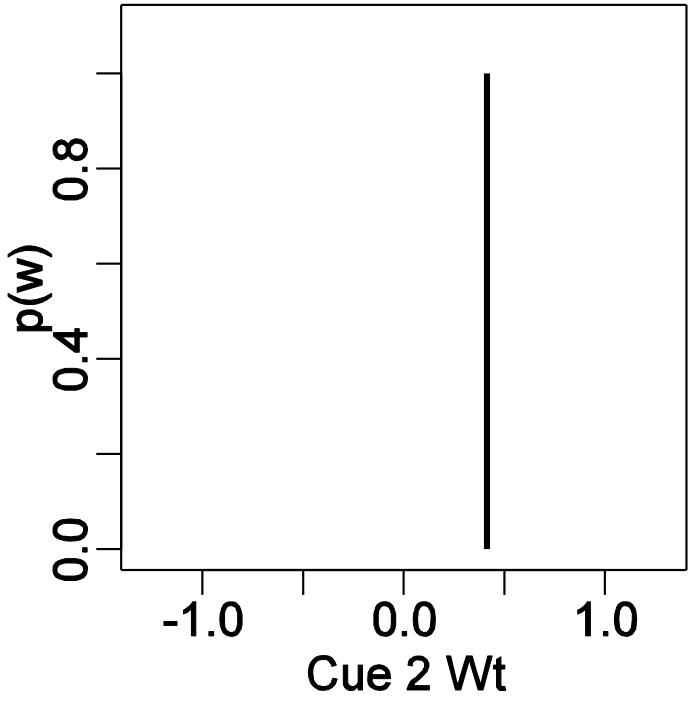
Training: BackwardBlocking

Beginning of Trial 14

Mean: 0.543 0.413



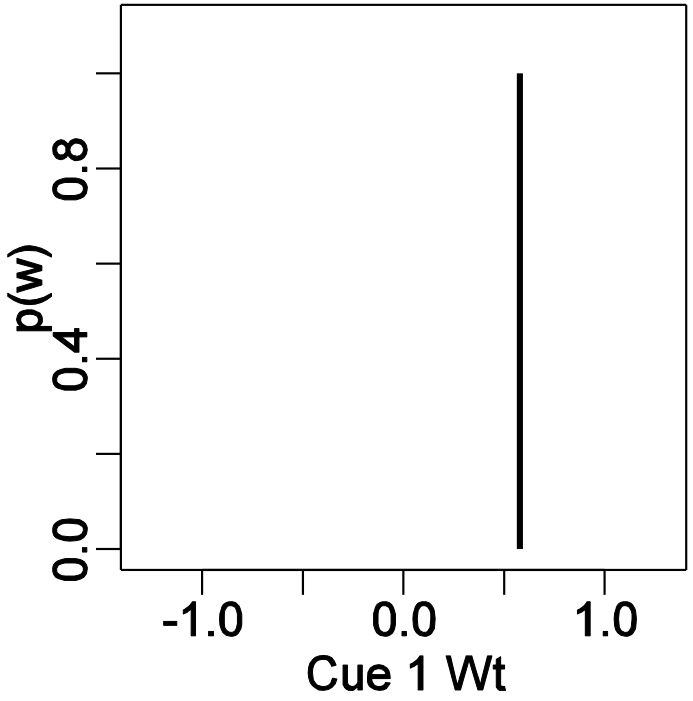


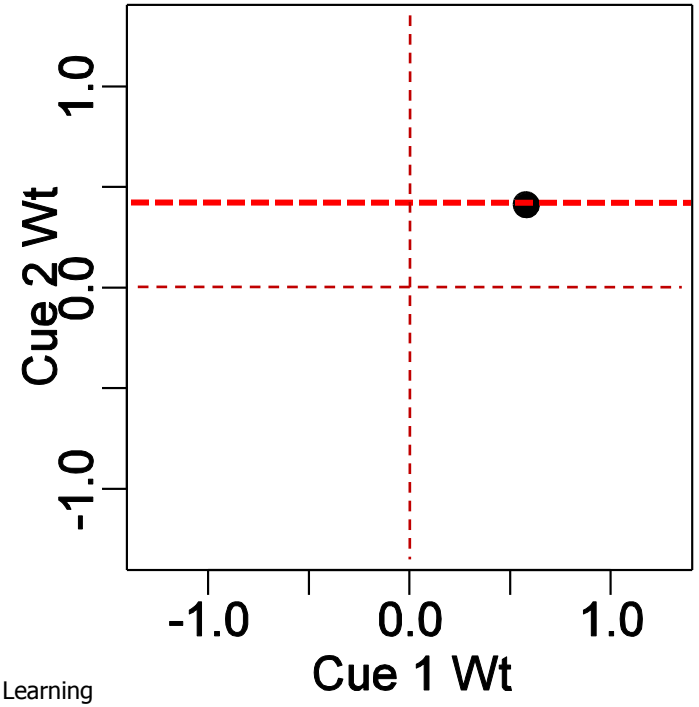


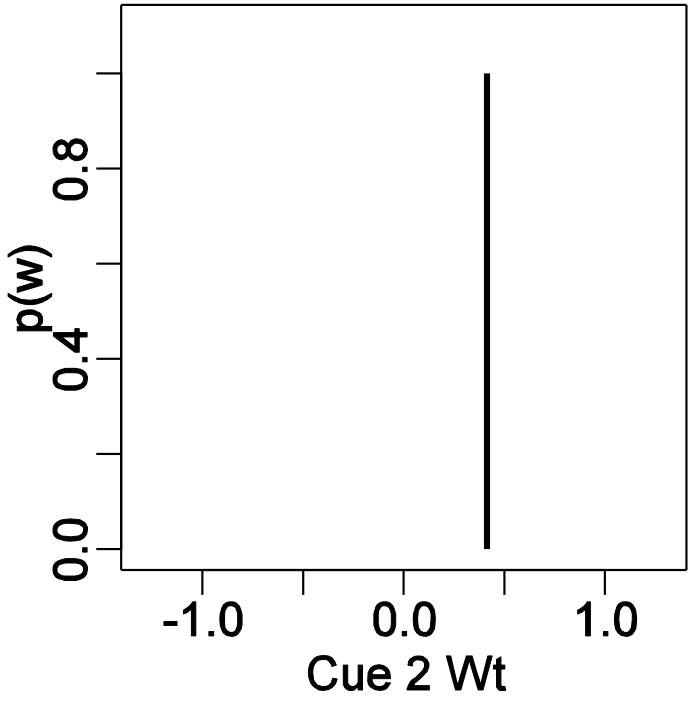
Training: BackwardBlocking

Beginning of Trial 15

Mean: 0.579 0.413



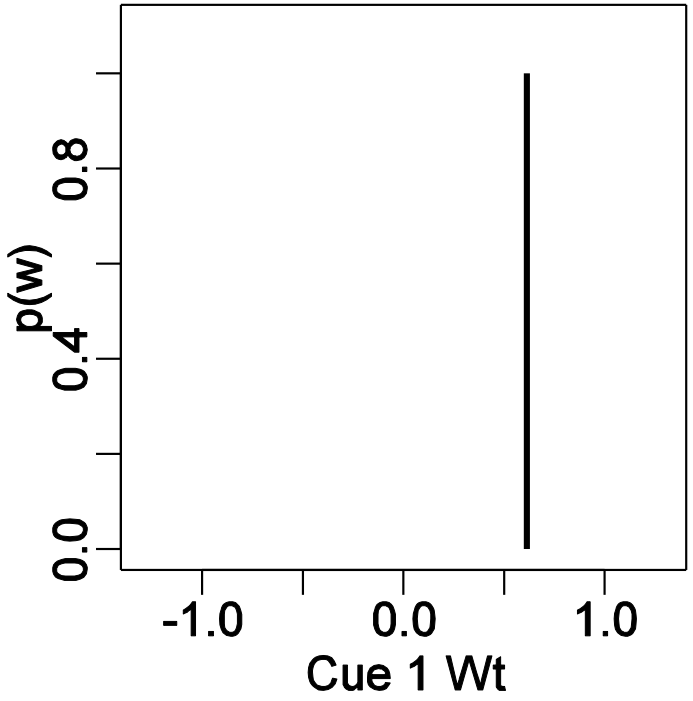


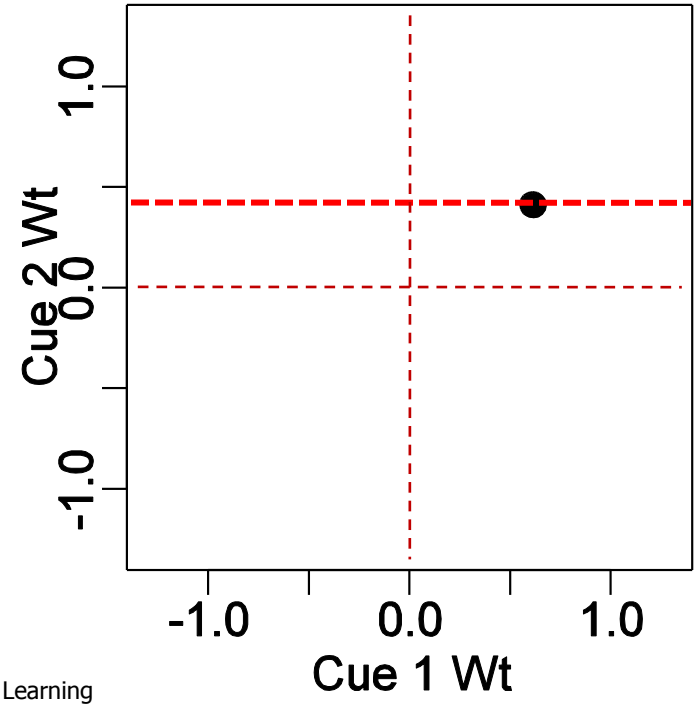


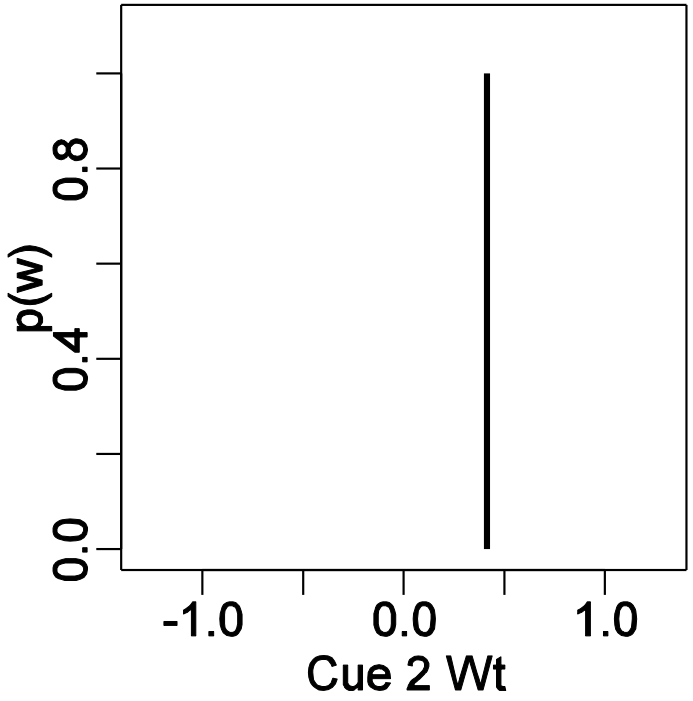
Training: BackwardBlocking

Beginning of Trial 16

Mean: 0.613 0.413



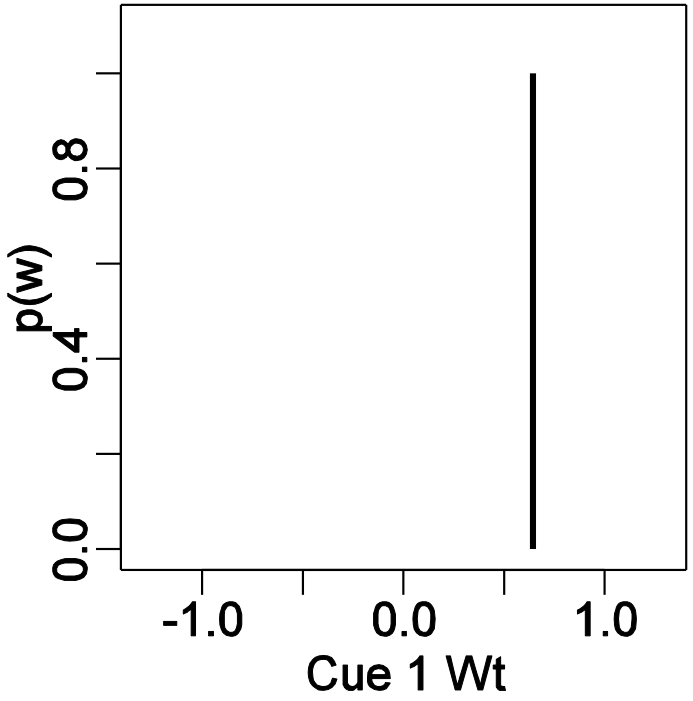


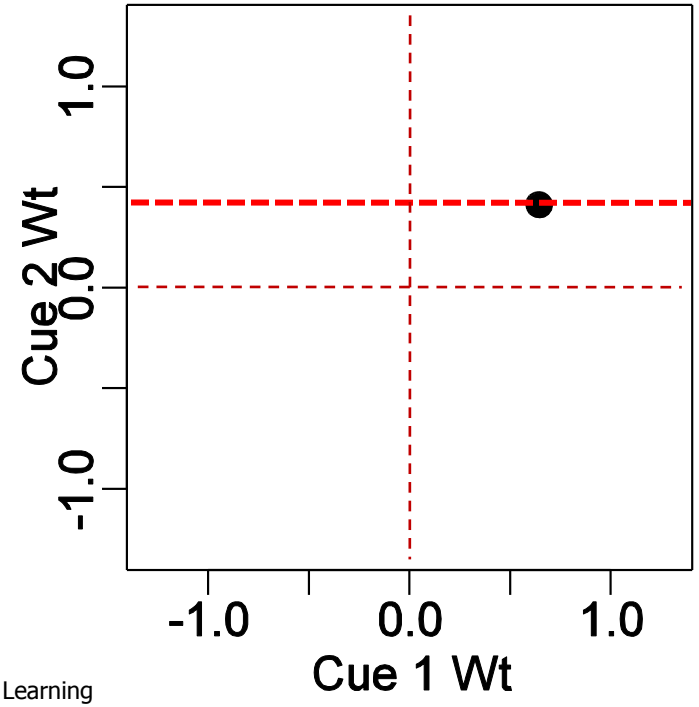


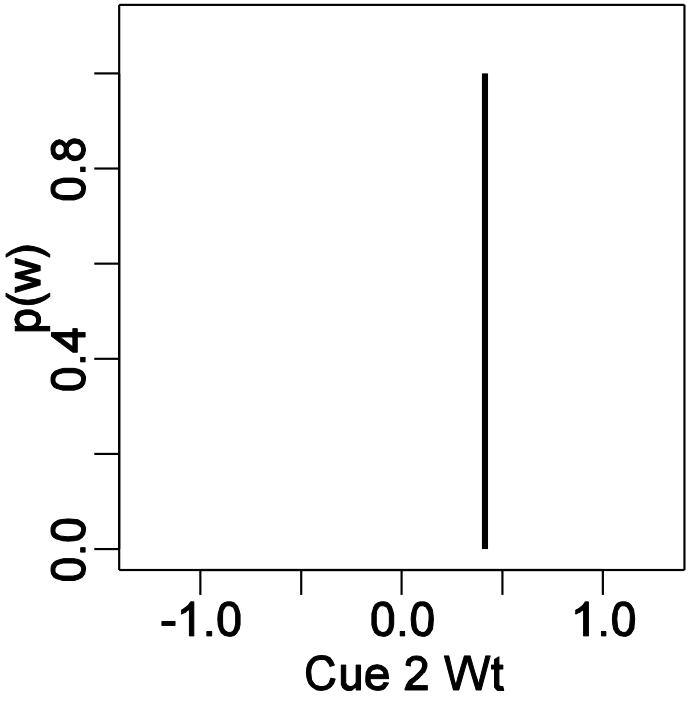
Training: BackwardBlocking

Beginning of Trial 17

Mean: 0.644 0.413



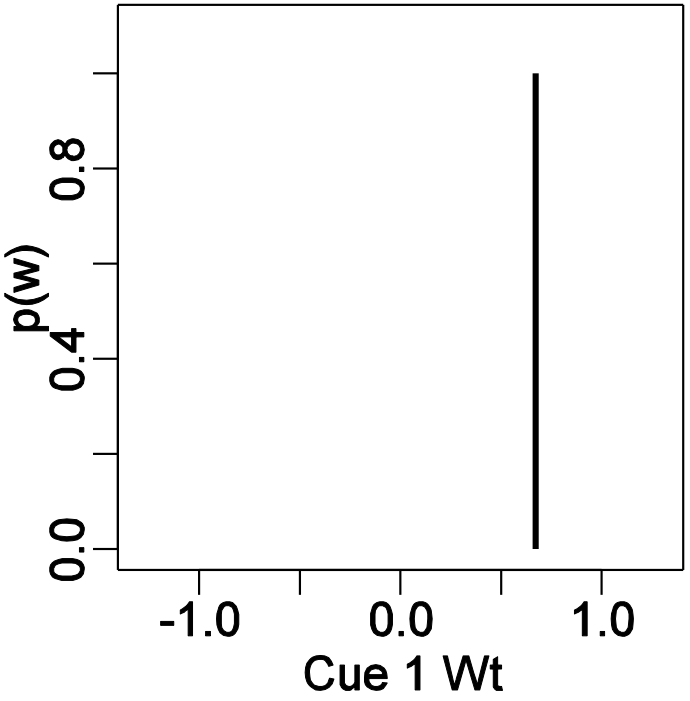


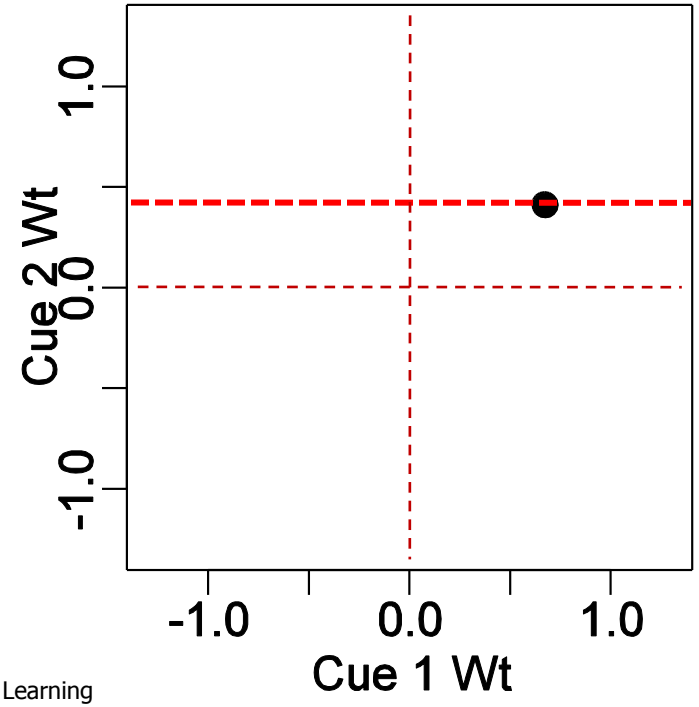


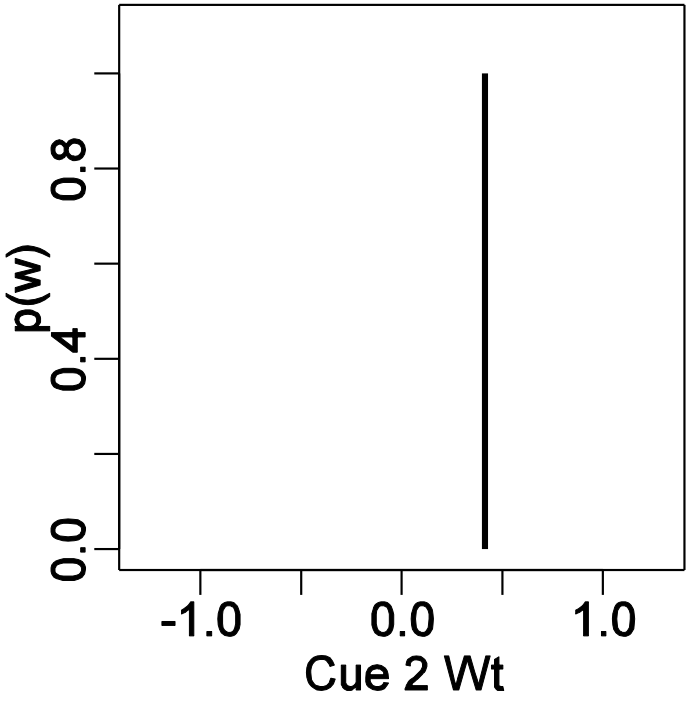
Training: BackwardBlocking

Beginning of Trial 18

Mean: 0.672 0.413



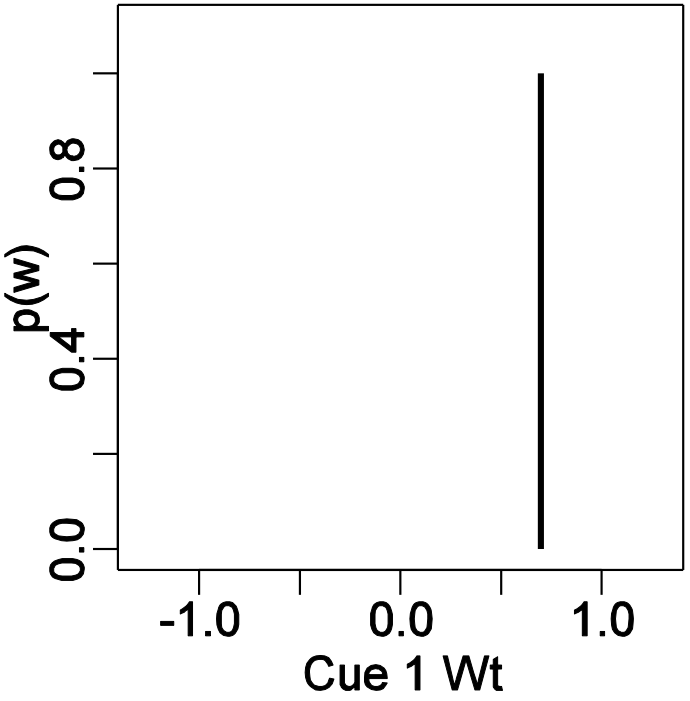


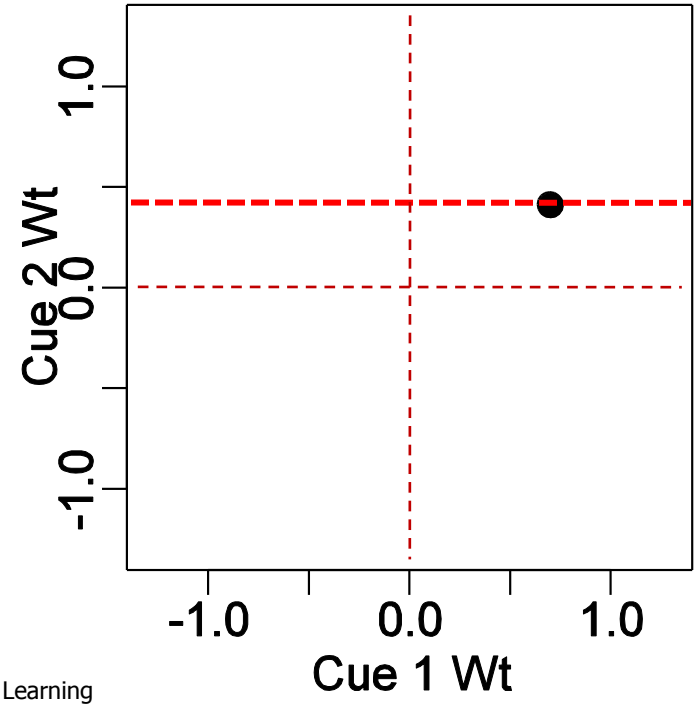


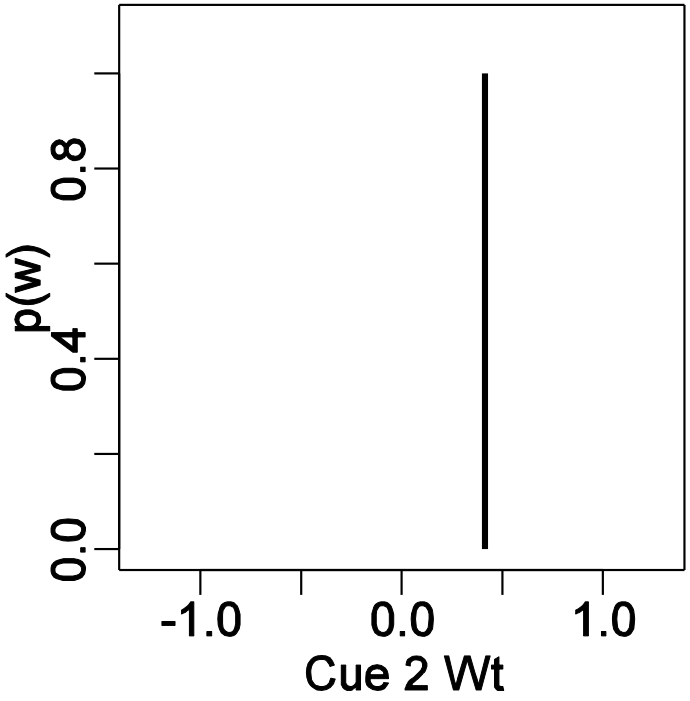
Training: BackwardBlocking

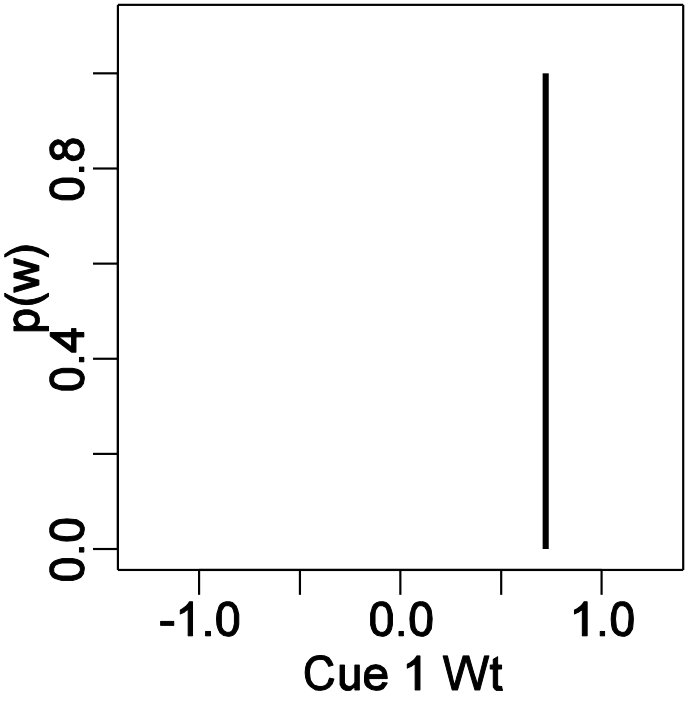
Beginning of Trial 19

Mean: 0.699 0.413





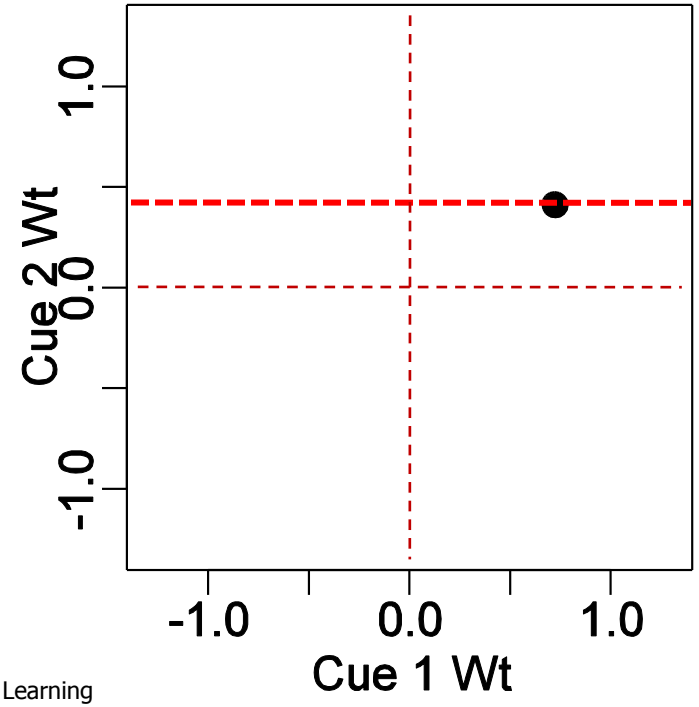


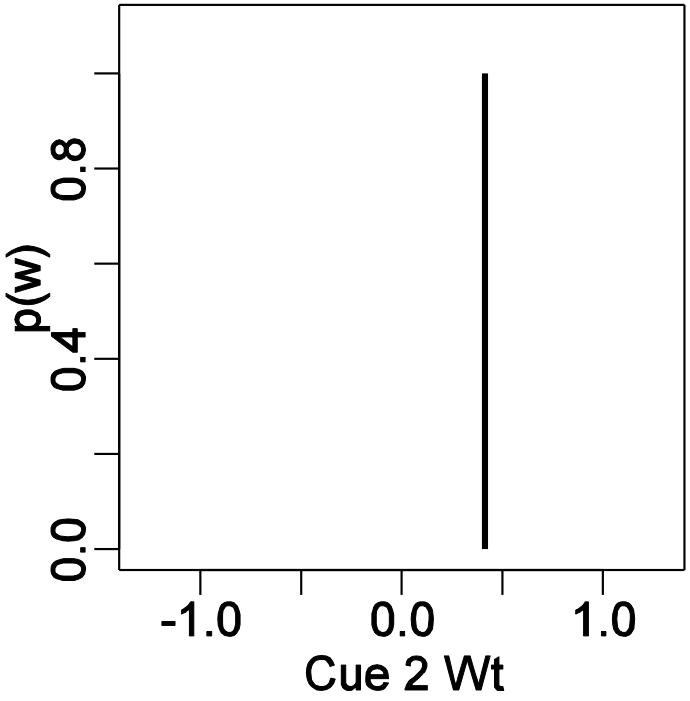


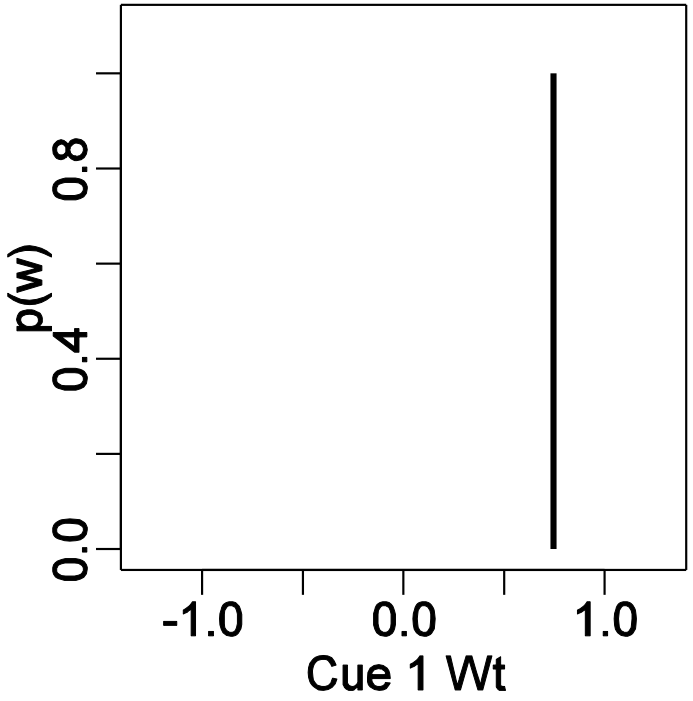
Training: BackwardBlocking

Beginning of Trial 20

Mean: 0.723 0.413



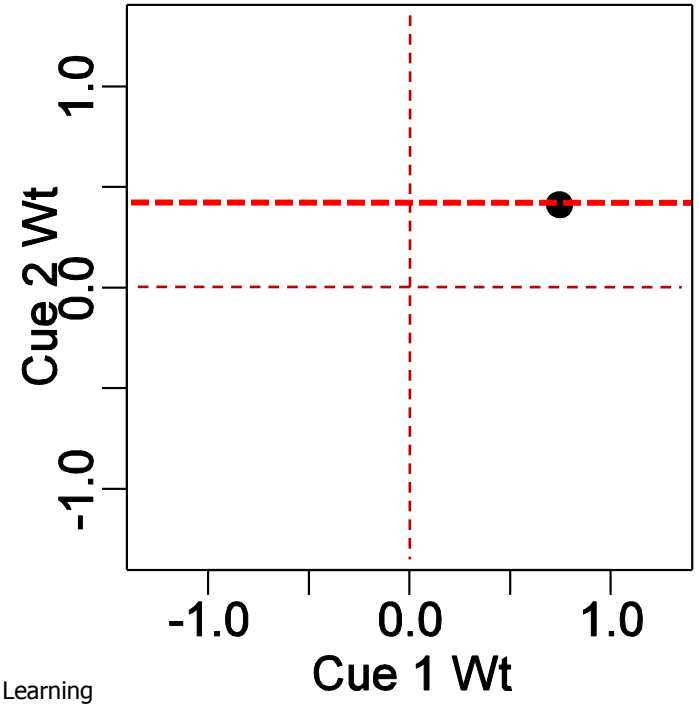


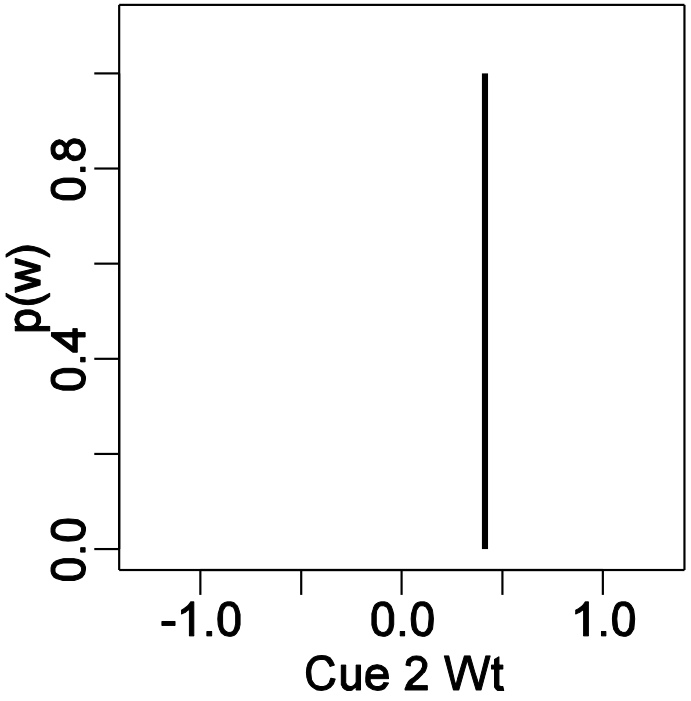


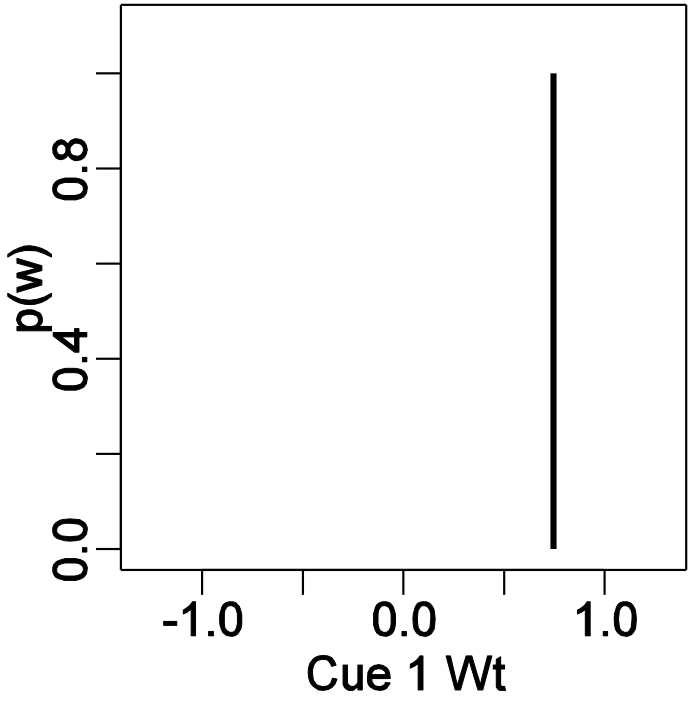
Training: BackwardBlocking

Beginning of Trial 21

Mean: 0.745 0.413





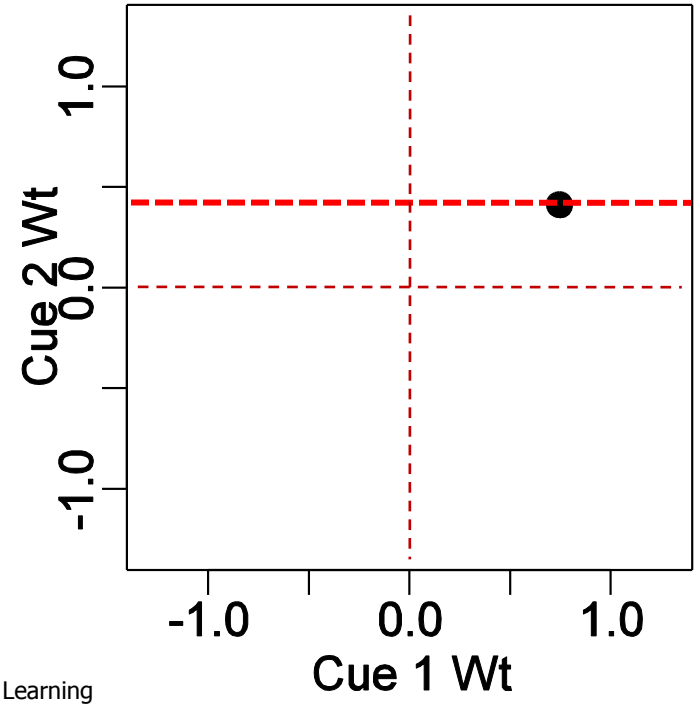


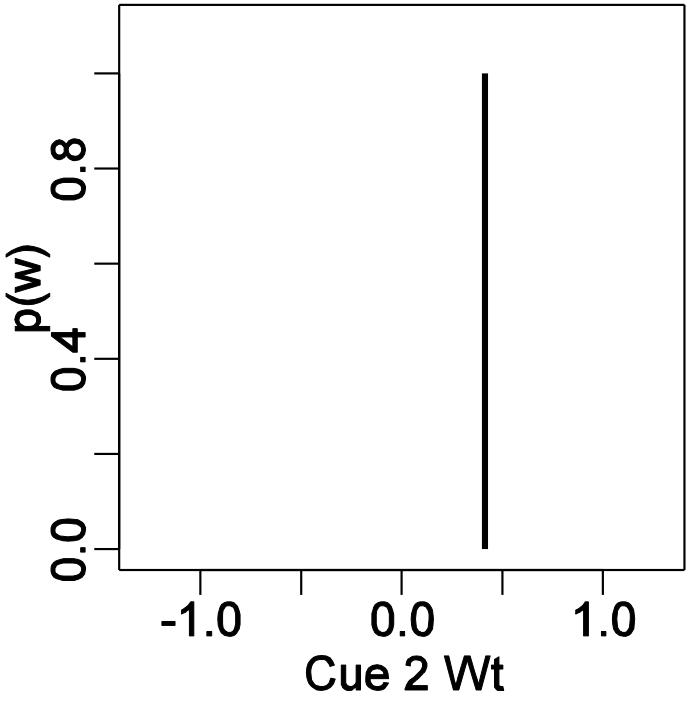
Training: BackwardBlocking

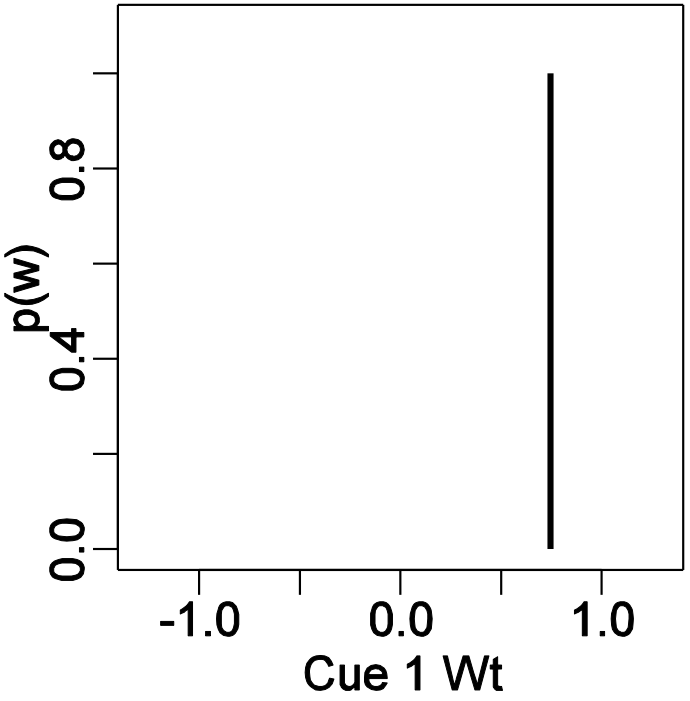
Beginning of Trial 21

Mean: 0.745 0.413

Rescorla-Wagner model does *not* show backward blocking.





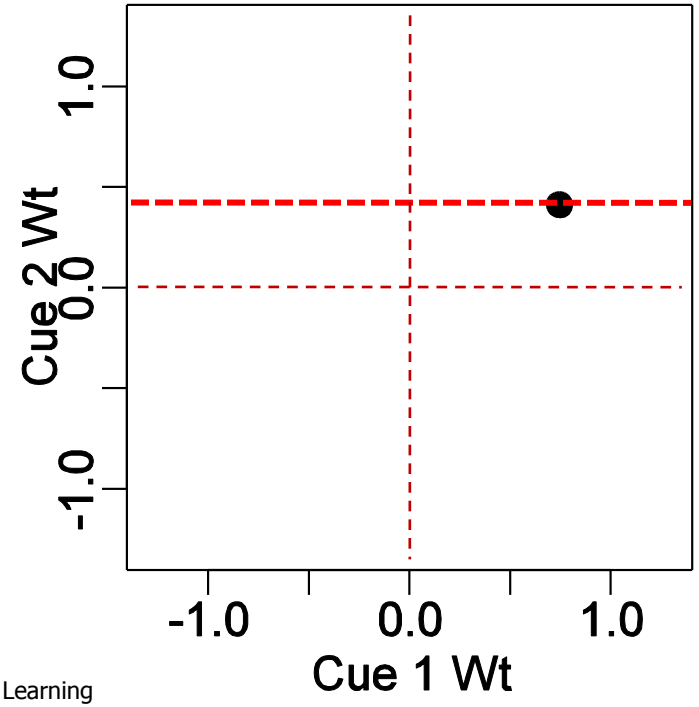


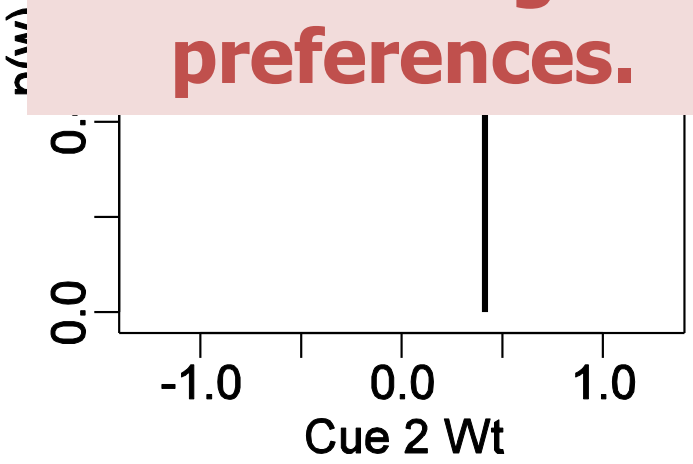
Training: BackwardBlocking

Beginning of Trial 21

Mean: 0.745 0.413

Rescorla-Wagner model does *not* show backward blocking.
And, has no way to predict active learning preferences.





Design of the (rest of the) Talk:
2 x 2 x 3 factorial

Bayesian models of associative learning:

- Kalman filter.
- Noisy-logic gate.

Goals for active learning:

- Minimize expected uncertainty.
- Maximize expected maximum believability.

Training structures:

- Backward blocking.
- Blocking and reduced overshadowing.
- Ambiguous cue.

The Bayesian Ontology

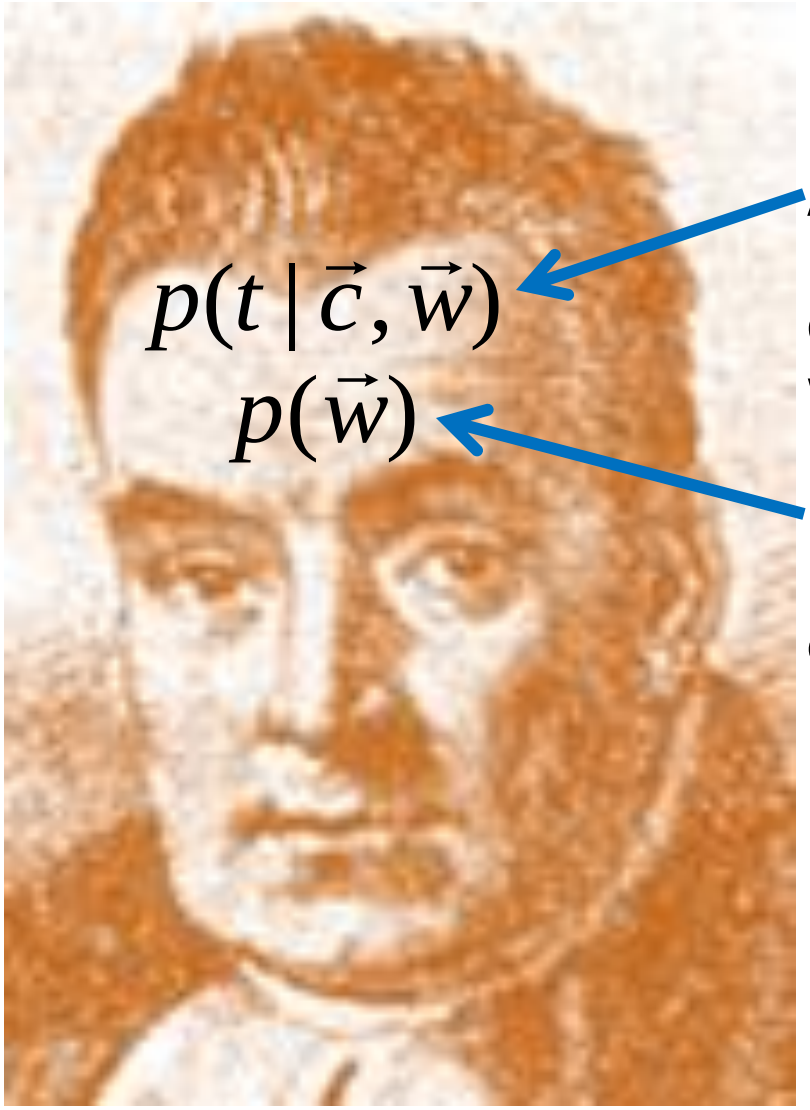
Multiple hypotheses entertained simultaneously. Not just a single state as in traditional models.

Graded *degree of belief* in each hypothesis.

Hypotheses specify the *probability* of each possible stimulus/outcome value. Not deterministic as in traditional models.

Learning is *shifting of beliefs* from less likely to more likely hypotheses. Bayes' rule specifies how.

Bayesian Learning In the Mind

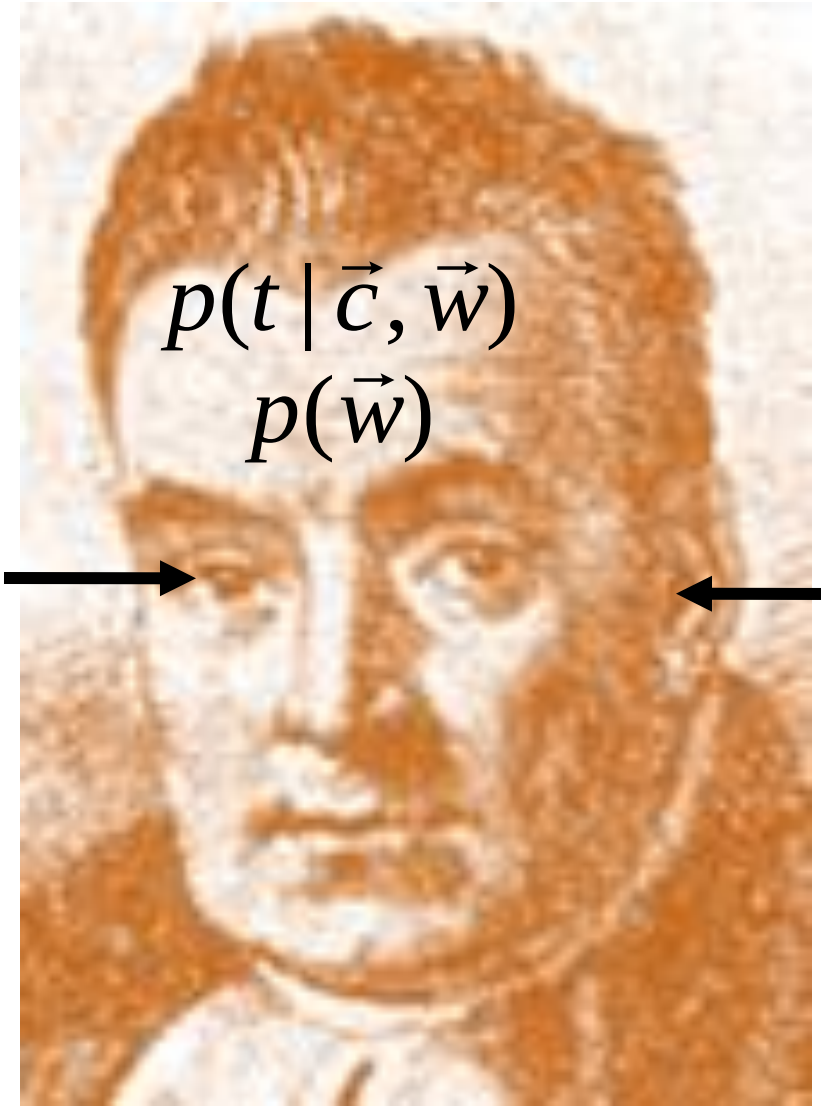


Start with:

A model that specifies the probability of outcome value t, given cues c and hypothetical weights w.

Prior believabilities of every possible hypothetical weight combination w.

Bayesian Learning In the Mind

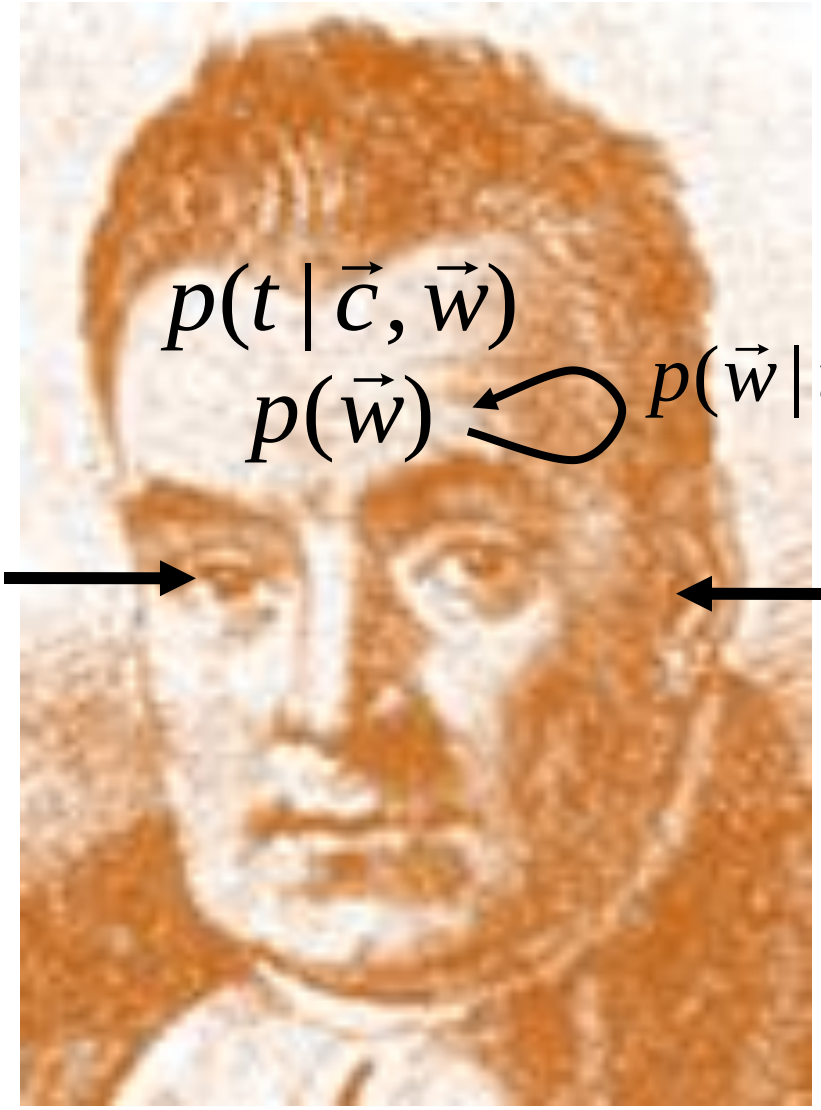


Observe:

Cues c.

Actual outcome t.

Bayesian Learning In the Mind

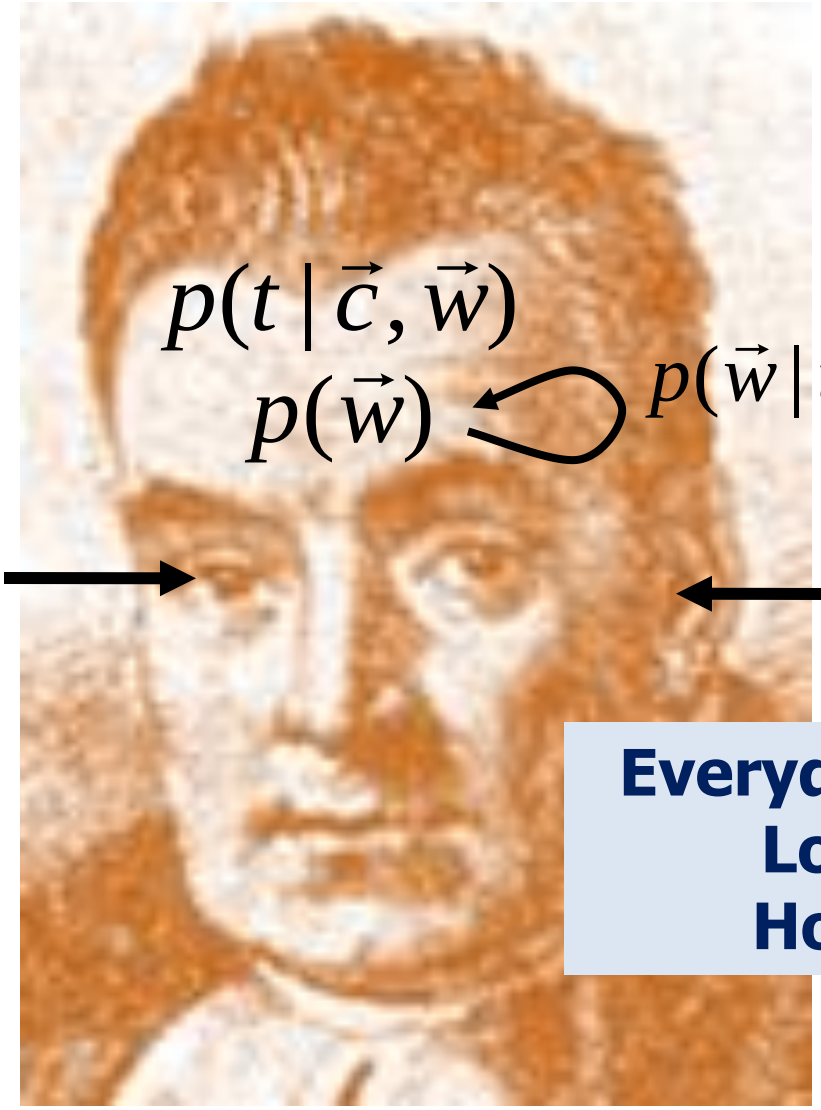


Learn:

Update beliefs by Bayes' rule.

$$p(\vec{w} \mid t, \vec{c}) = \frac{p(t \mid \vec{c}, \vec{w})\ p(\vec{w})}{\int d\vec{w}\ p(t \mid \vec{c}, \vec{w})\ p(\vec{w})}$$

Bayesian Learning In the Mind



Learn:

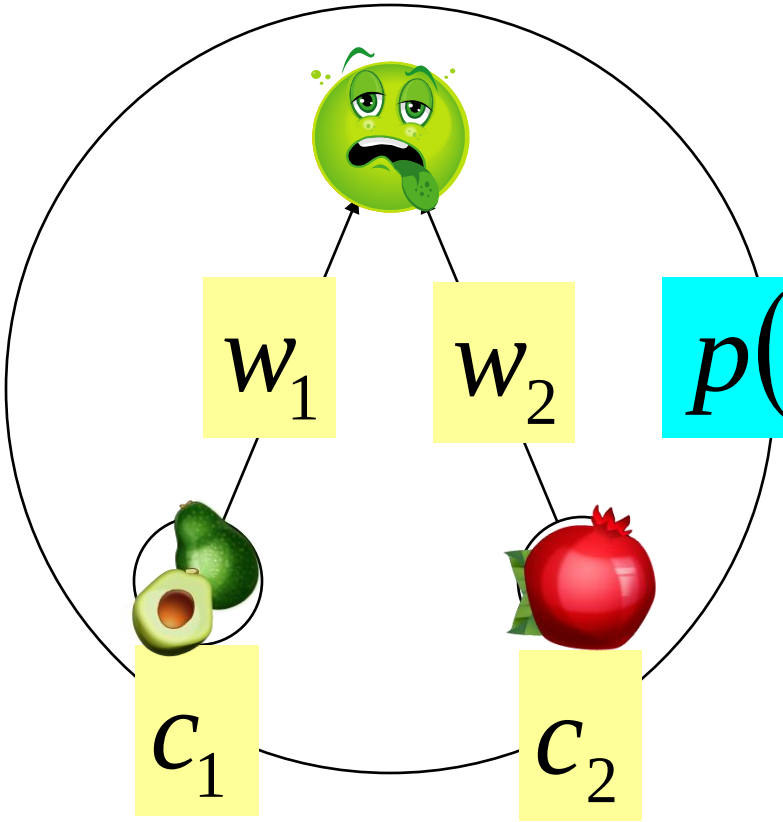
Update beliefs by Bayes' rule.

$$p(\vec{w} \mid t, \vec{c}) = \frac{p(t \mid \vec{c}, \vec{w})\ p(\vec{w})}{\int d\vec{w}\ p(t \mid \vec{c}, \vec{w})\ p(\vec{w})}$$

Everyday Bayesian reasoning:
Logic of exoneration.
Holmesian deduction.

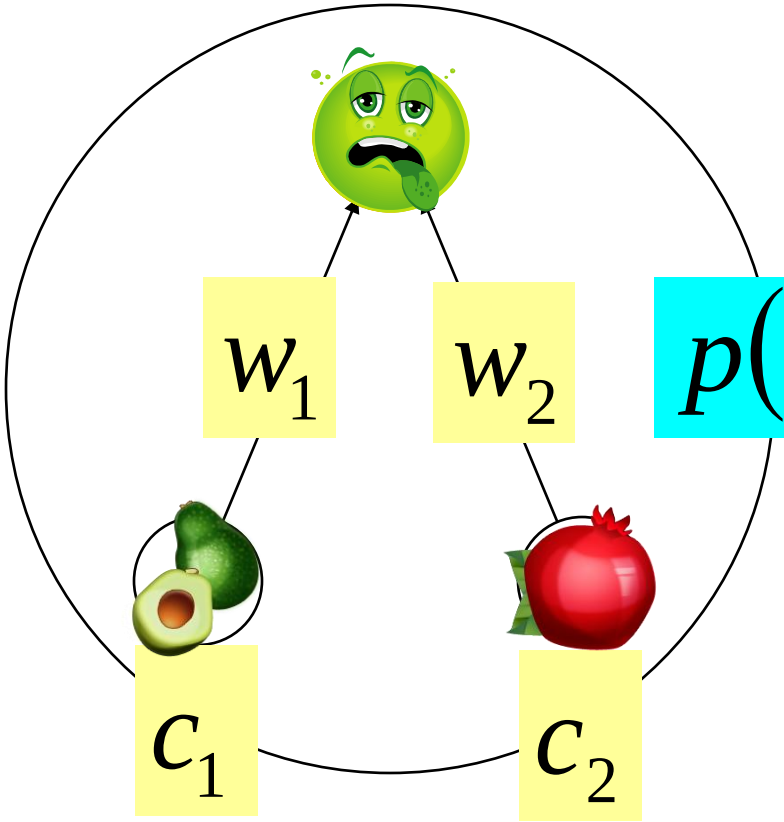
Kalman Filter

$$p(t \mid \vec{c}, \vec{w}) \sim N(t \mid \vec{w}^T \vec{c}, v)$$



$$p(\vec{w}) \sim N(\vec{w} \mid \vec{\mu}, C)$$

Kalman Filter Updating: Step 1. Linear Dynamics

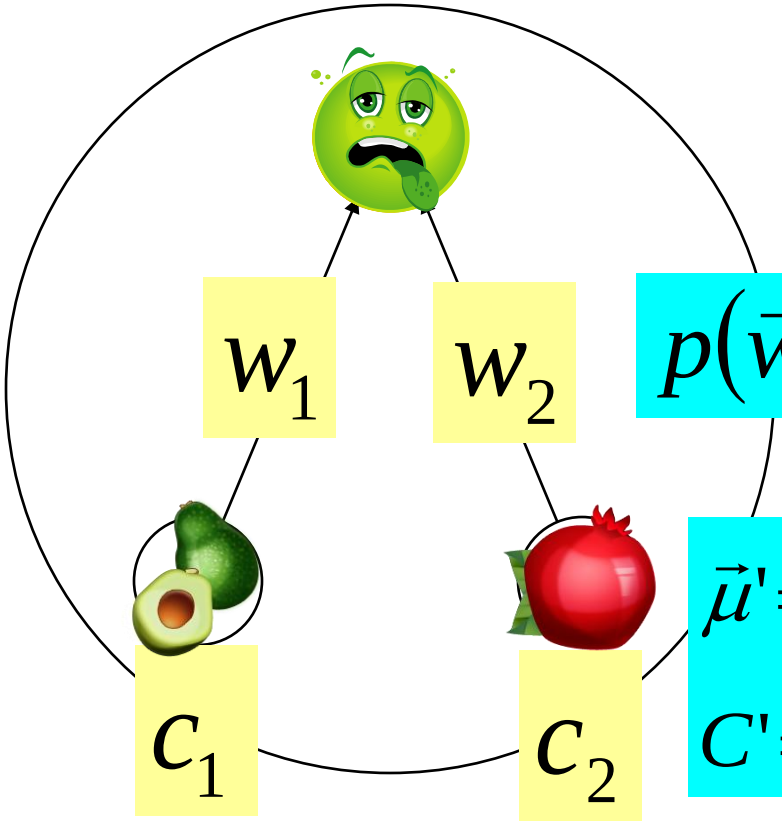


$$p(\vec{w}) \sim N(\vec{w} \mid \vec{\mu}, C)$$

$$\vec{\mu}^* = D\vec{\mu}$$

$$C^* = DCD^T + U$$

Kalman Filter Updating:
Step 2. **Bayesian Learning**



$$p(\vec{w}) \sim N(\vec{w} \mid \vec{\mu}^*, C^*)$$

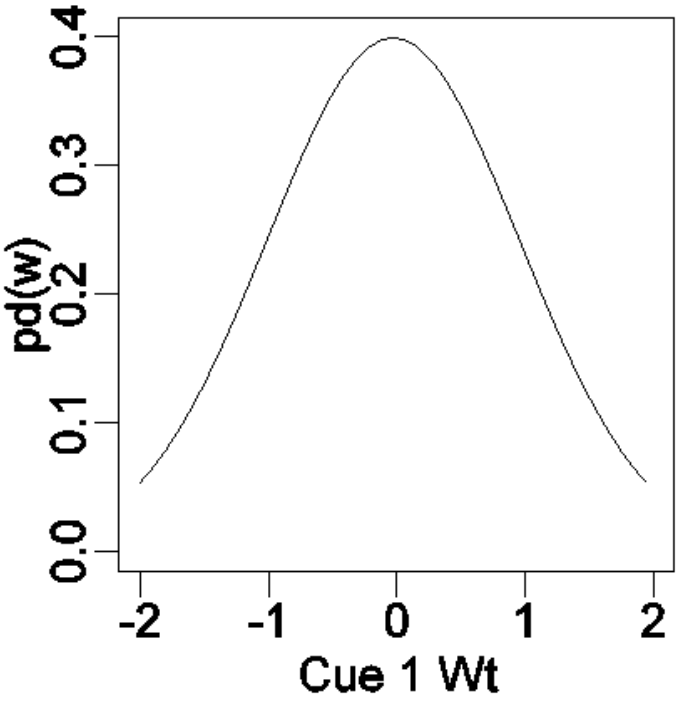
$$\vec{\mu}' = \vec{\mu}^* + \left[v + \vec{c}^T C^* \vec{c}\right]^{-1} \left(t - \vec{\mu}^{*T} \vec{c}\right) C^* \vec{c}$$

$$C' = C^* - \left[v + \vec{c}^T C^* \vec{c}\right]^{-1} C^* \vec{c} \vec{c}^T C^*$$

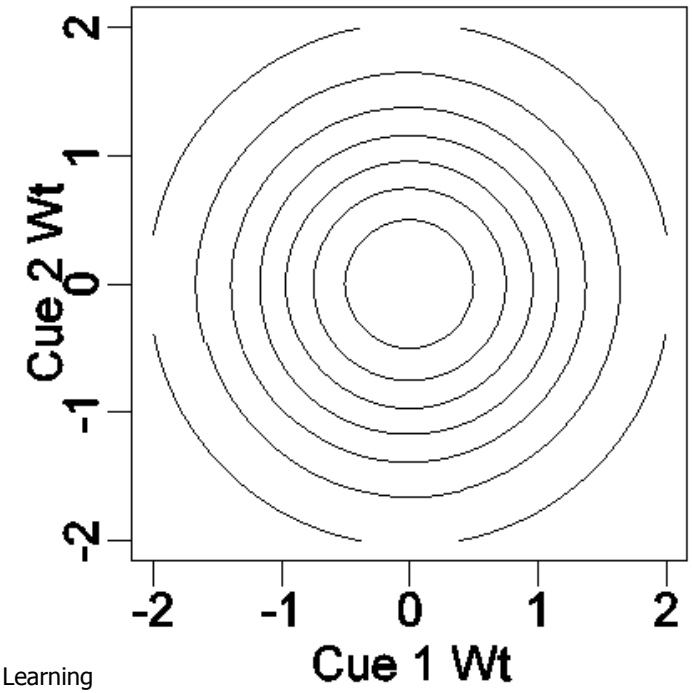
Backward Blocking

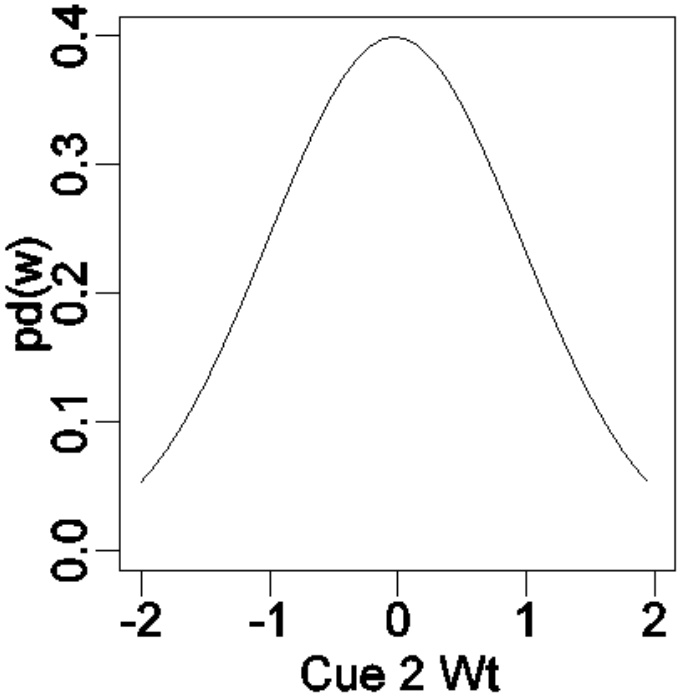
Phase	Frequency	Cue 1	Cue 2	Outcome
I	10	1 	1 	1 
II	10	1 	0	1 

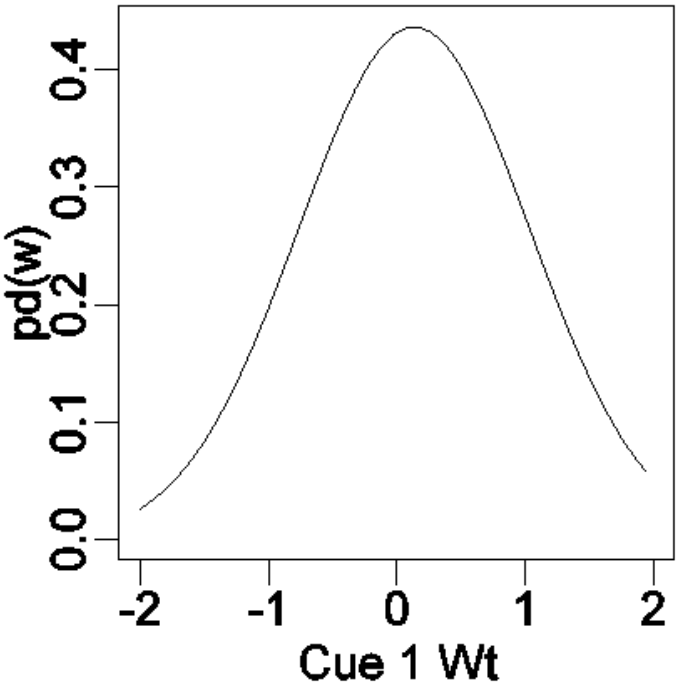
Note: Cell with a 1 indicates presence of cue or outcome. Cell with a 0 indicates absence of cue or outcome.



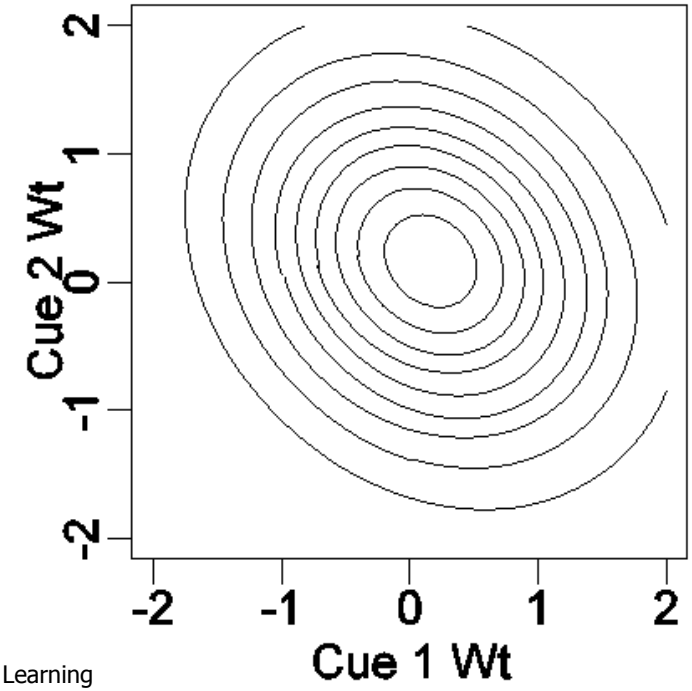
Train: BackwardBlocking
Beginning of Trial 1
Mean: 0 0
Covariance matrix:
1 0
0 1
Current Uncertainty: 2.838
Probe: 1 0 => EU: 2.726
Probe: 0 1 => EU: 2.726

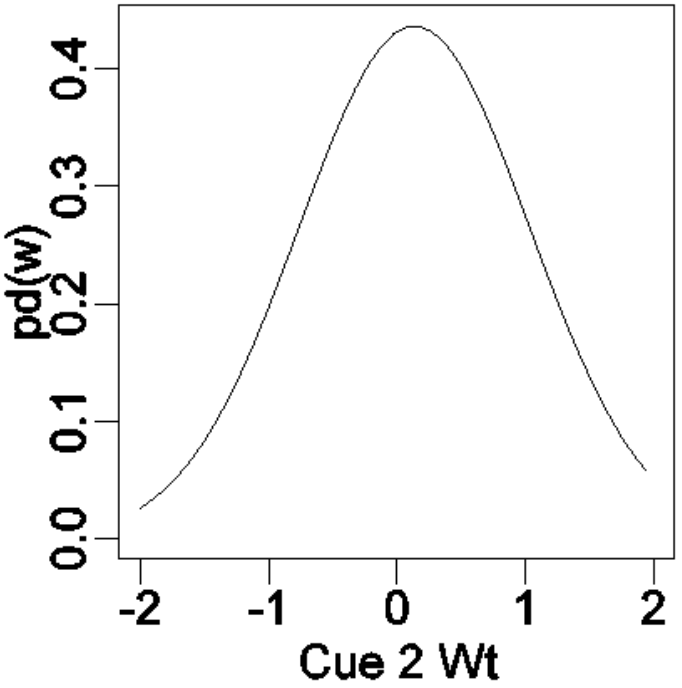


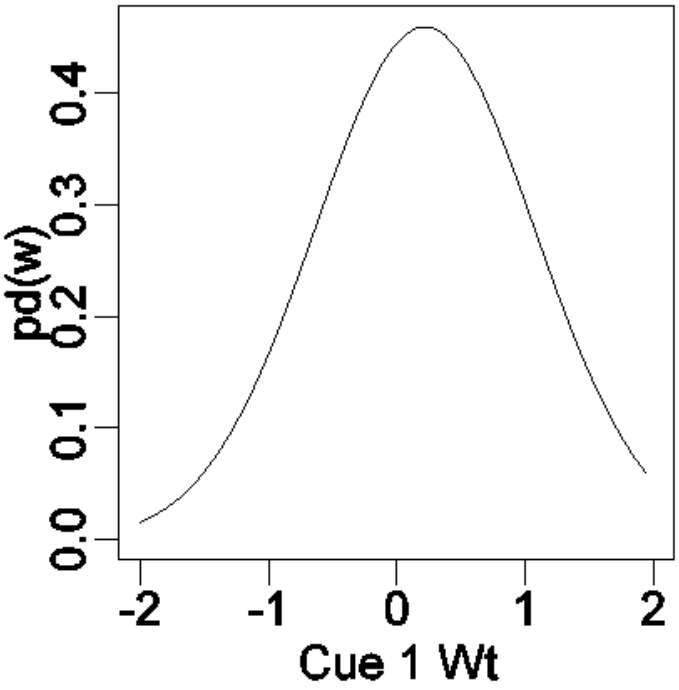




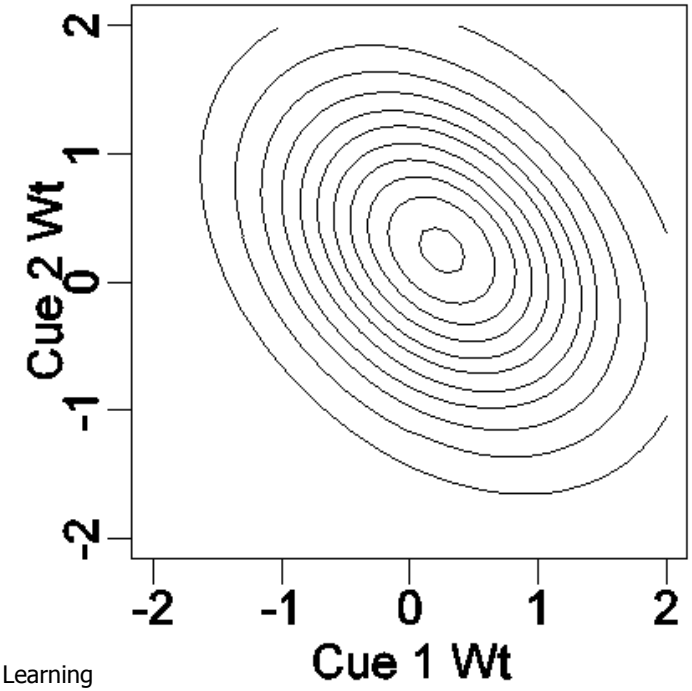
Train: BackwardBlocking
Beginning of Trial 2
Mean: 0.167 0.167
Covariance matrix:
0.833 -0.167
-0.167 0.833
Current Uncertainty: 2.635
Probe: 1 0 => EU: 2.541
Probe: 0 1 => EU: 2.541

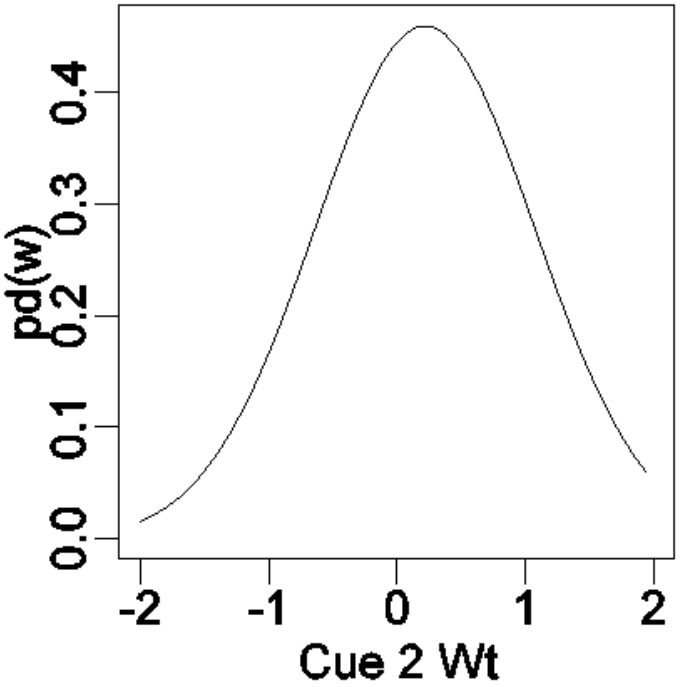


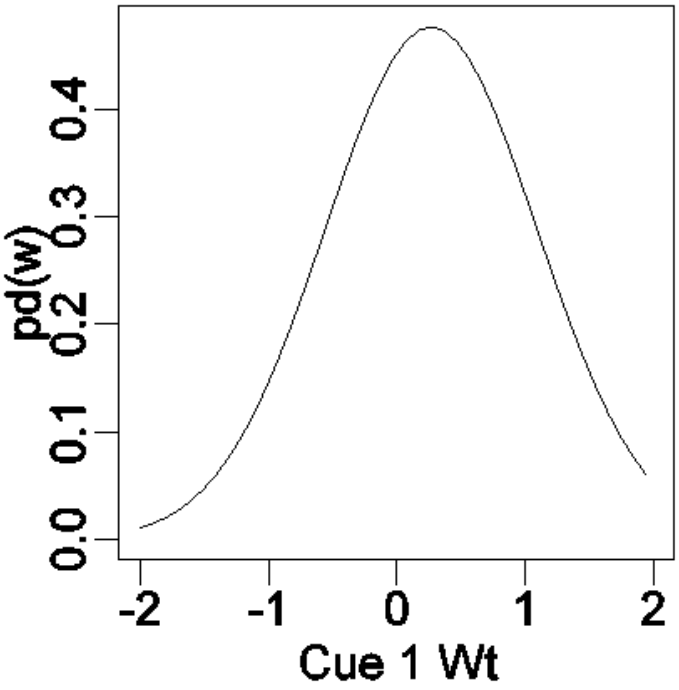




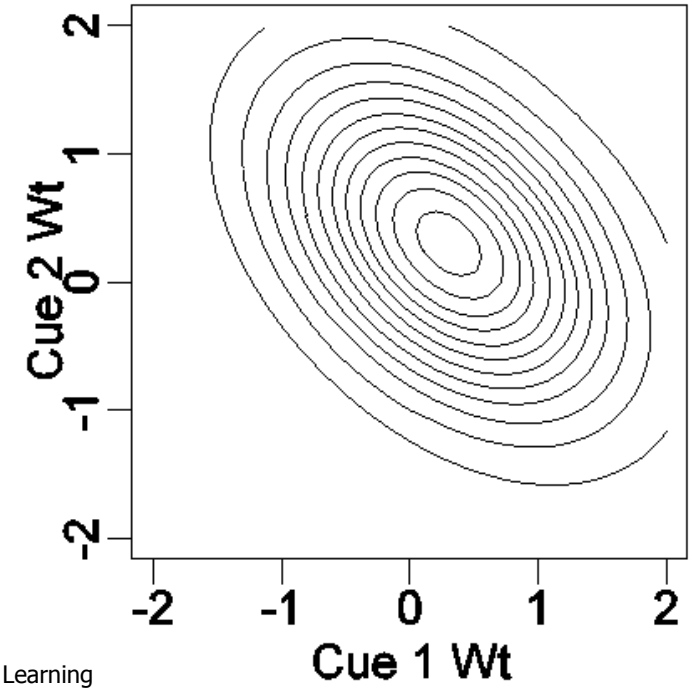
Train: BackwardBlocking
Beginning of Trial 3
Mean: 0.25 0.25
Covariance matrix:
0.75 -0.25
-0.25 0.75
Current Uncertainty: 2.491
Probe: 1 0 => EU: 2.405
Probe: 0 1 => EU: 2.405

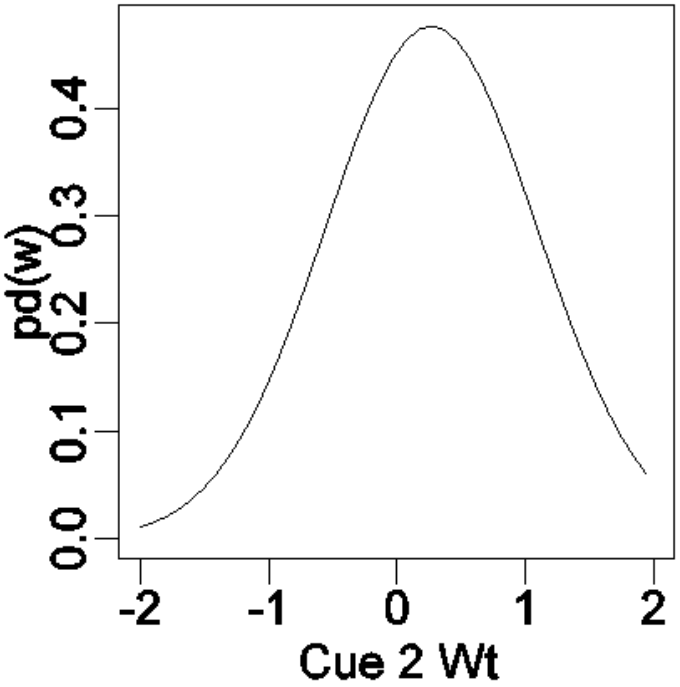


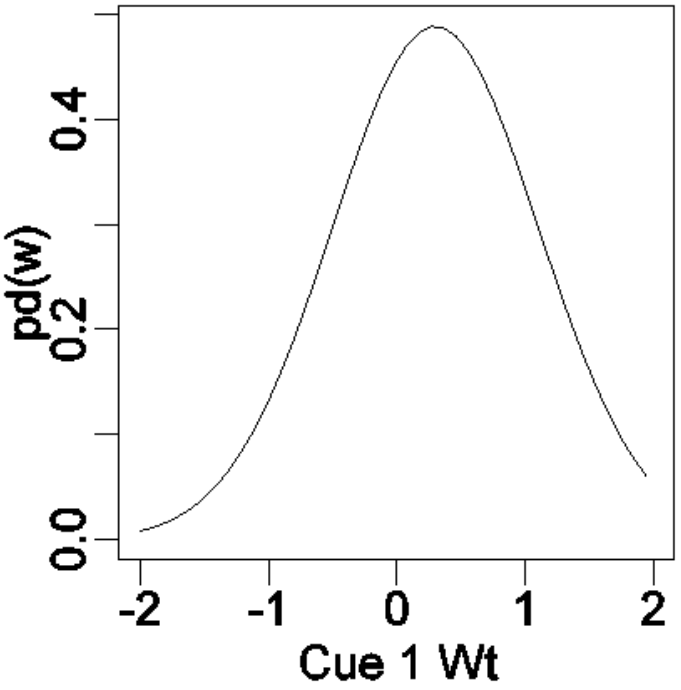




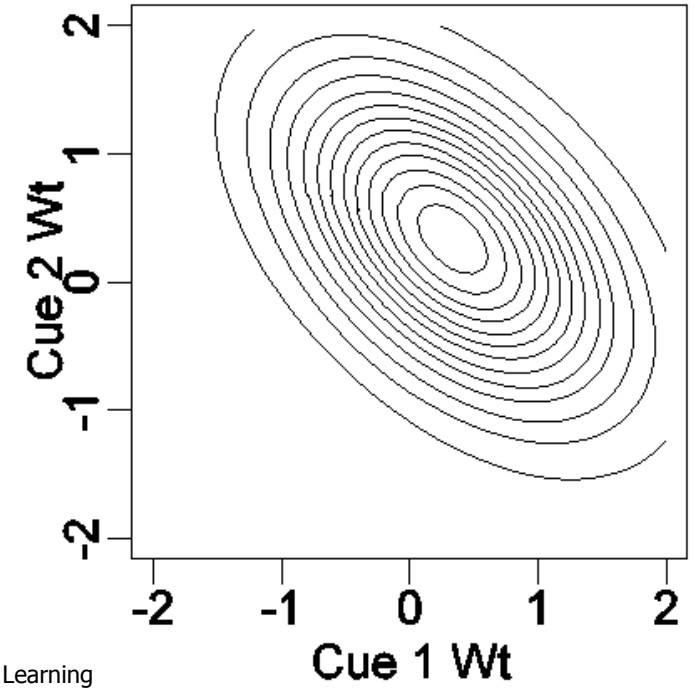
Train: BackwardBlocking
Beginning of Trial 4
Mean: 0.3 0.3
Covariance matrix:
0.7 -0.3
-0.3 0.7
Current Uncertainty: 2.38
Probe: 1 0 => EU: 2.299
Probe: 0 1 => EU: 2.299

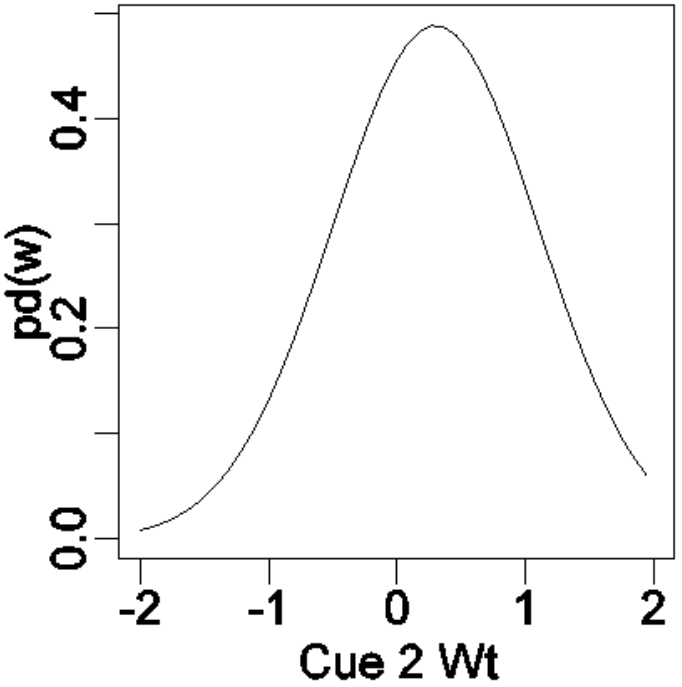


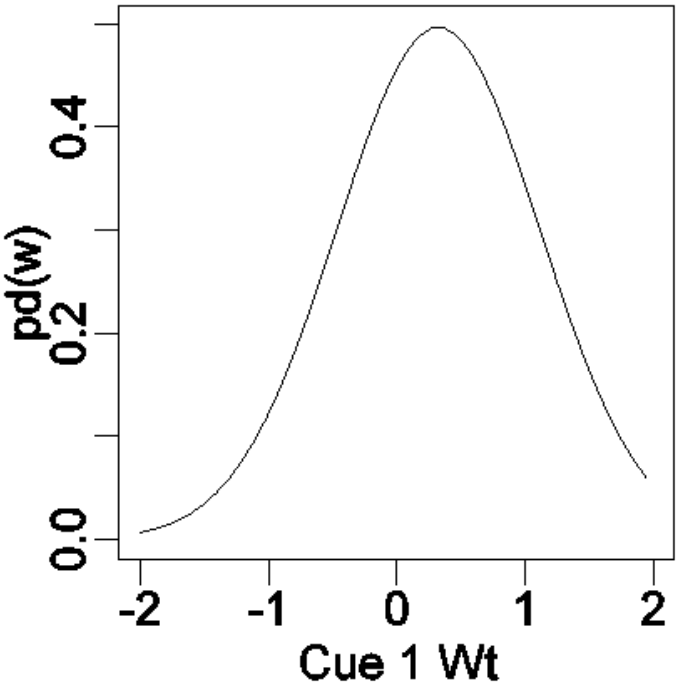




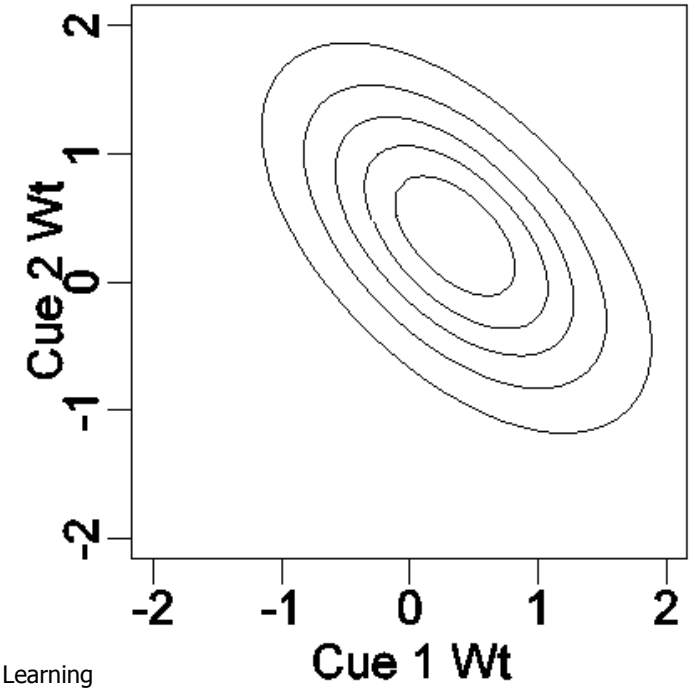
Train: BackwardBlocking
Beginning of Trial 5
Mean: 0.333 0.333
Covariance matrix:
0.667 -0.333
-0.333 0.667
Current Uncertainty: 2.289
Probe: 1 0 => EU: 2.211
Probe: 0 1 => EU: 2.211

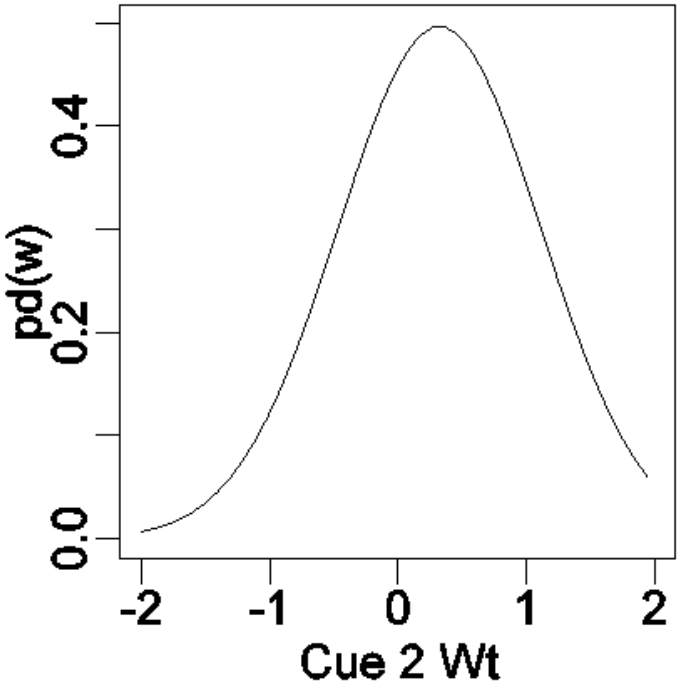


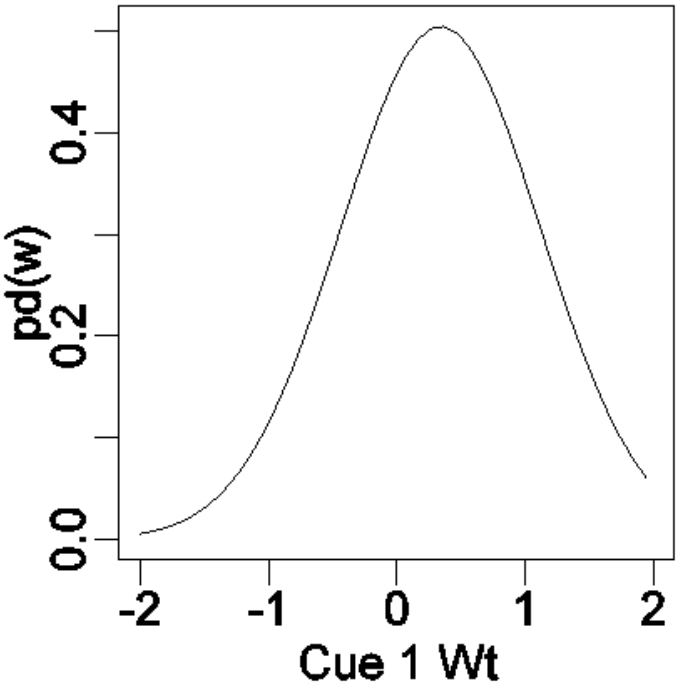




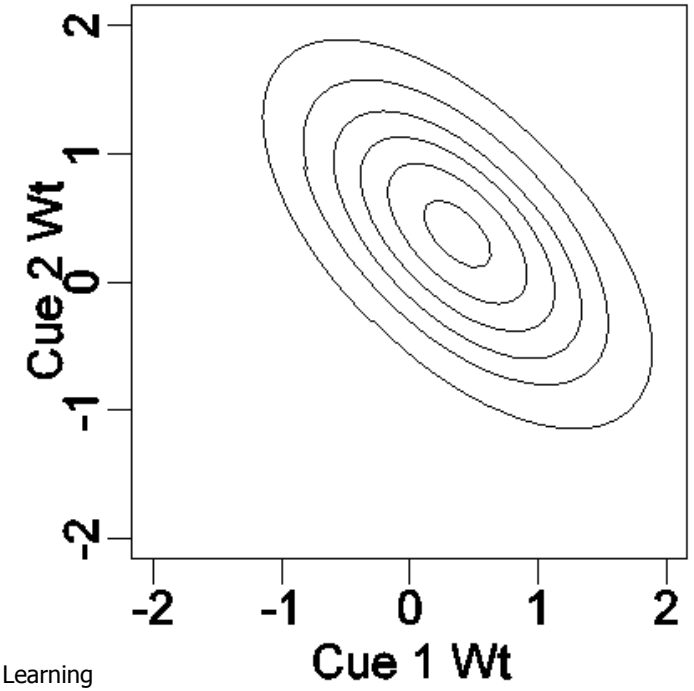
Train: BackwardBlocking
Beginning of Trial 6
Mean: 0.357 0.357
Covariance matrix:
0.643 -0.357
-0.357 0.643
Current Uncertainty: 2.211
Probe: 1 0 => EU: 2.137
Probe: 0 1 => EU: 2.137

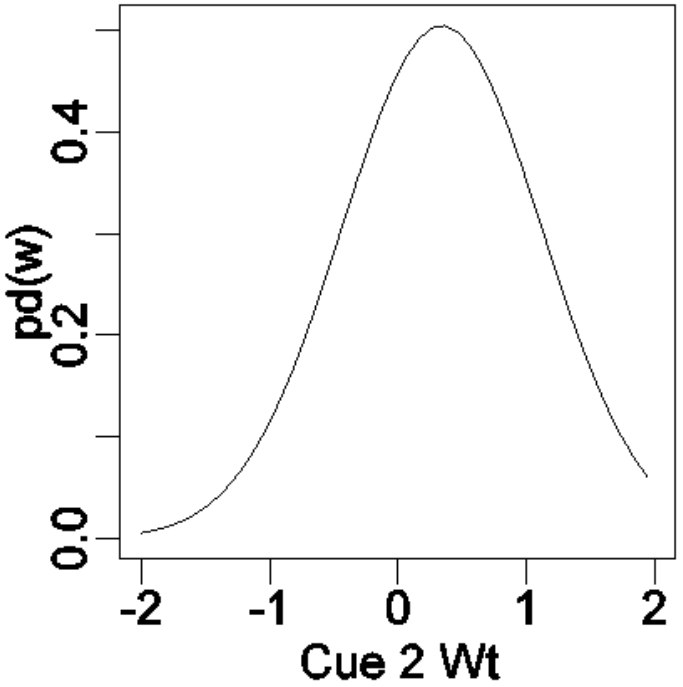


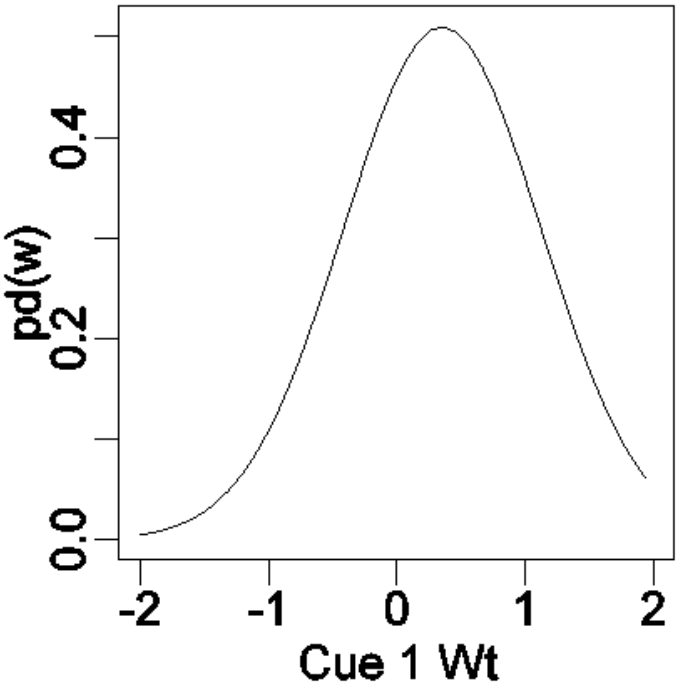




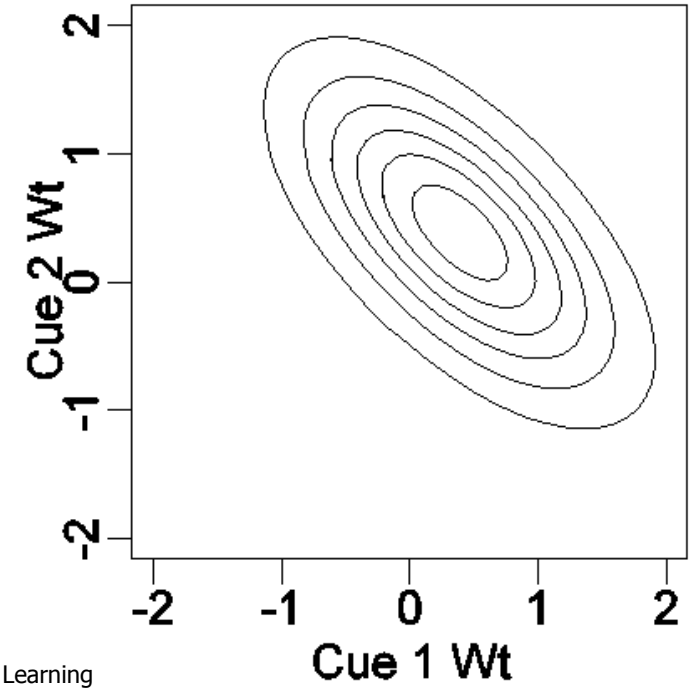
Train: BackwardBlocking
Beginning of Trial 7
Mean: 0.375 0.375
Covariance matrix:
0.625 -0.375
-0.375 0.625
Current Uncertainty: 2.145
Probe: 1 0 => EU: 2.072
Probe: 0 1 => EU: 2.072

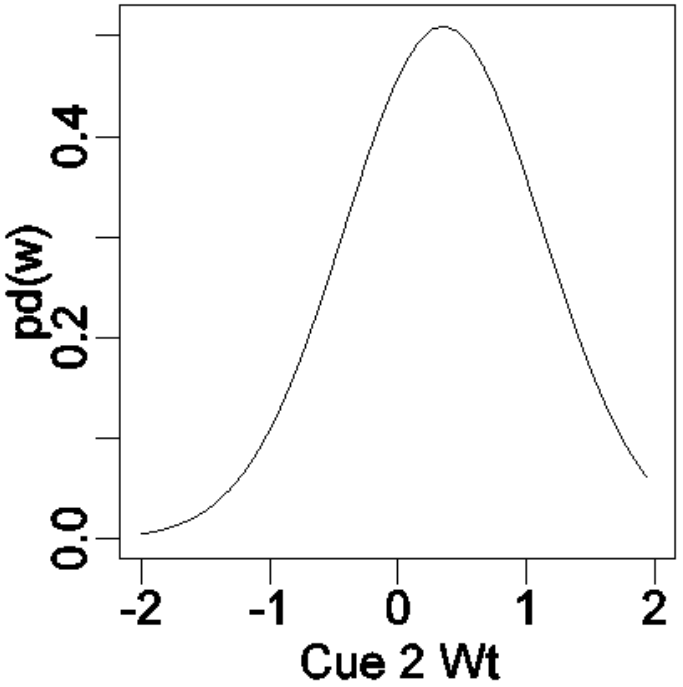


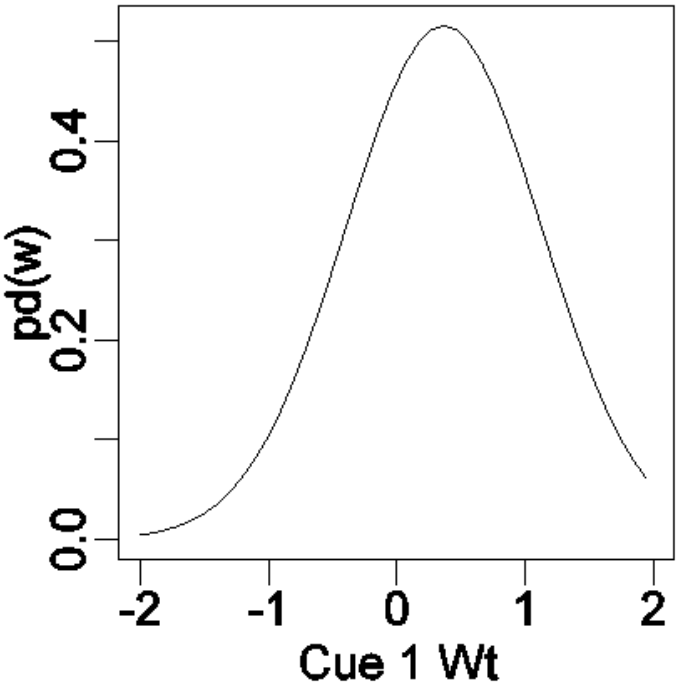




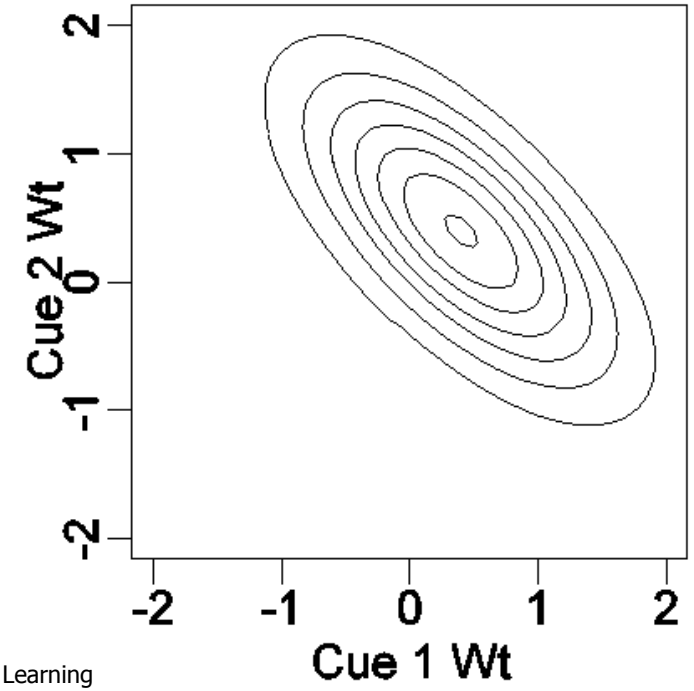
Train: BackwardBlocking
Beginning of Trial 8
Mean: 0.389 0.389
Covariance matrix:
0.611 -0.389
-0.389 0.611
Current Uncertainty: 2.086
Probe: 1 0 => EU: 2.015
Probe: 0 1 => EU: 2.015

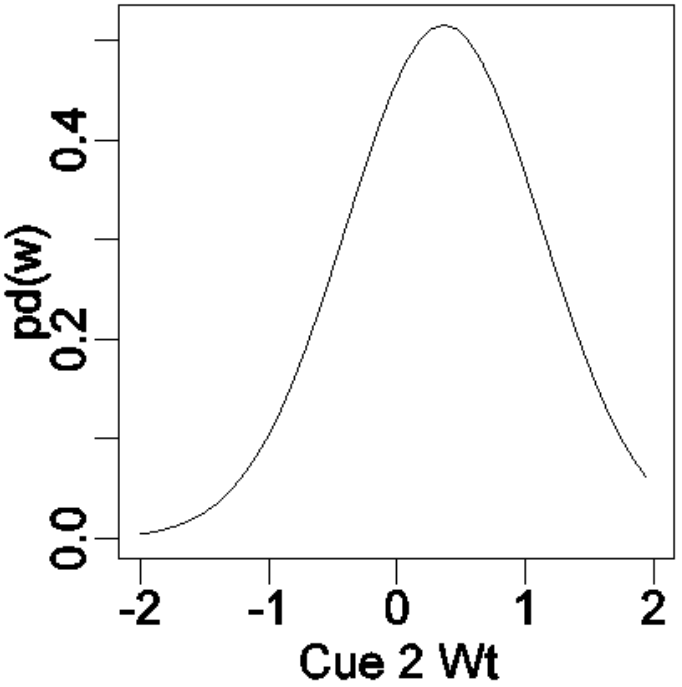


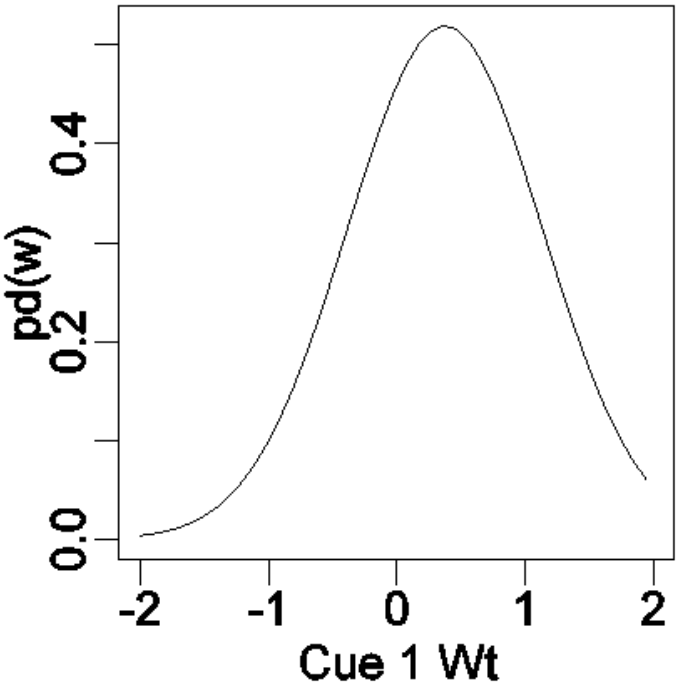




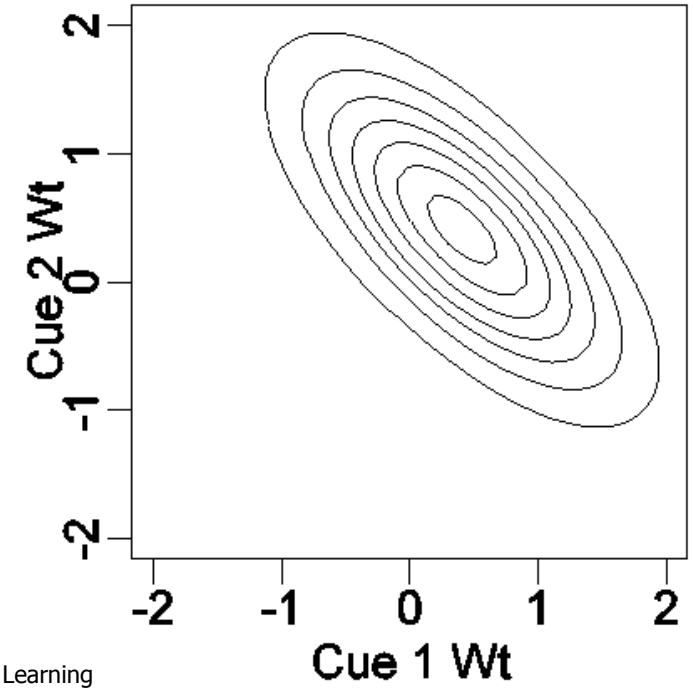
Train: BackwardBlocking
Beginning of Trial 9
Mean: 0.4 0.4
Covariance matrix:
0.6 -0.4
-0.4 0.6
Current Uncertainty: 2.033
Probe: 1 0 => EU: 1.963
Probe: 0 1 => EU: 1.963

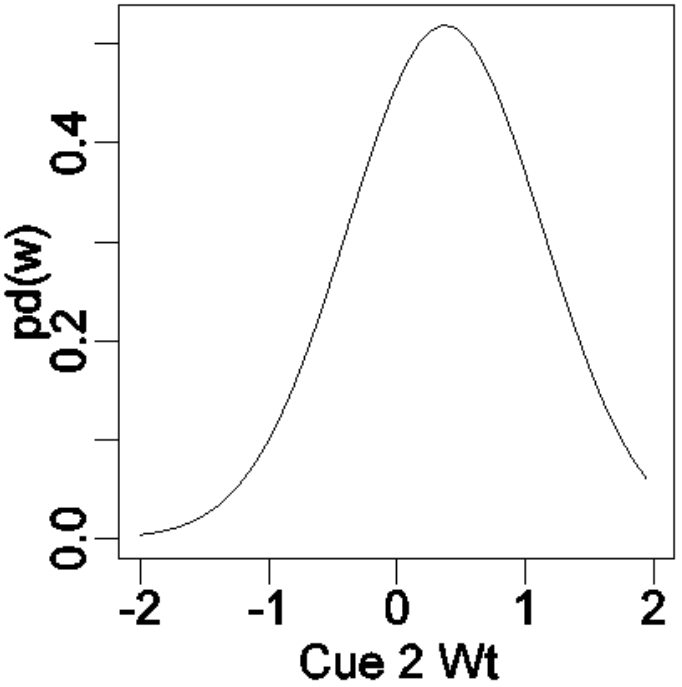


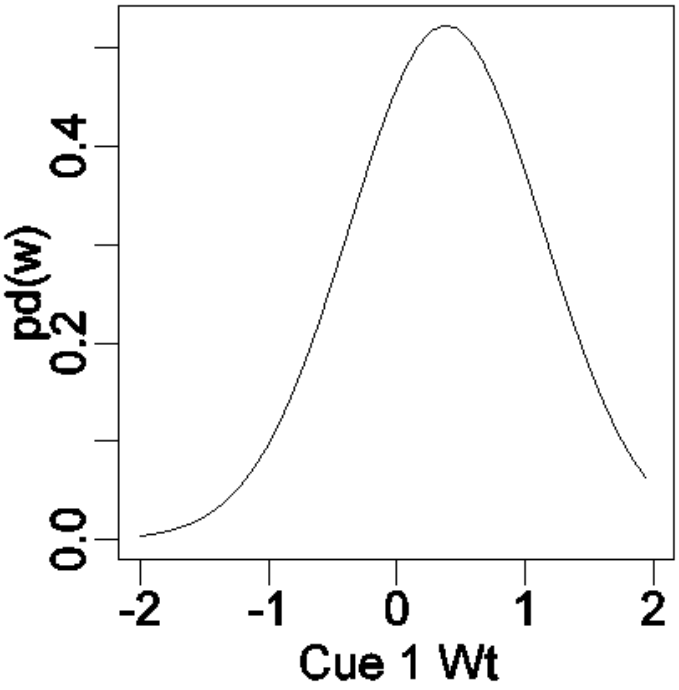




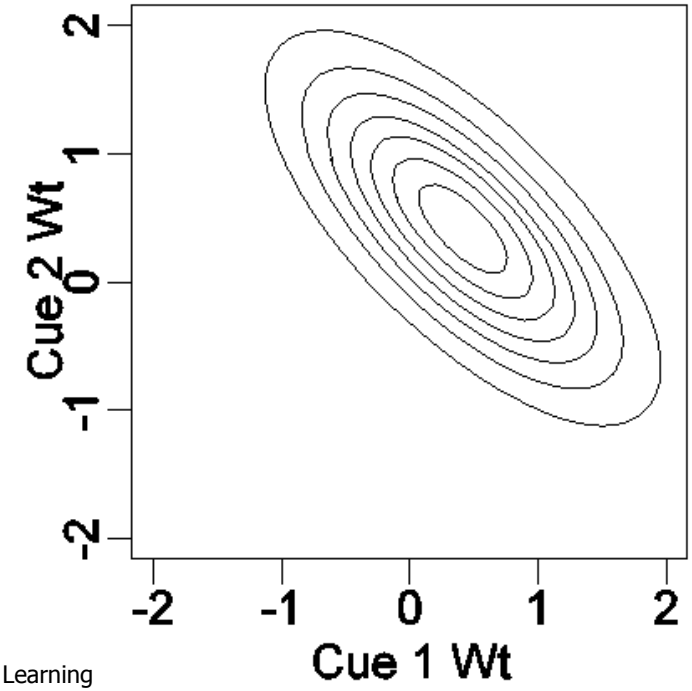
Train: BackwardBlocking
Beginning of Trial 10
Mean: 0.409 0.409
Covariance matrix:
0.591 -0.409
-0.409 0.591
Current Uncertainty: 1.986
Probe: 1 0 => EU: 1.917
Probe: 0 1 => EU: 1.917

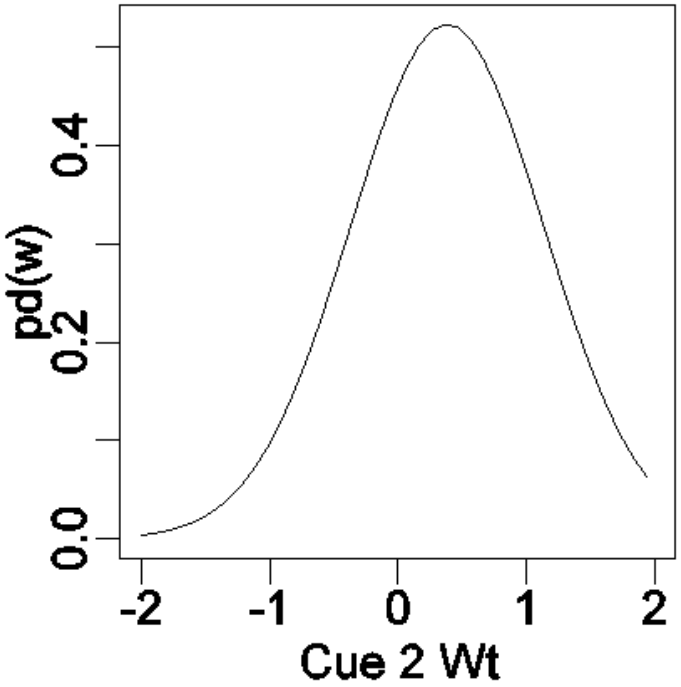


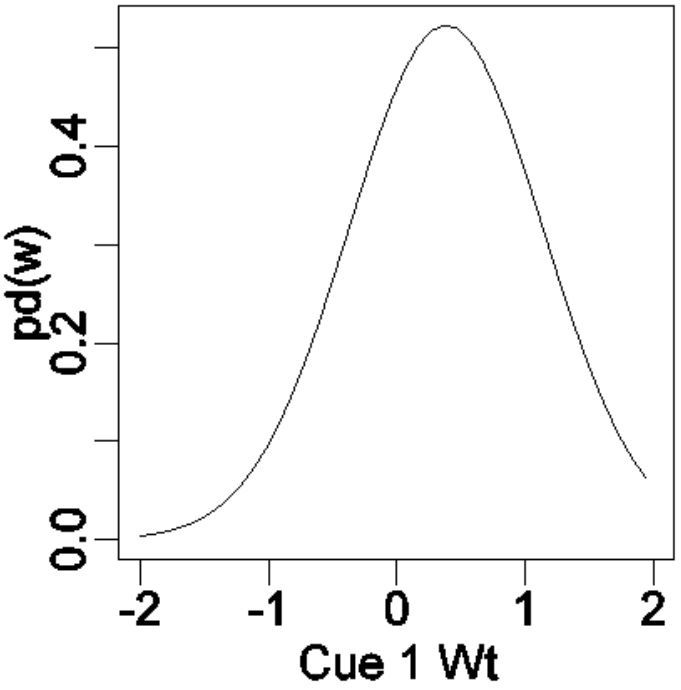




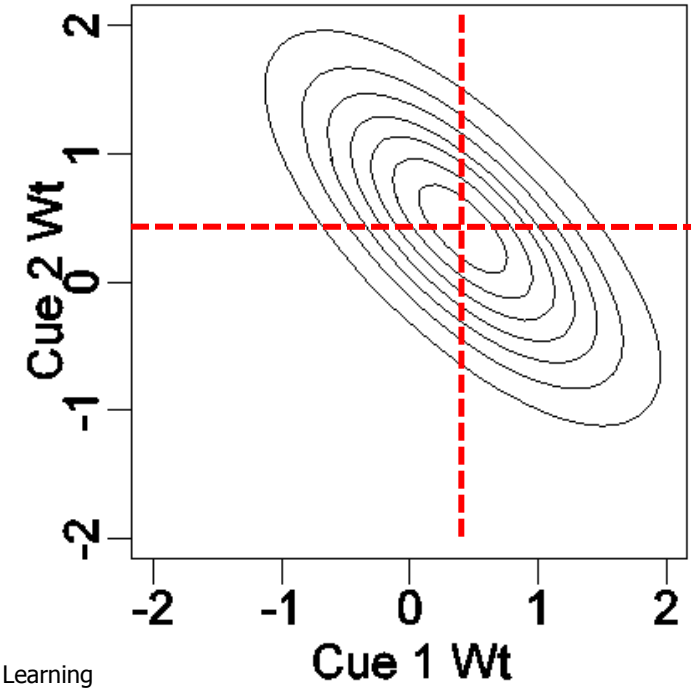
Train: BackwardBlocking
Beginning of Trial 11
Mean: 0.417 0.417
Covariance matrix:
0.583 -0.417
-0.417 0.583
Current Uncertainty: 1.942
Probe: 1 0 => EU: 1.874
Probe: 0 1 => EU: 1.874

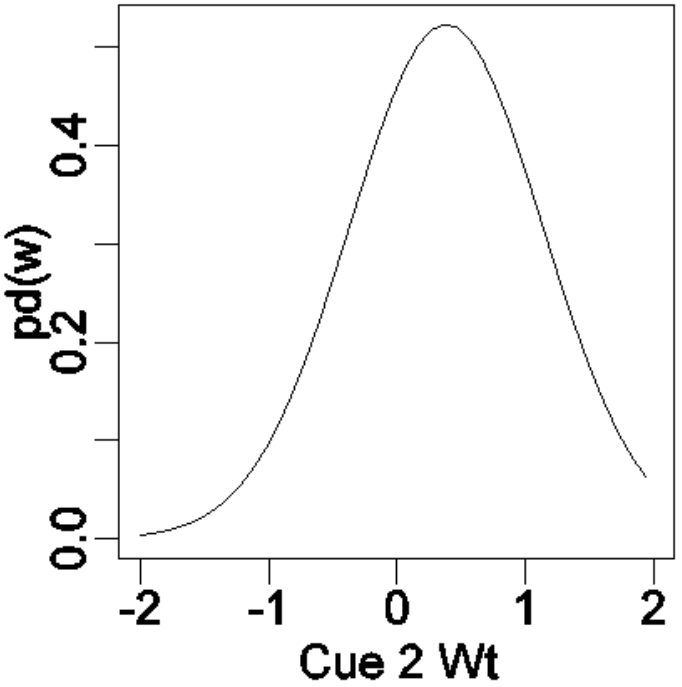


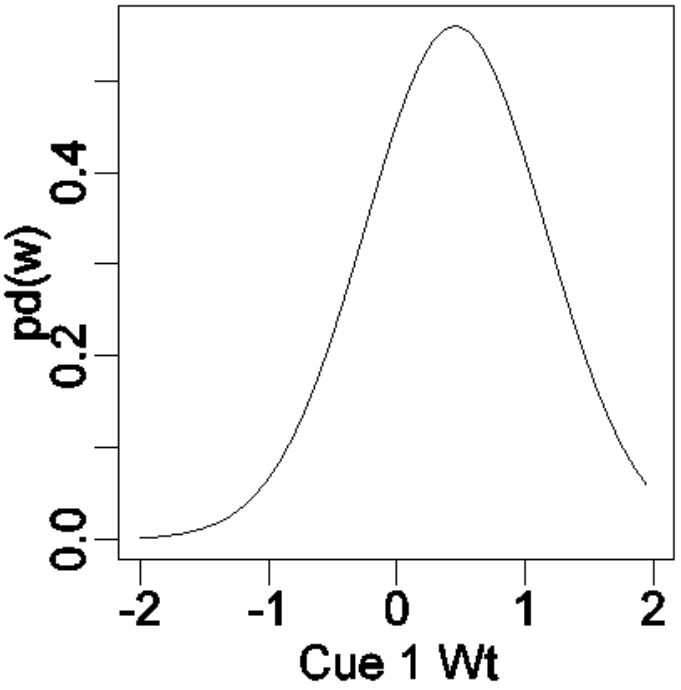




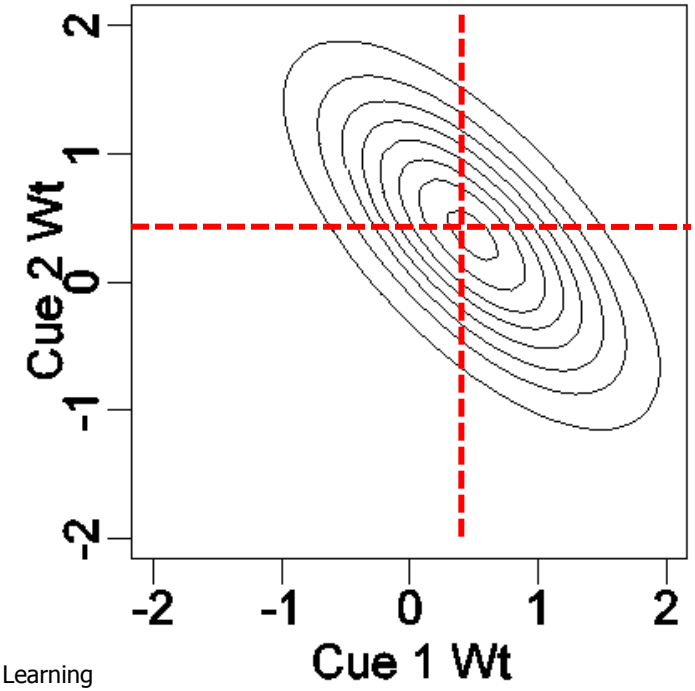
Train: BackwardBlocking
Beginning of Trial 11
Mean: 0.417 0.417
Covariance matrix:
0.583 -0.417
-0.417 0.583
Current Uncertainty: 1.942
Probe: 1 0 => EU: 1.874
Probe: 0 1 => EU: 1.874

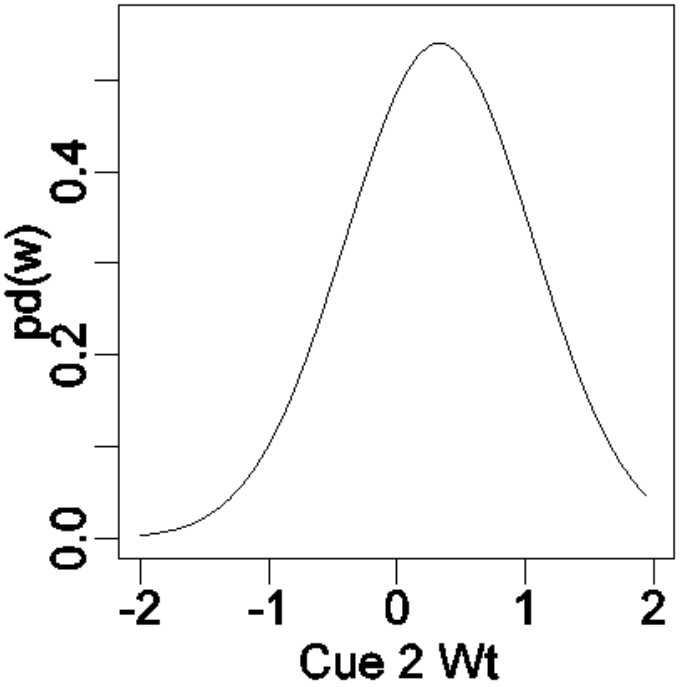


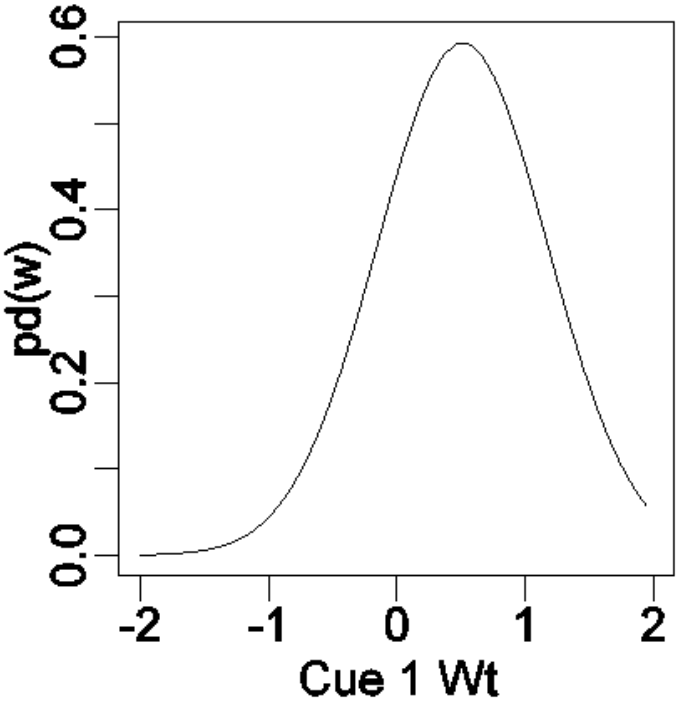




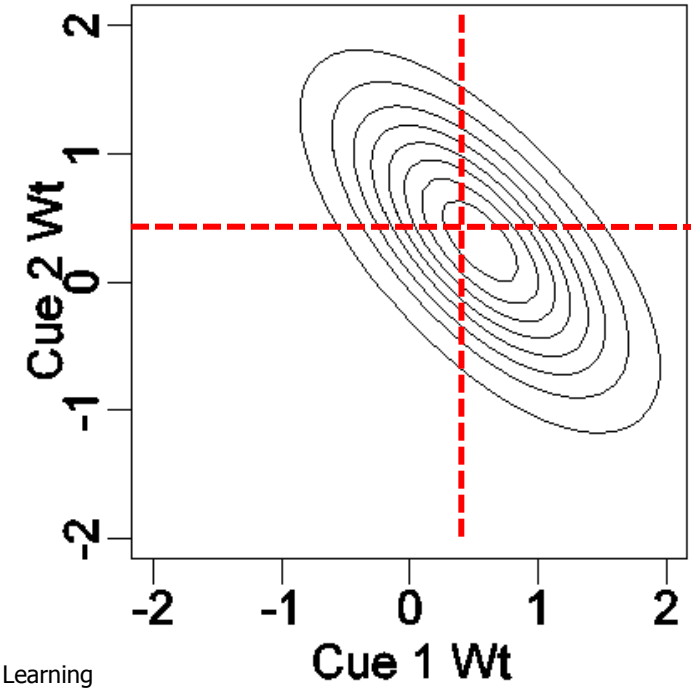
Train: BackwardBlocking
Beginning of Trial 12
Mean: 0.491 0.364
Covariance matrix:
0.509 -0.364
-0.364 0.545
Current Uncertainty: 1.874
Probe: 1 0 => EU: 1.814
Probe: 0 1 => EU: 1.81

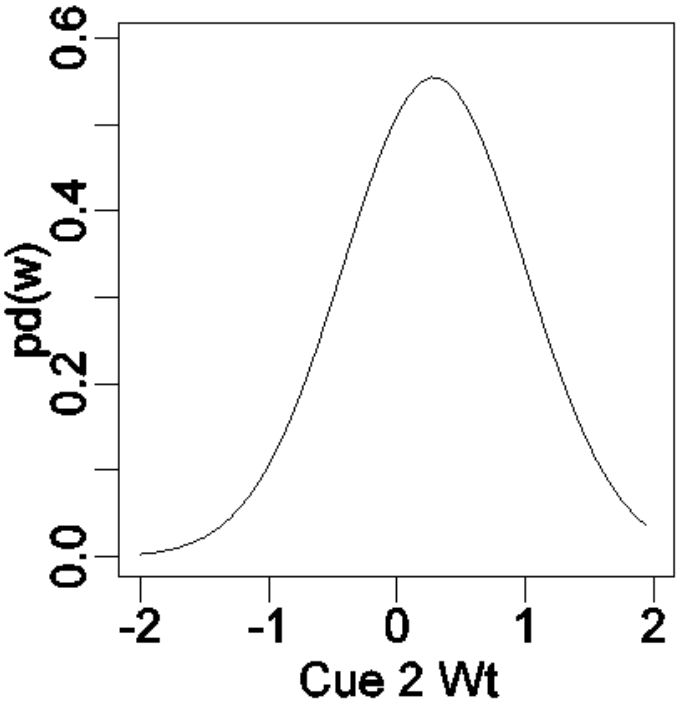


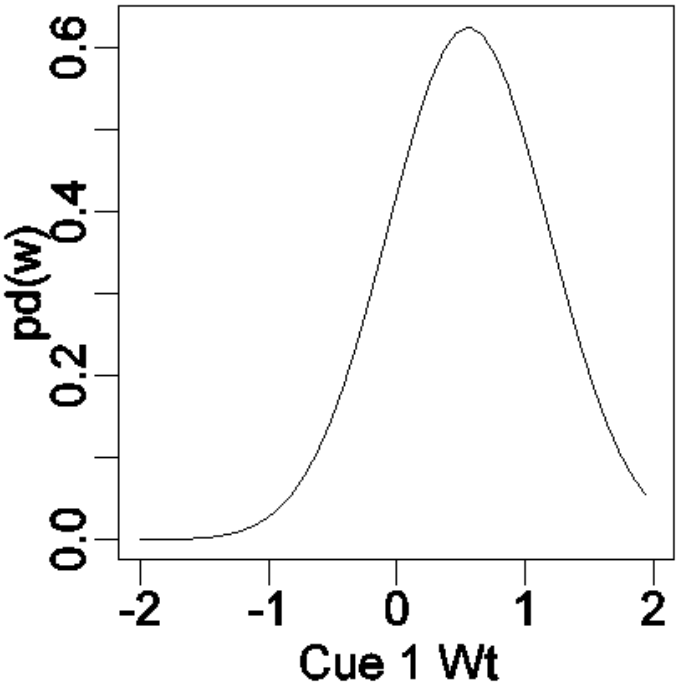




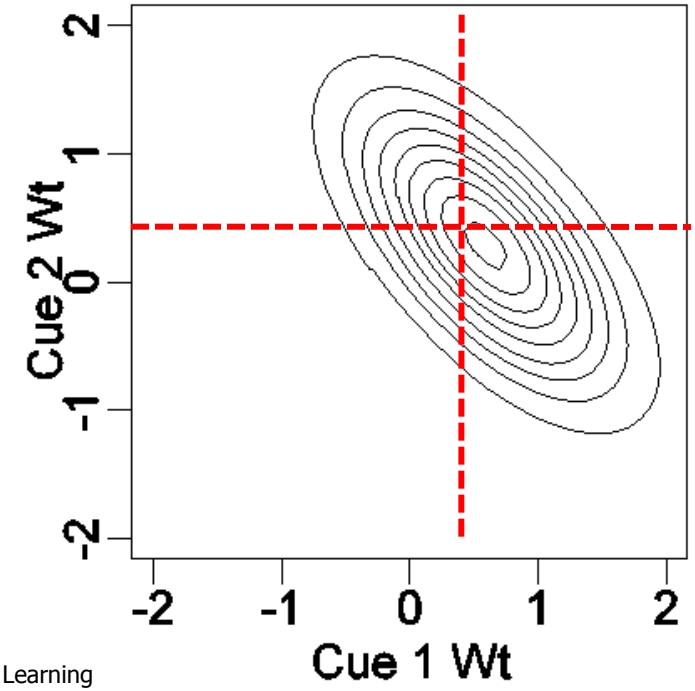
Train: BackwardBlocking
Beginning of Trial 13
Mean: 0.548 0.323
Covariance matrix:
0.452 -0.323
-0.323 0.516
Current Uncertainty: 1.814
Probe: 1 0 => EU: 1.761
Probe: 0 1 => EU: 1.753

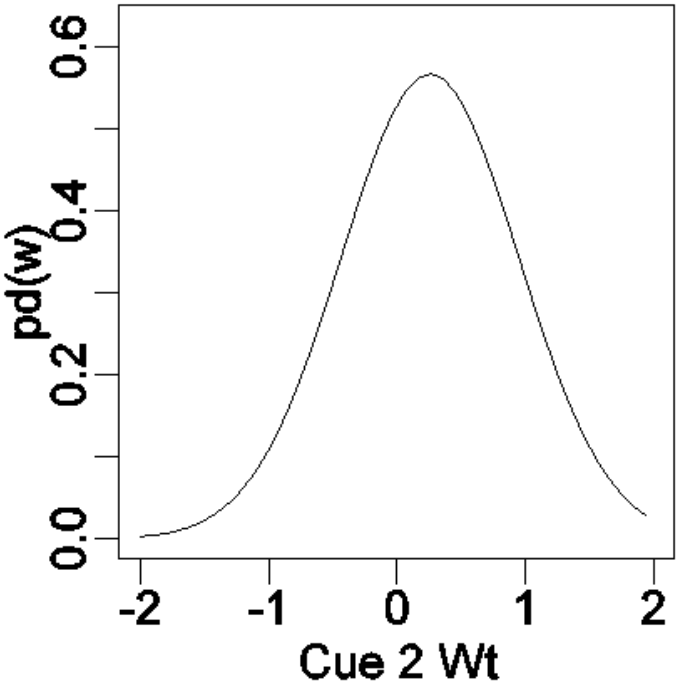


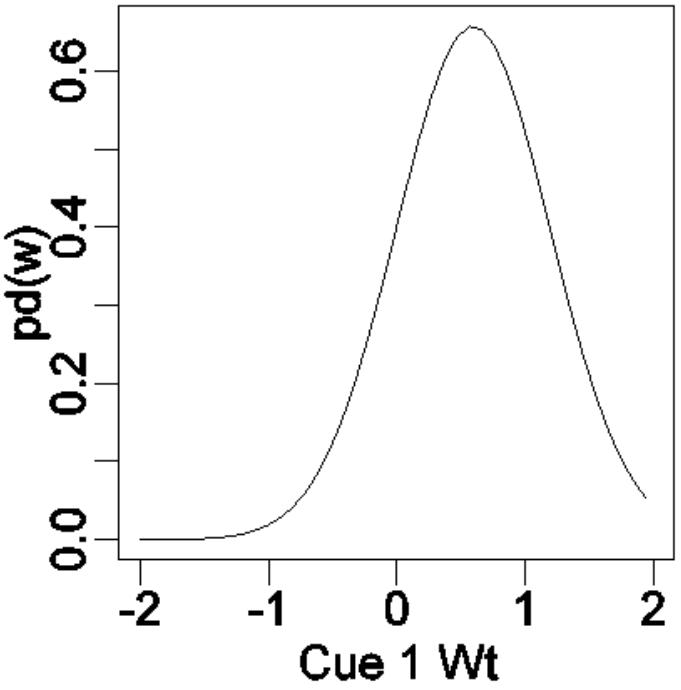




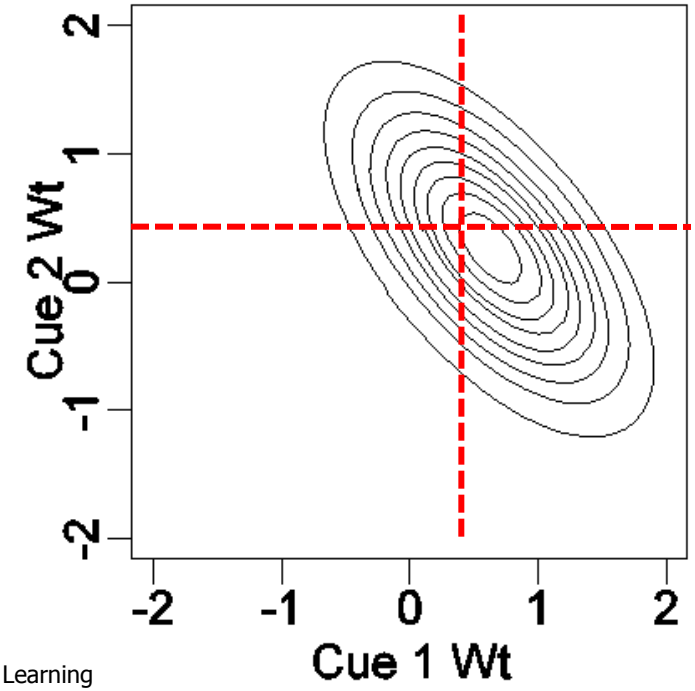
Train: BackwardBlocking
Beginning of Trial 14
Mean: 0.594 0.29
Covariance matrix:
0.406 -0.29
-0.29 0.493
Current Uncertainty: 1.761
Probe: 1 0 => EU: 1.712
Probe: 0 1 => EU: 1.702

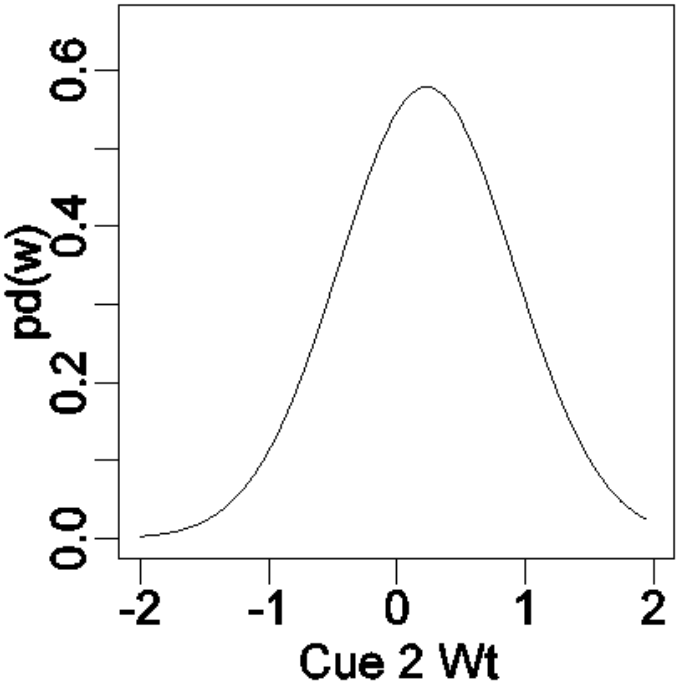


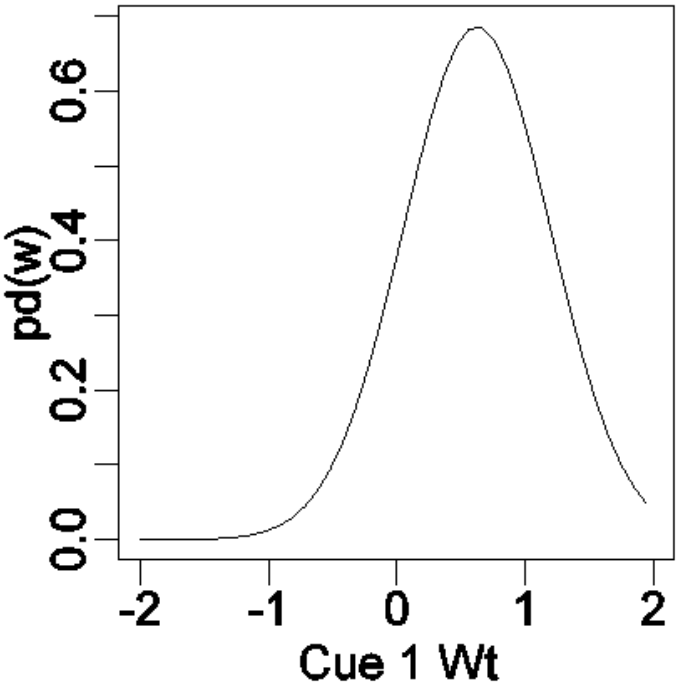




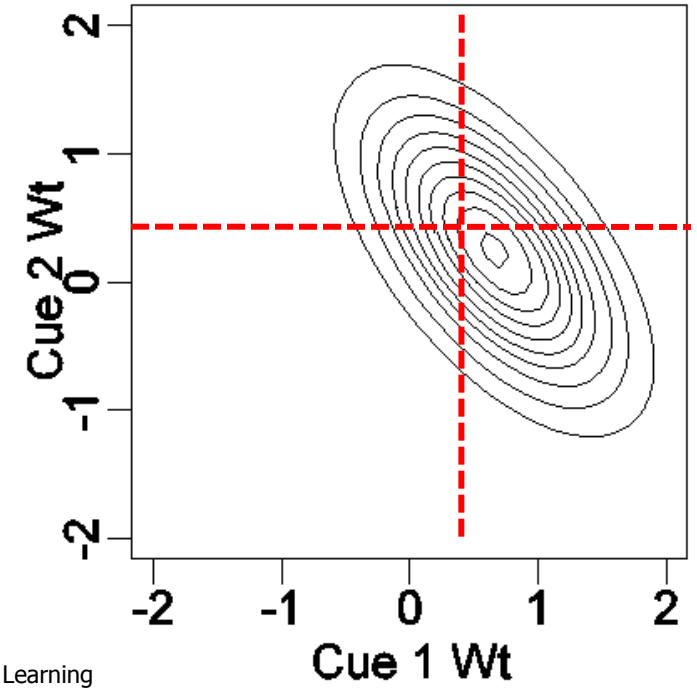
Train: BackwardBlocking
Beginning of Trial 15
Mean: 0.632 0.263
Covariance matrix:
0.368 -0.263
-0.263 0.474
Current Uncertainty: 1.712
Probe: 1 0 => EU: 1.668
Probe: 0 1 => EU: 1.656

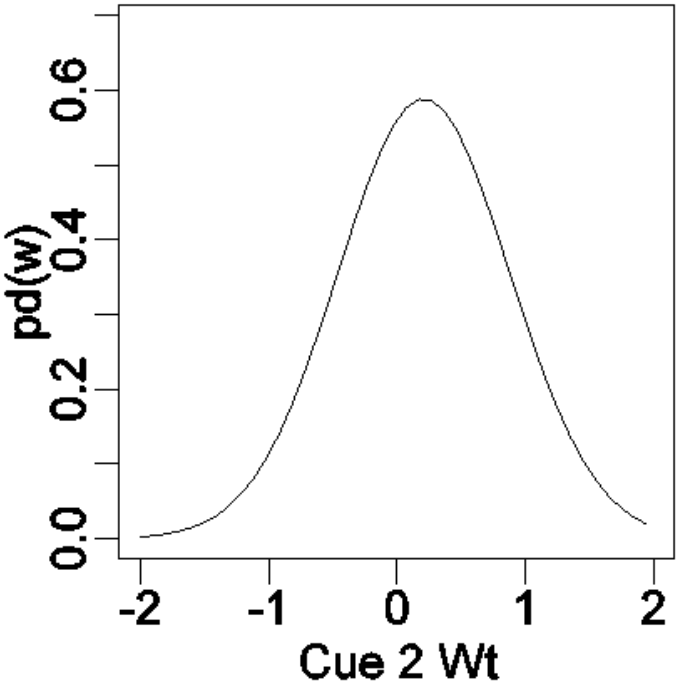


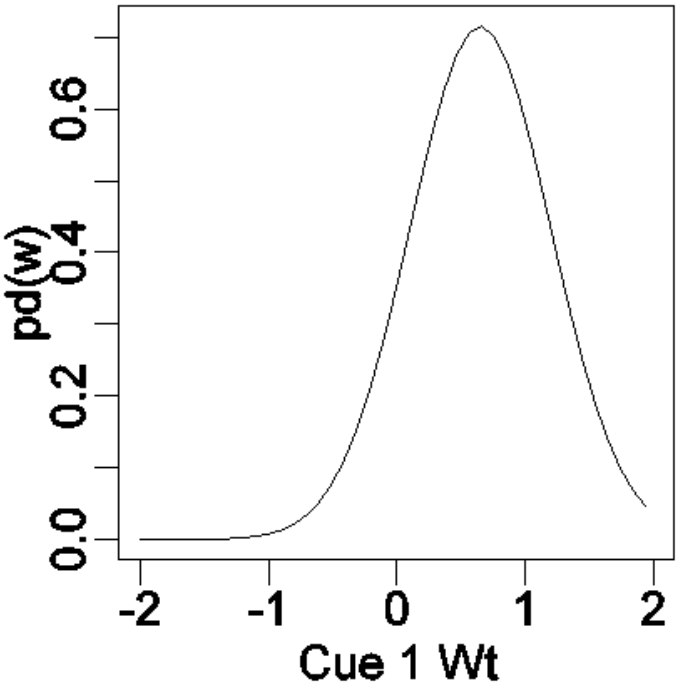




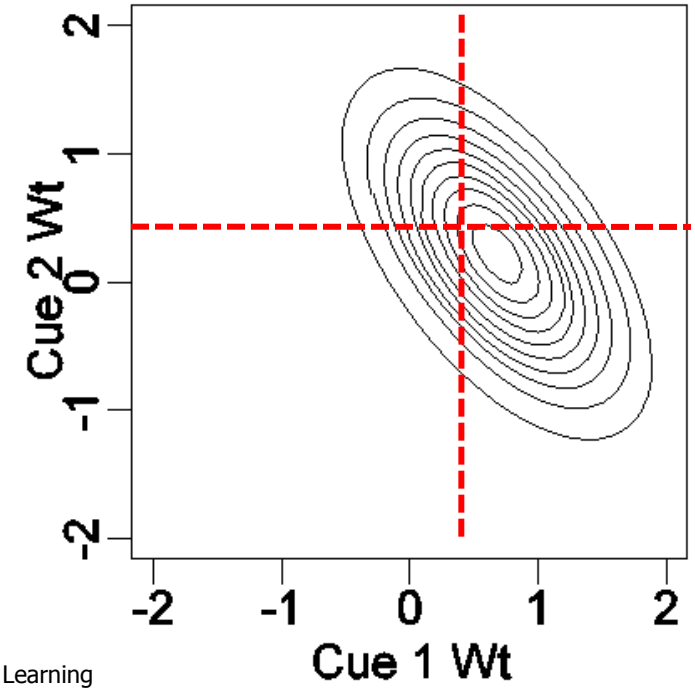
Train: BackwardBlocking
Beginning of Trial 16
Mean: 0.663 0.241
Covariance matrix:
0.337 -0.241
-0.241 0.458
Current Uncertainty: 1.668
Probe: 1 0 => EU: 1.628
Probe: 0 1 => EU: 1.614

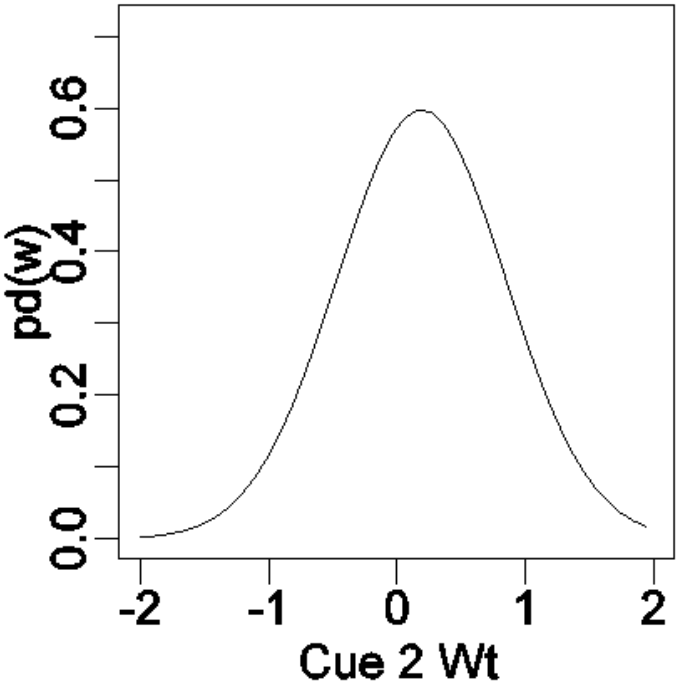


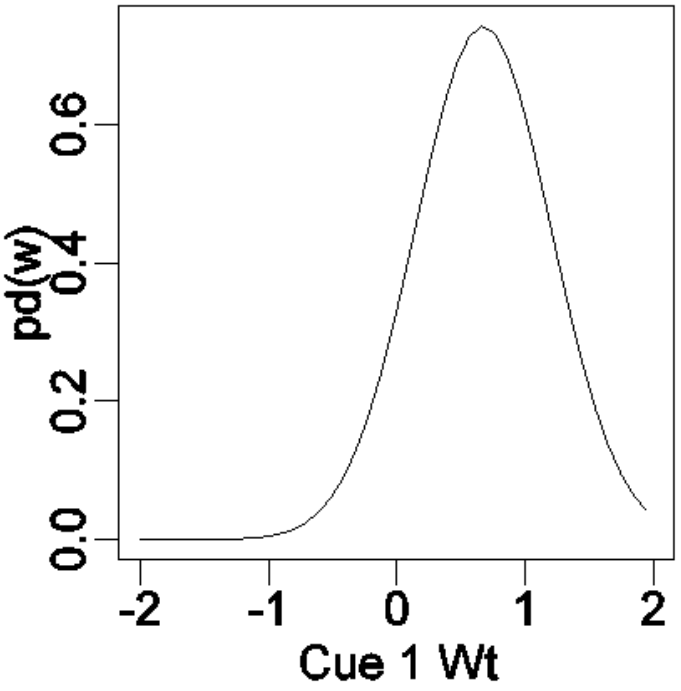




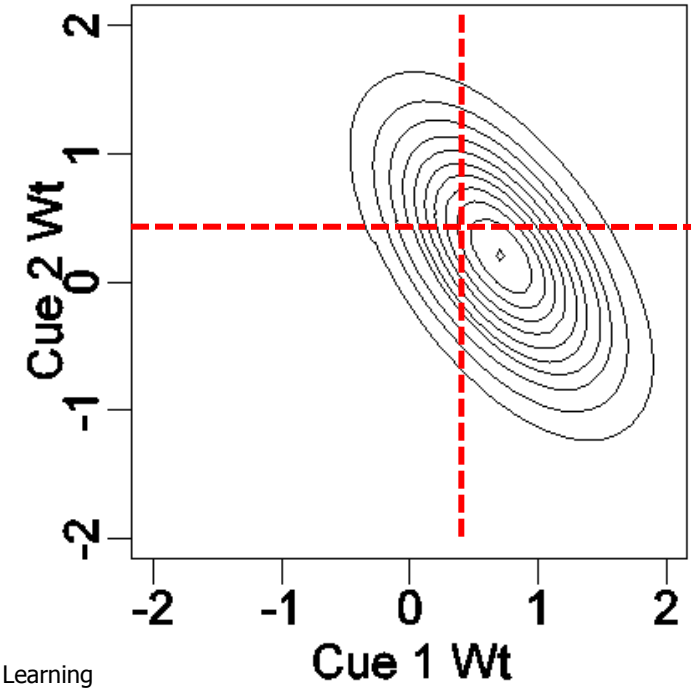
Train: BackwardBlocking
Beginning of Trial 17
Mean: 0.689 0.222
Covariance matrix:
0.311 -0.222
-0.222 0.444
Current Uncertainty: 1.628
Probe: 1 0 => EU: 1.59
Probe: 0 1 => EU: 1.575

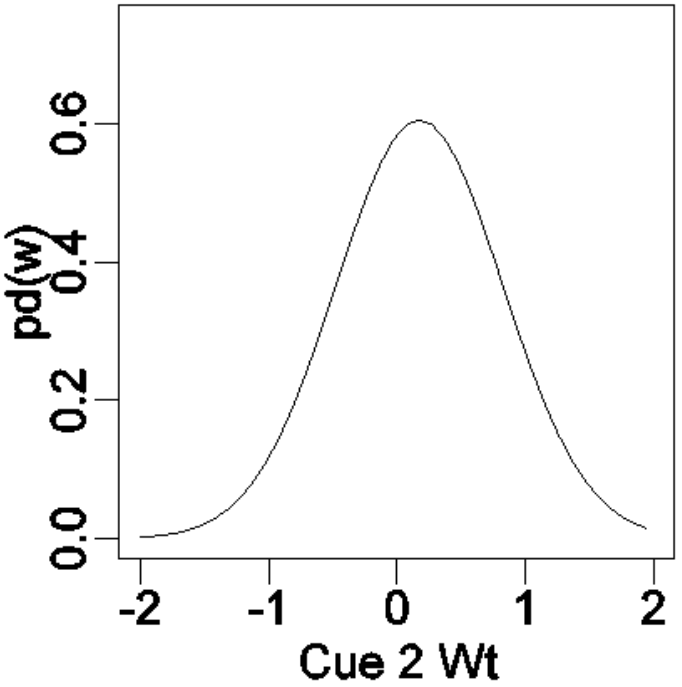


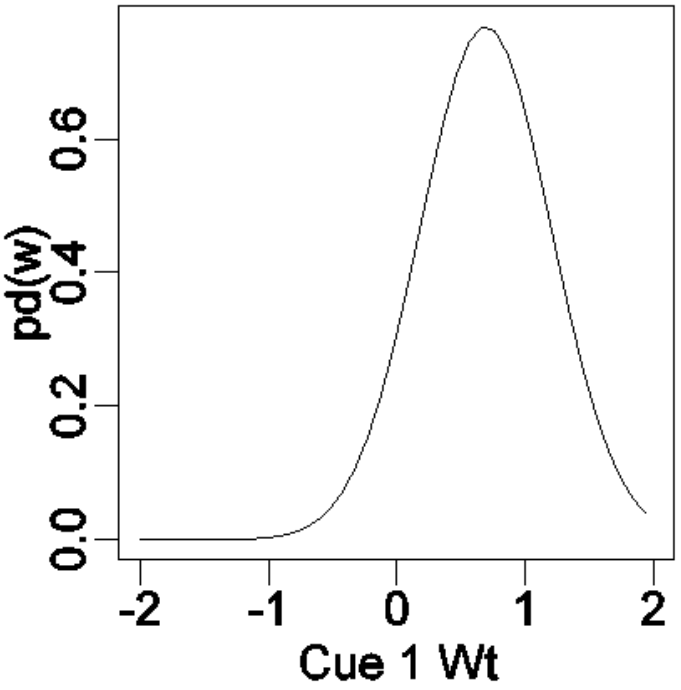




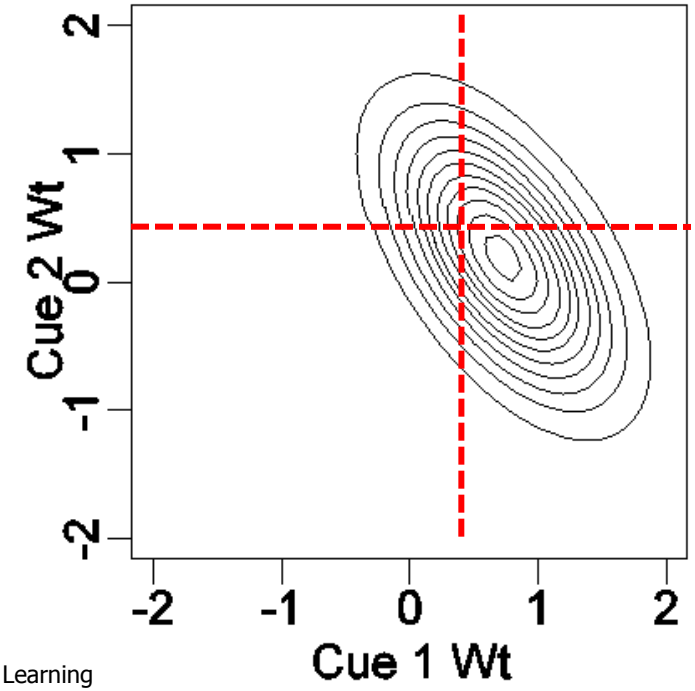
Train: BackwardBlocking
Beginning of Trial 18
Mean: 0.711 0.206
Covariance matrix:
0.289 -0.206
-0.206 0.433
Current Uncertainty: 1.59
Probe: 1 0 => EU: 1.555
Probe: 0 1 => EU: 1.539

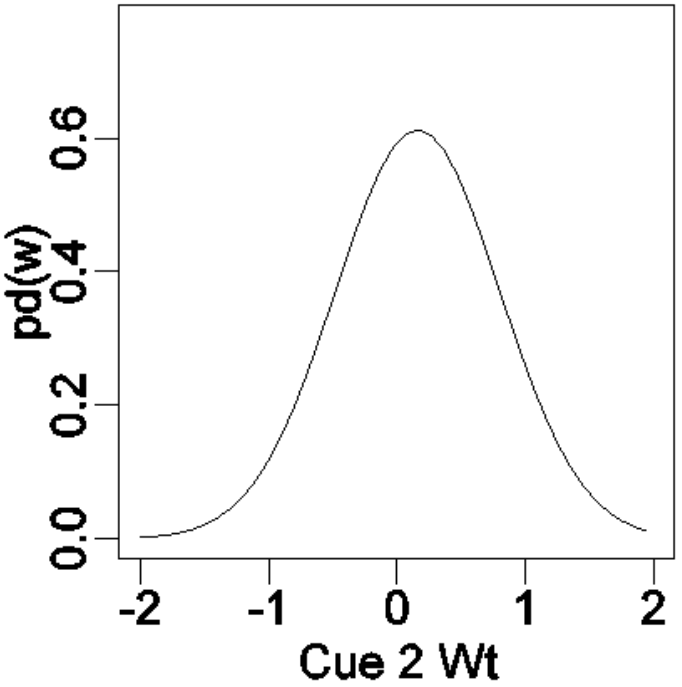


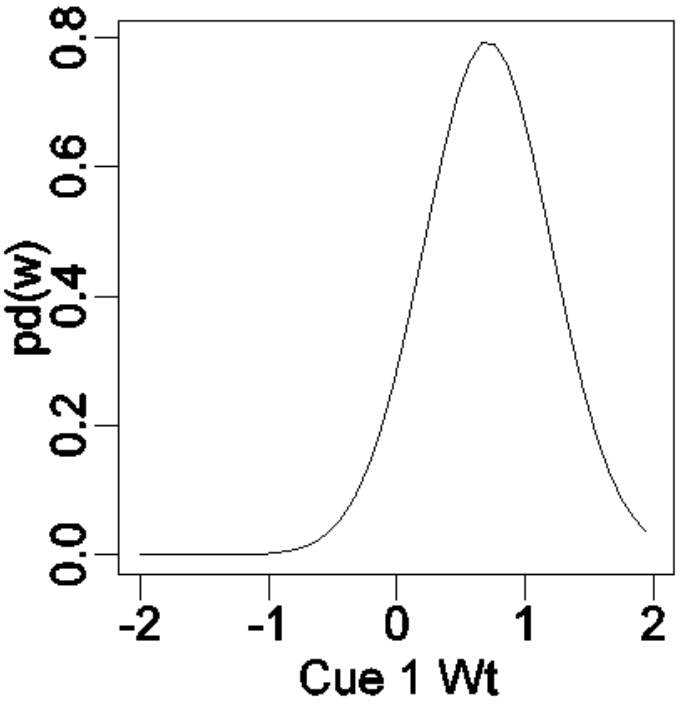




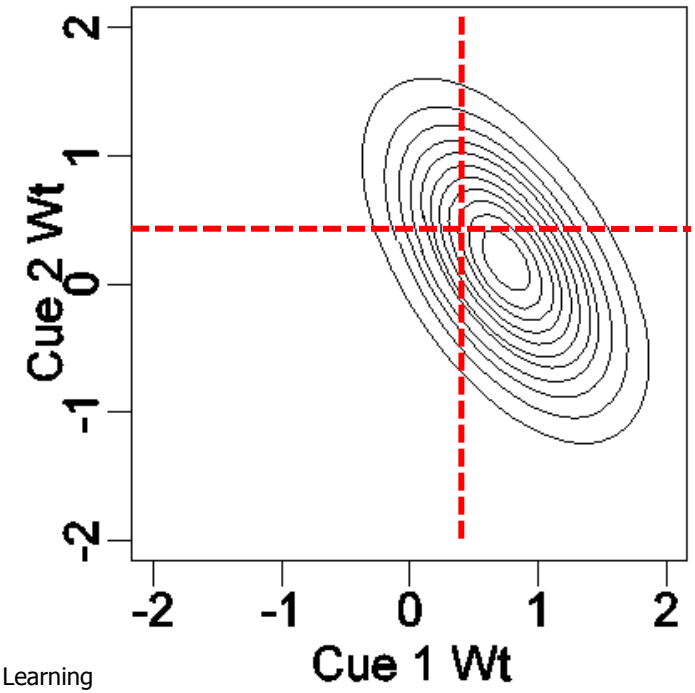
Train: BackwardBlocking
Beginning of Trial 19
Mean: 0.731 0.192
Covariance matrix:
0.269 -0.192
-0.192 0.423
Current Uncertainty: 1.555
Probe: 1 0 => EU: 1.523
Probe: 0 1 => EU: 1.505

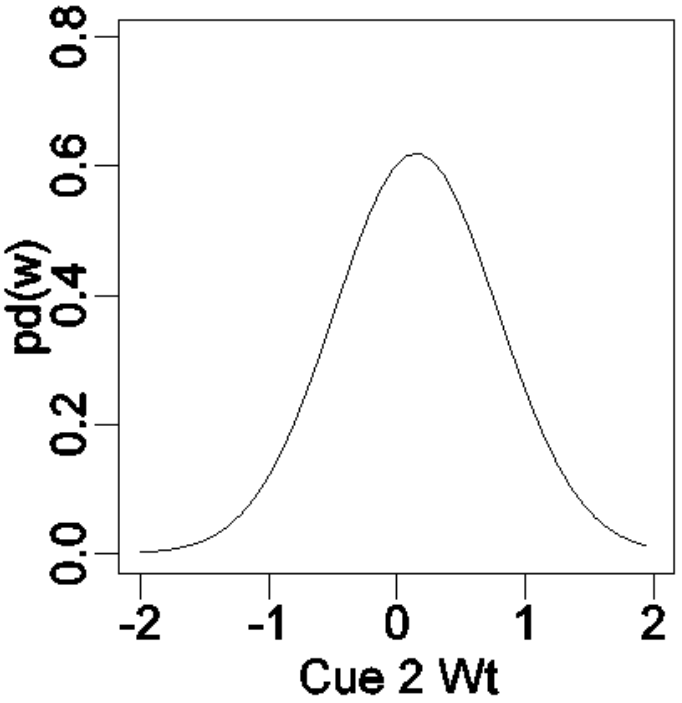


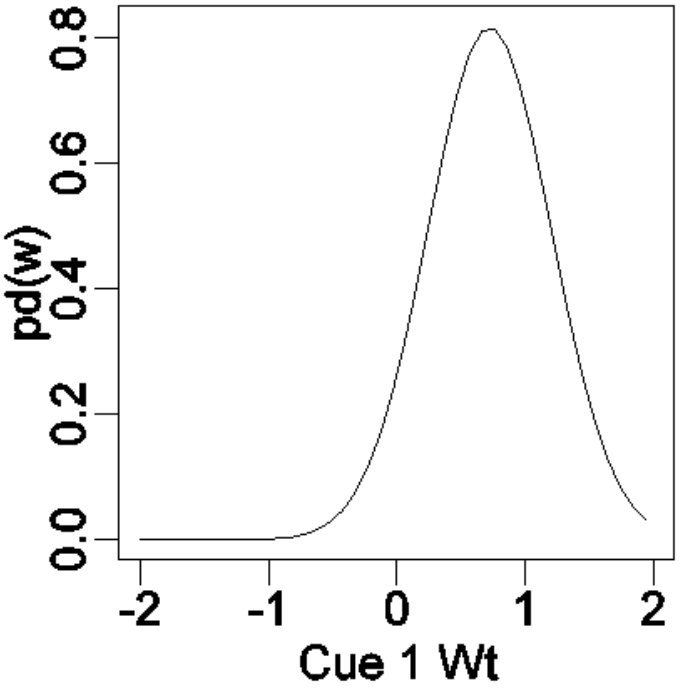




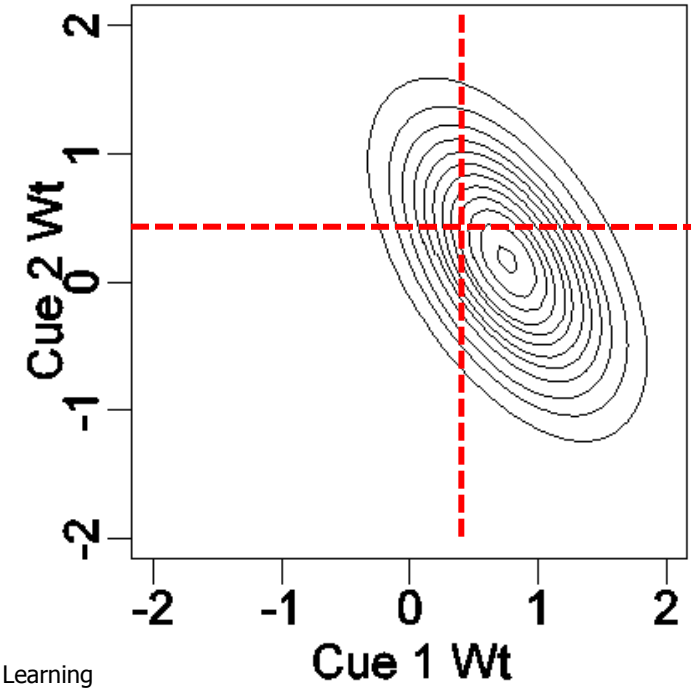
Train: BackwardBlocking
Beginning of Trial 20
Mean: 0.748 0.18
Covariance matrix:
0.252 -0.18
-0.18 0.414
Current Uncertainty: 1.523
Probe: 1 0 => EU: 1.492
Probe: 0 1 => EU: 1.474

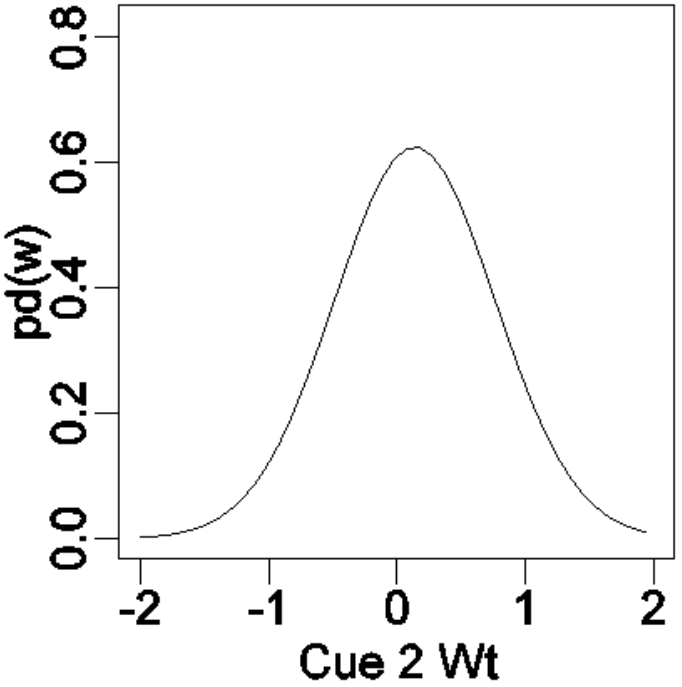


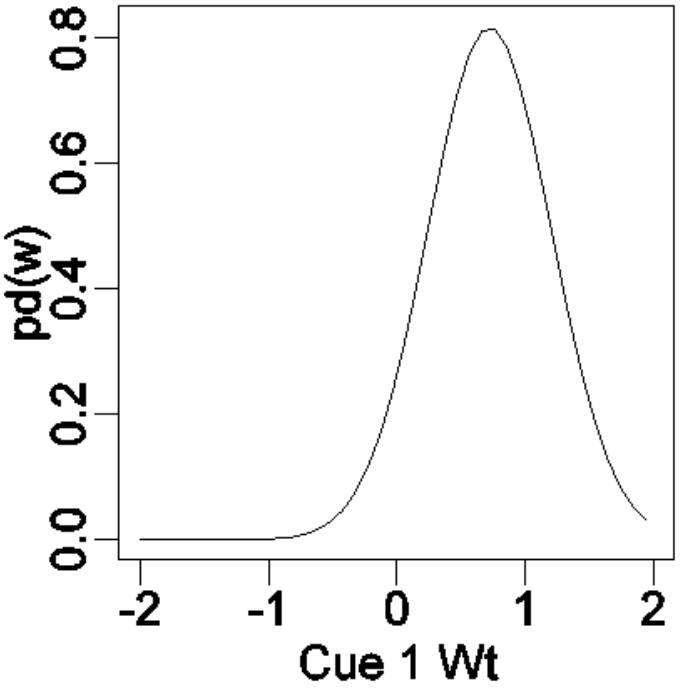




Train: BackwardBlocking
Beginning of Trial 21
Mean: 0.763 0.169
Covariance matrix:
0.237 -0.169
-0.169 0.407
Current Uncertainty: 1.492
Probe: 1 0 => EU: 1.463
Probe: 0 1 => EU: 1.444

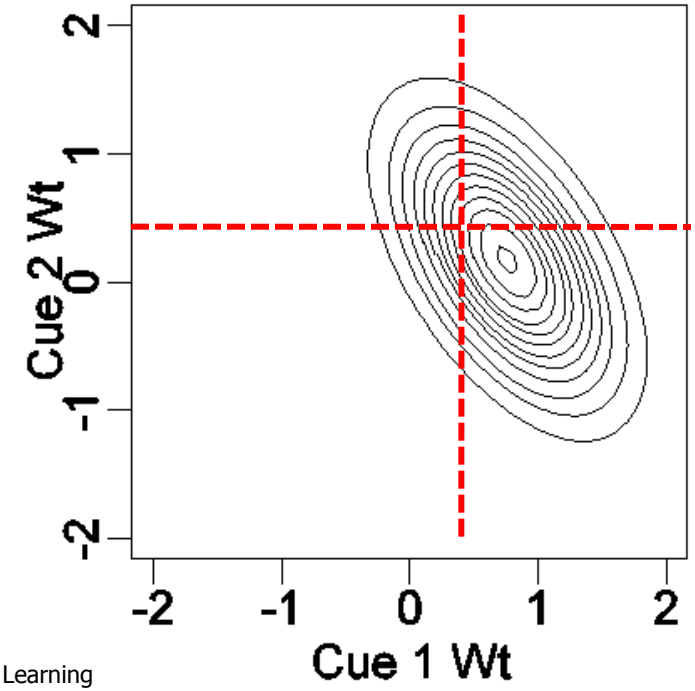


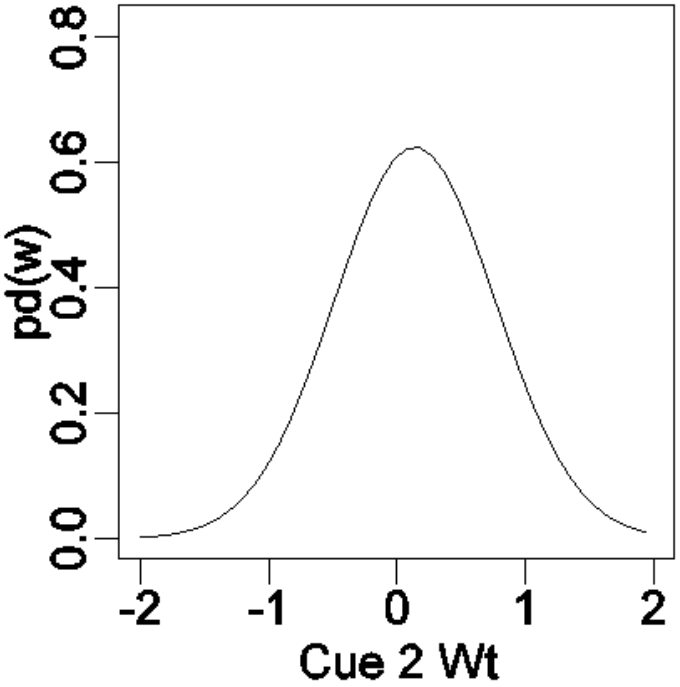




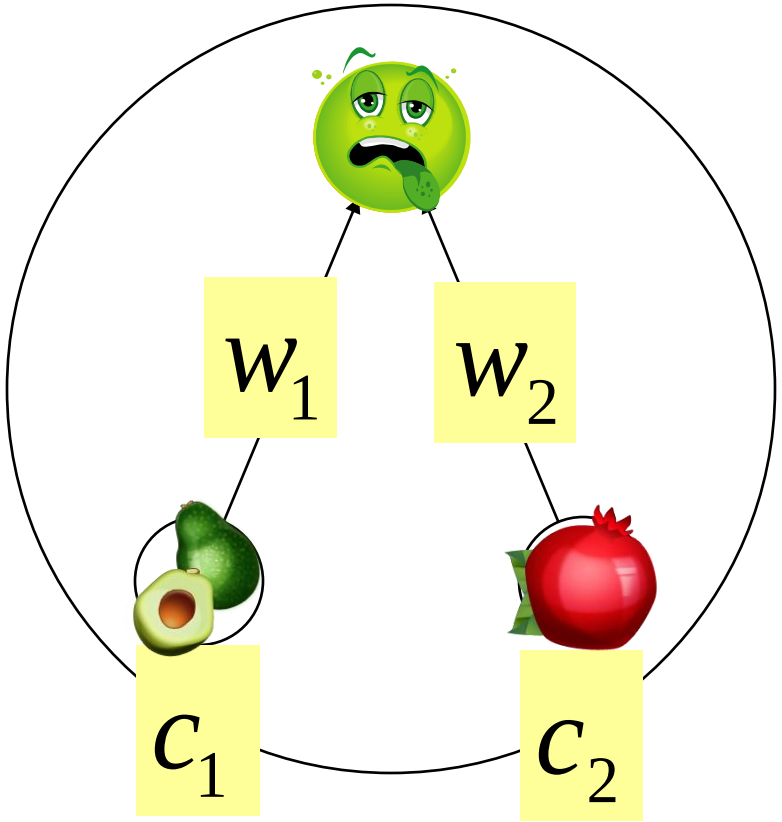
Train: BackwardBlocking
Beginning of Trial 21
Mean: 0.763 0.169

Kalman filter *does* show backward blocking.





Noisy-logic gate



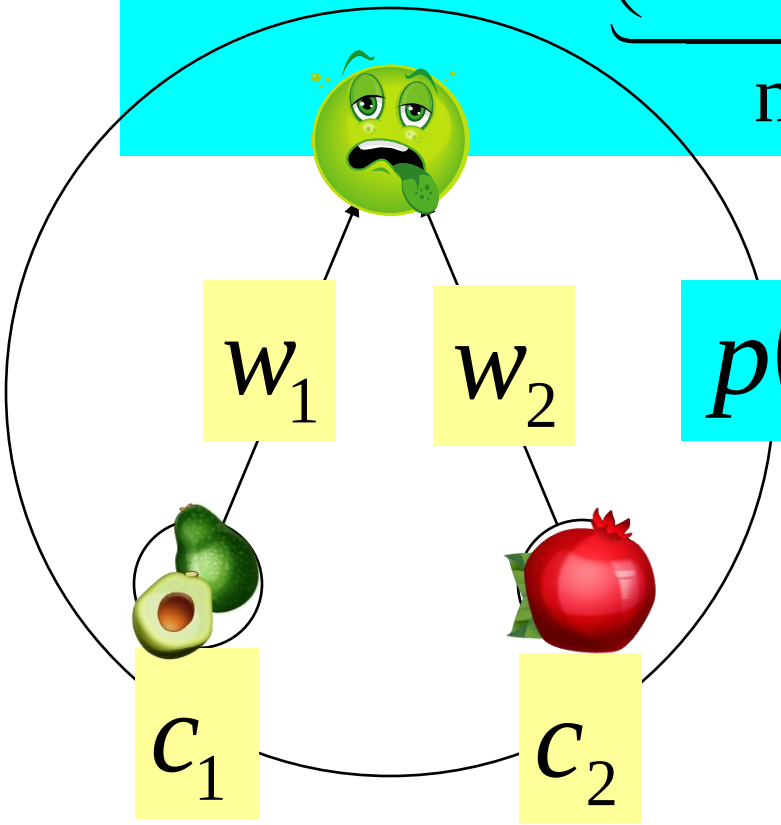
Outcome is present/absent.

Positive weight is probability that outcome occurs if that cue alone is present.

Negative weight is probability that outcome does not occur, when otherwise it would have, if that cue alone is present.

Noisy-logic gate

$$p(t=1|\vec{c},\vec{w}) = \underbrace{\left(1 - \prod_{i\ s.t.\ w_i \geq 0} (1-w_i)^{c_i}\right)}_{\text{noisy OR}} \underbrace{\prod_{j\ s.t.\ w_j < 0} (1+w_j)^{c_j}}_{\text{noisy NOT AND}}$$

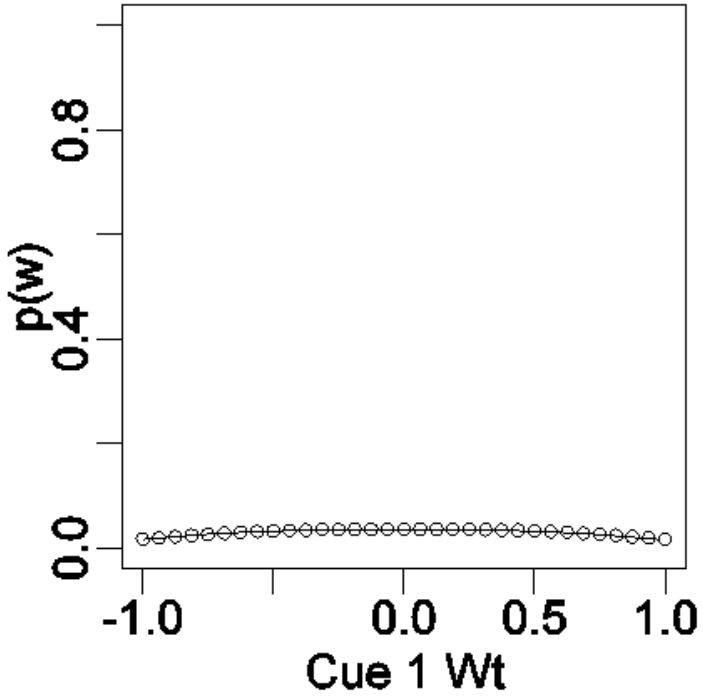


$$p(\vec{w}) \sim \ldots$$

Backward Blocking

Phase	Frequency	Cue 1	Cue 2	Outcome
I	10	1 	1 	1 
II	10	1 	0	1 

Note: Cell with a 1 indicates presence of cue or outcome. Cell with a 0 indicates absence of cue or outcome.



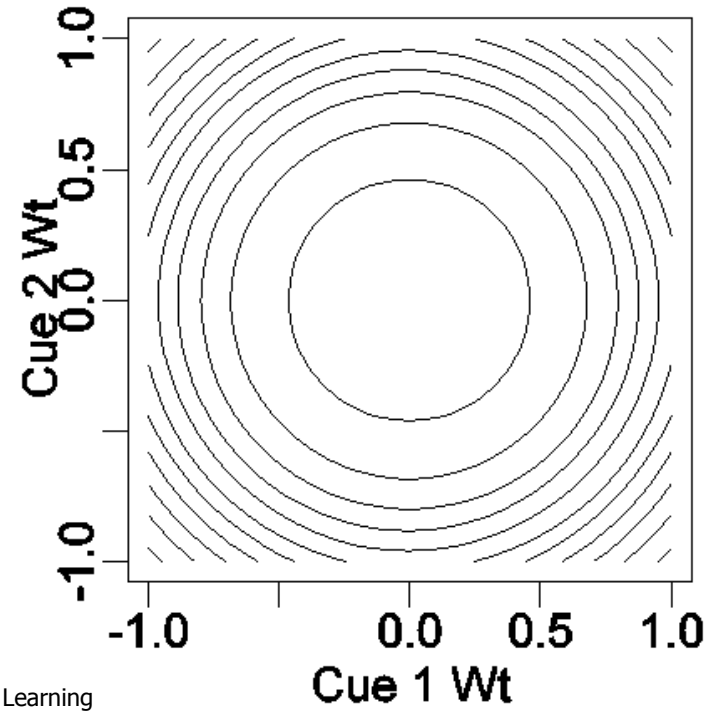
Training: BackwardBlocking

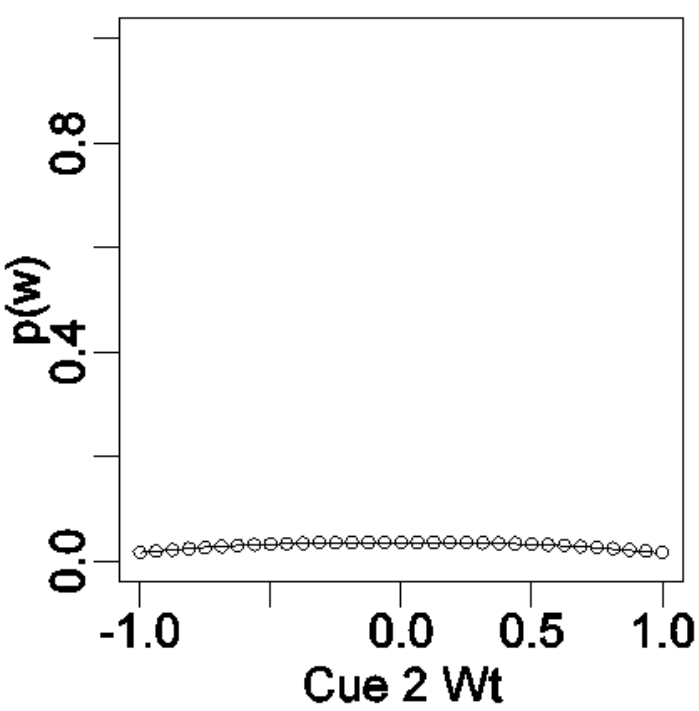
Beginning of Trial 1

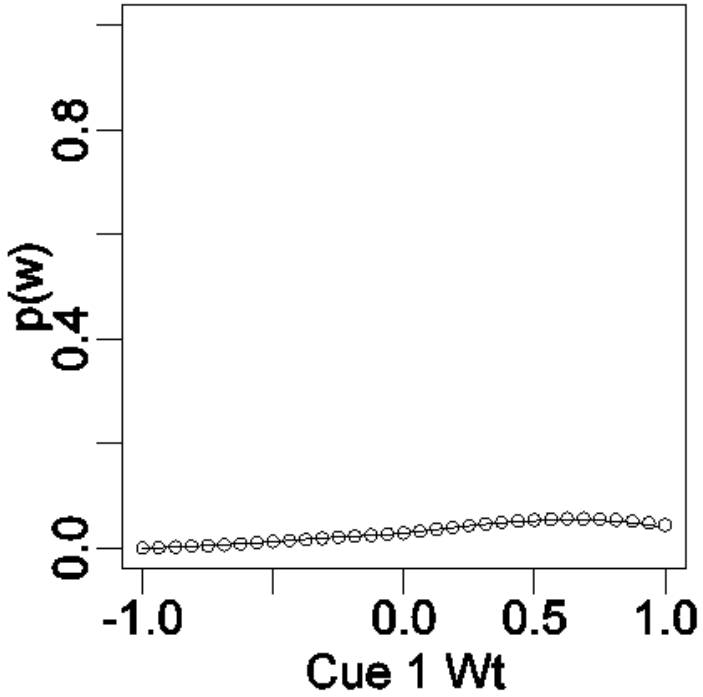
Current Uncertainty: 0.993

Probe: 1 0 => EU: 0.952

Probe: 0 1 => EU: 0.952







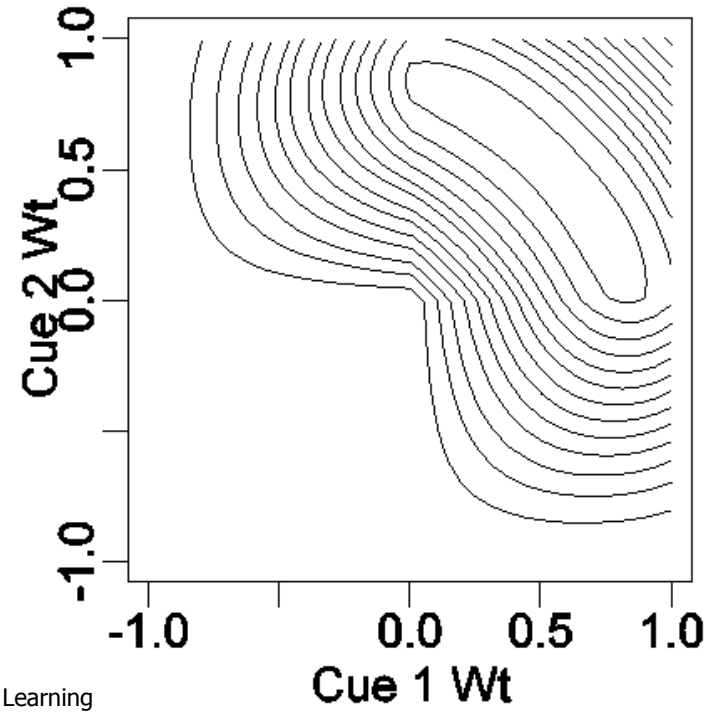
Training: BackwardBlocking

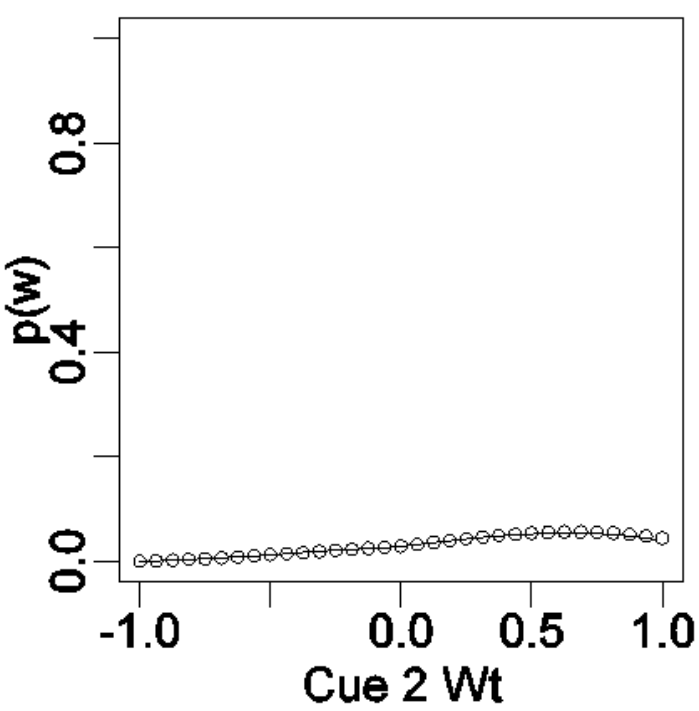
Beginning of Trial 2

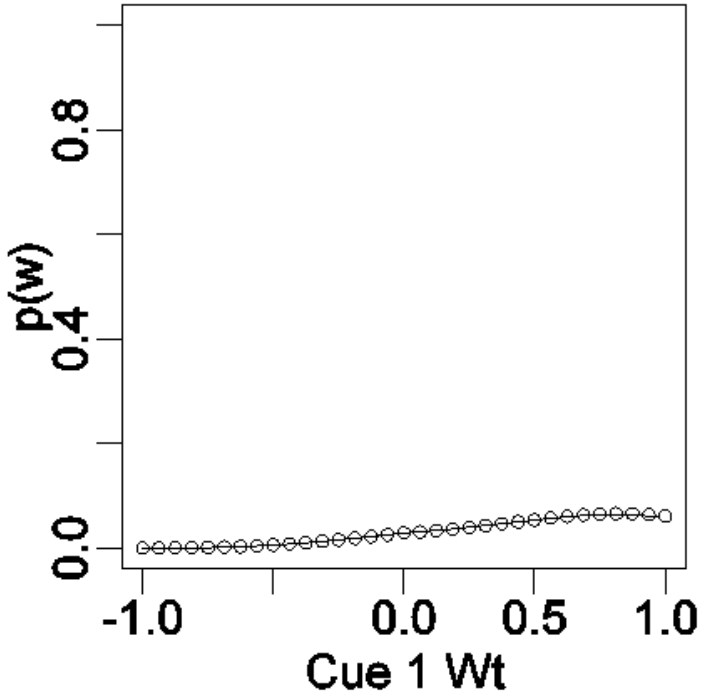
Current Uncertainty: 0.911

Probe: 1 0 => EU: 0.87

Probe: 0 1 => EU: 0.87







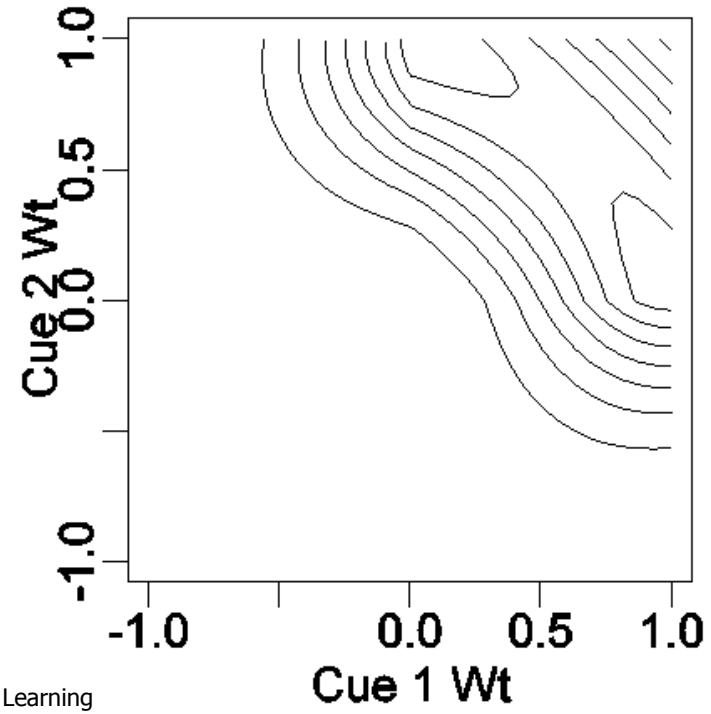
Training: BackwardBlocking

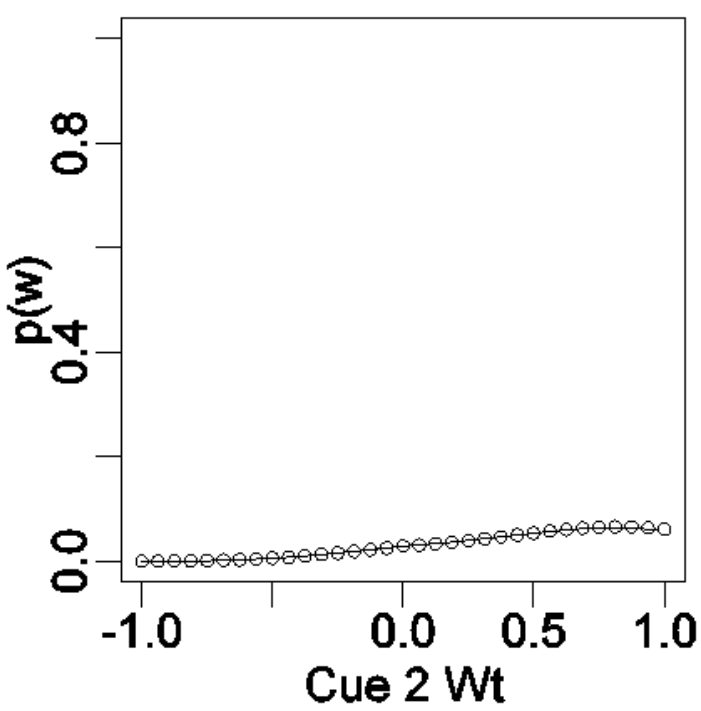
Beginning of Trial 3

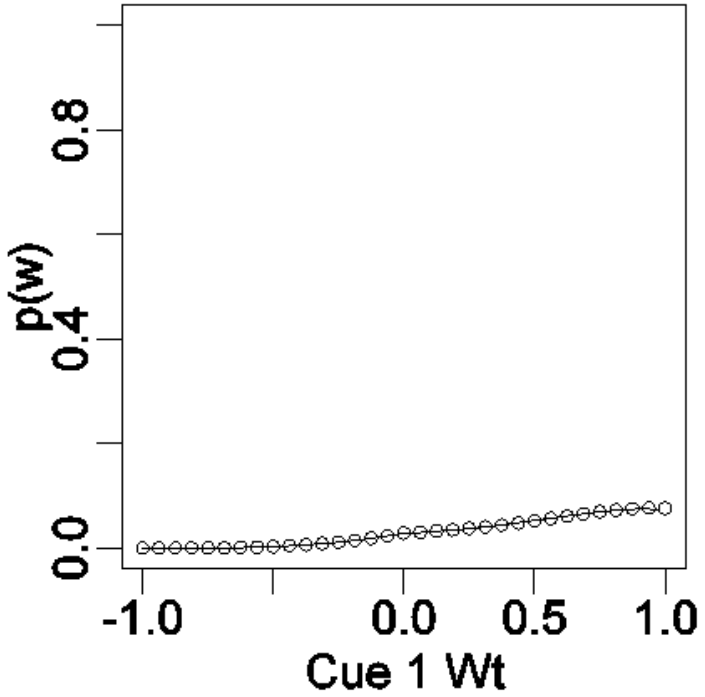
Current Uncertainty: 0.871

Probe: 1 0 => EU: 0.832

Probe: 0 1 => EU: 0.832







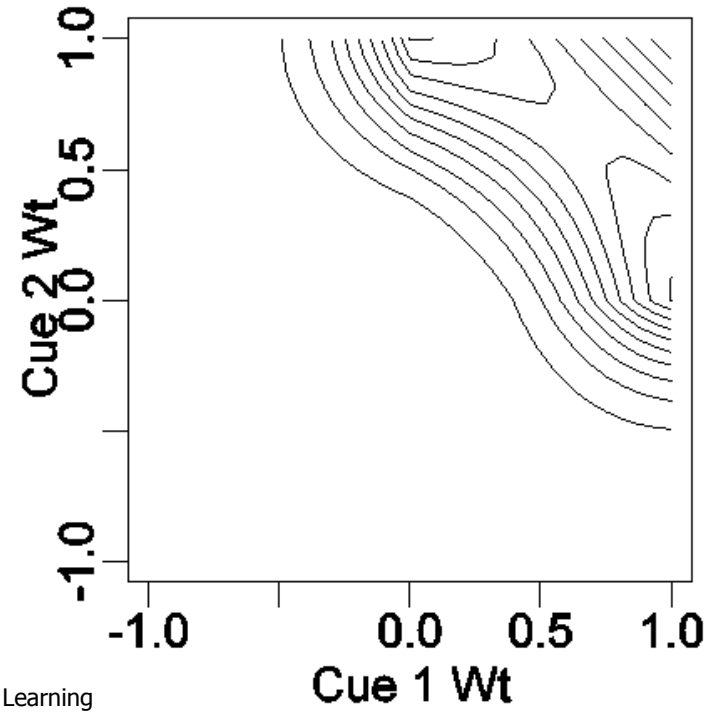
Training: BackwardBlocking

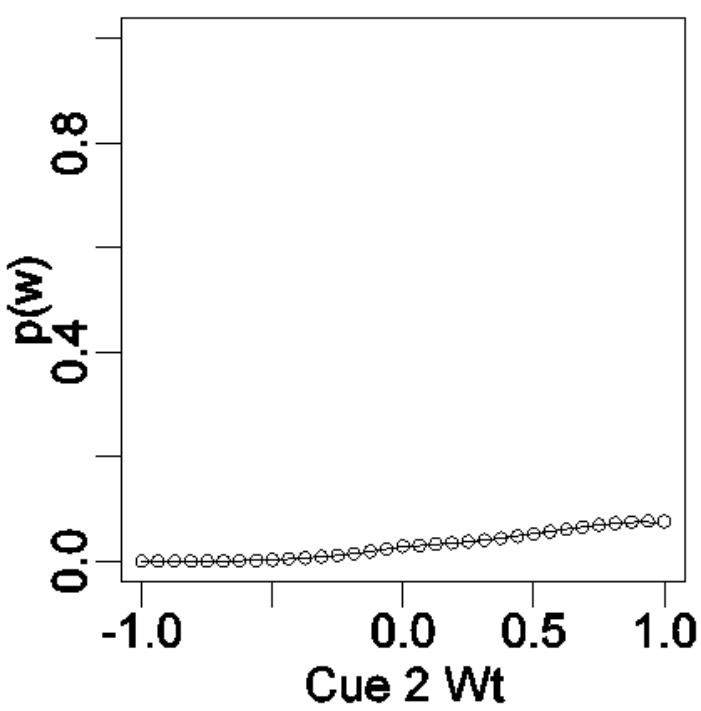
Beginning of Trial 4

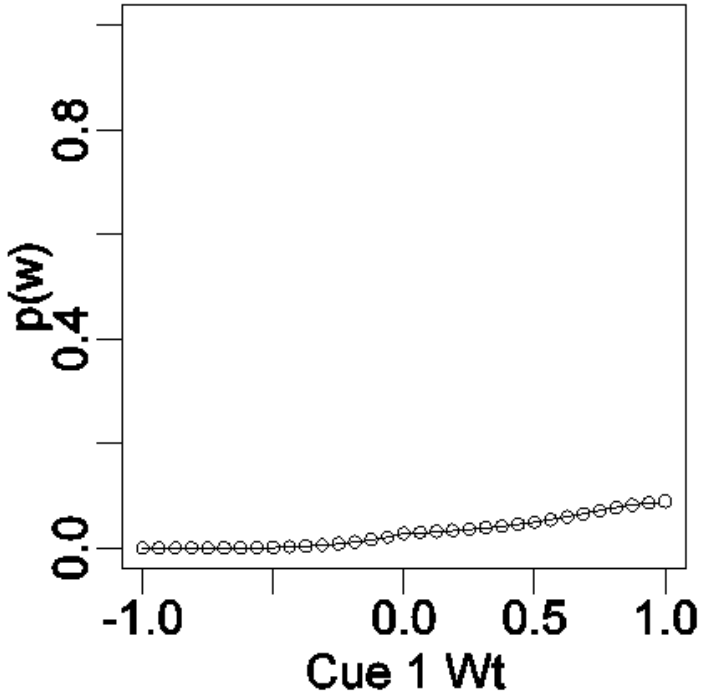
Current Uncertainty: 0.843

Probe: 1 0 => EU: 0.804

Probe: 0 1 => EU: 0.804







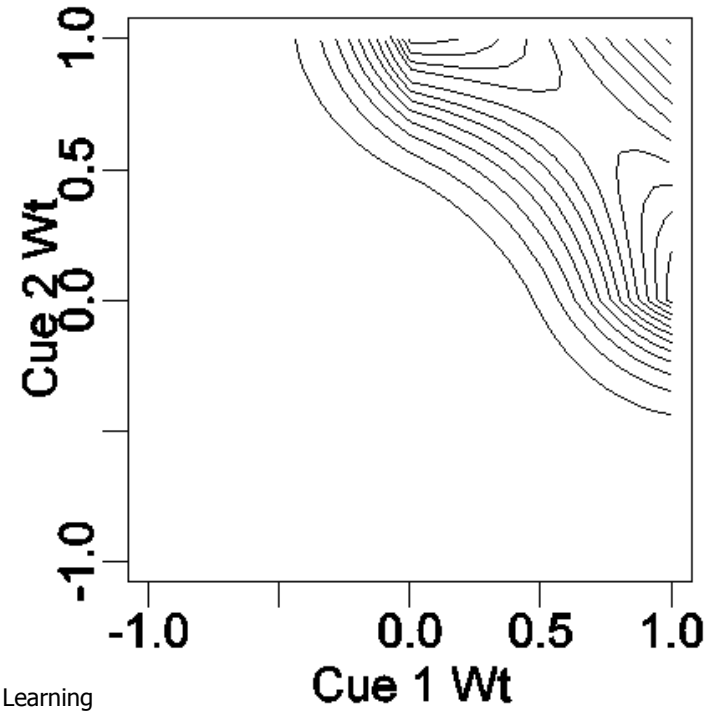
Training: BackwardBlocking

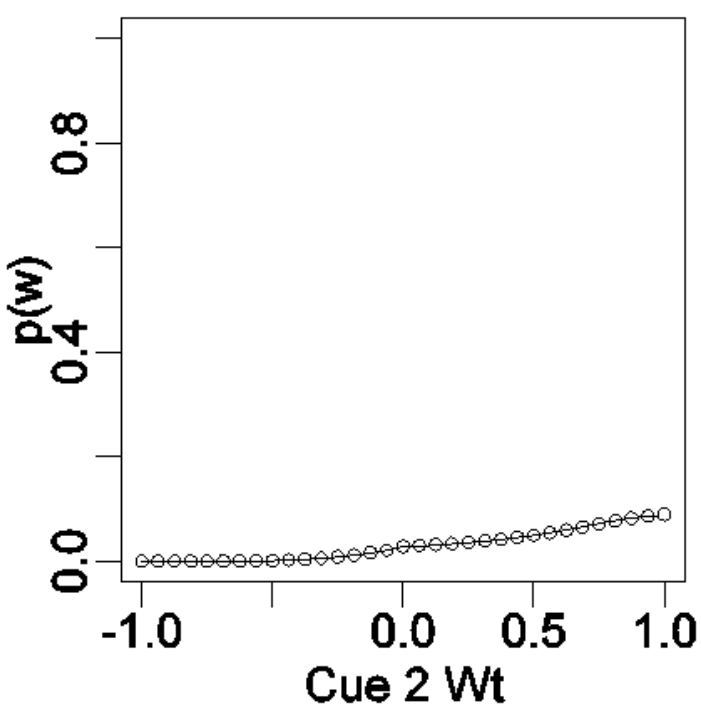
Beginning of Trial 5

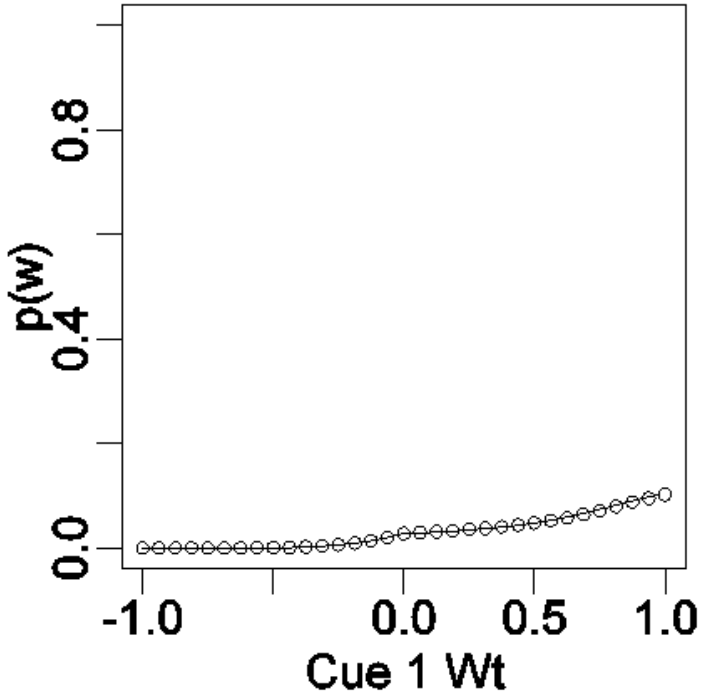
Current Uncertainty: 0.821

Probe: 1 0 => EU: 0.783

Probe: 0 1 => EU: 0.783







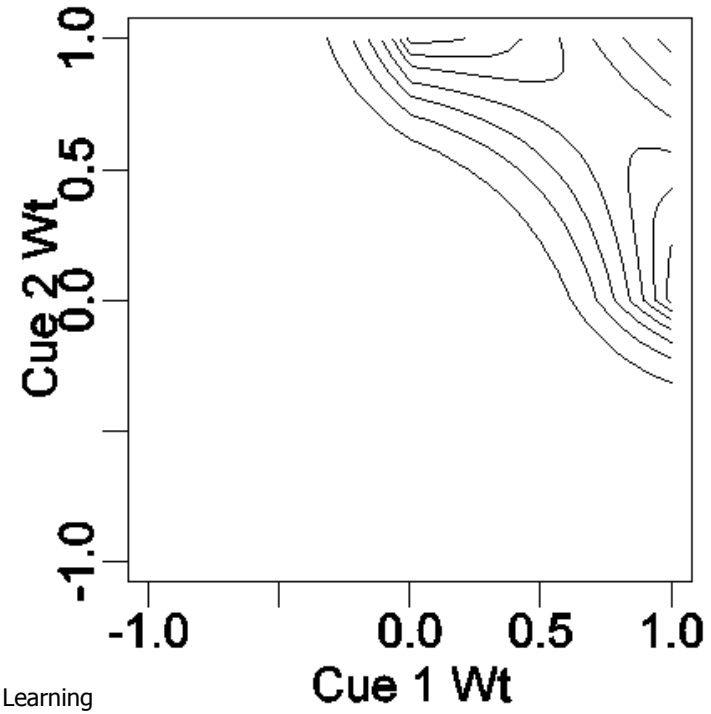
Training: BackwardBlocking

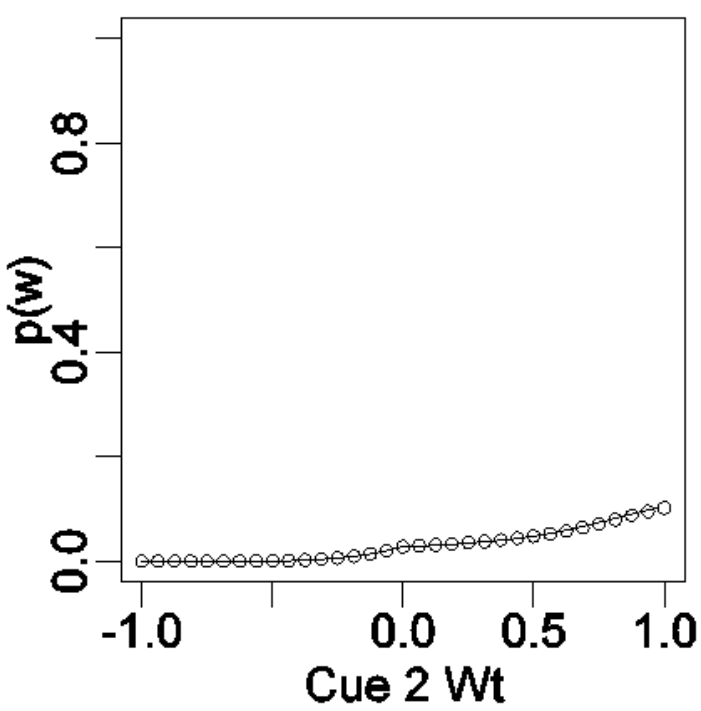
Beginning of Trial 6

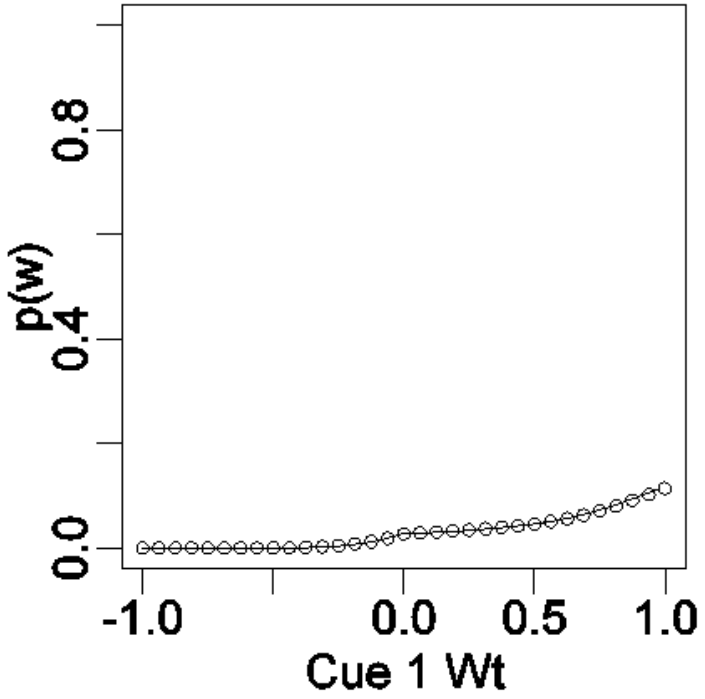
Current Uncertainty: 0.803

Probe: 1 0 => EU: 0.766

Probe: 0 1 => EU: 0.766







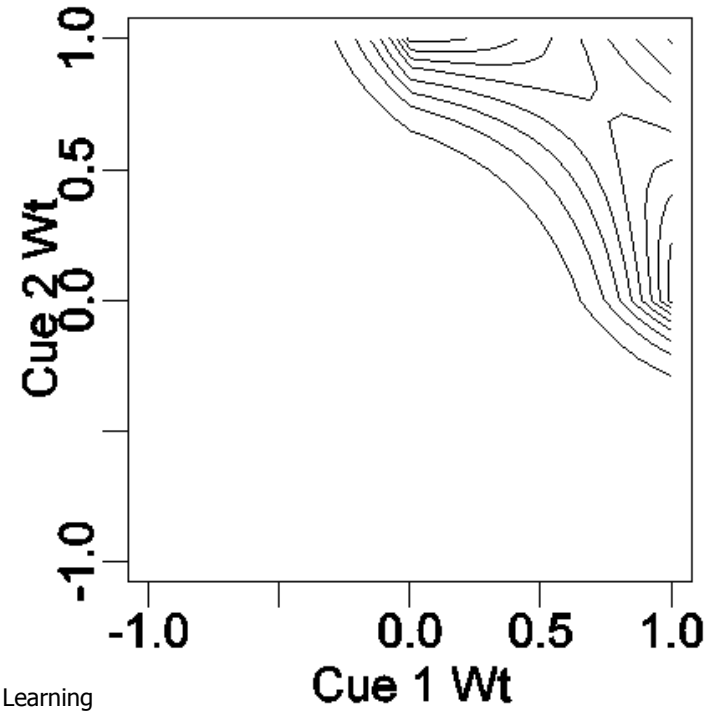
Training: BackwardBlocking

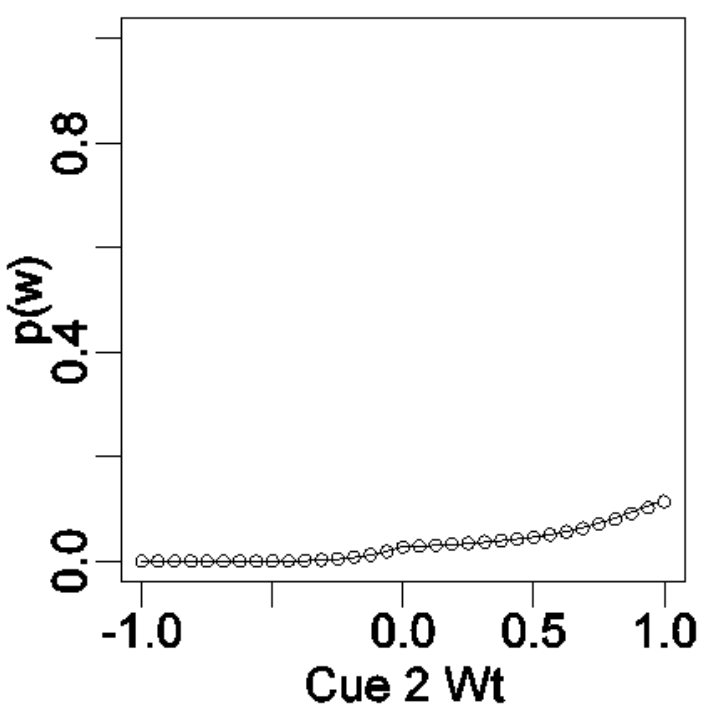
Beginning of Trial 7

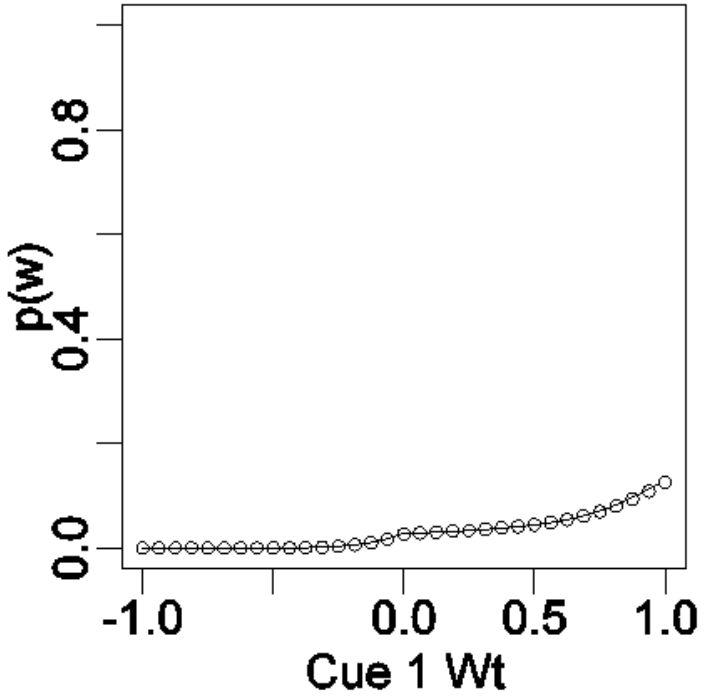
Current Uncertainty: 0.789

Probe: 1 0 => EU: 0.752

Probe: 0 1 => EU: 0.752







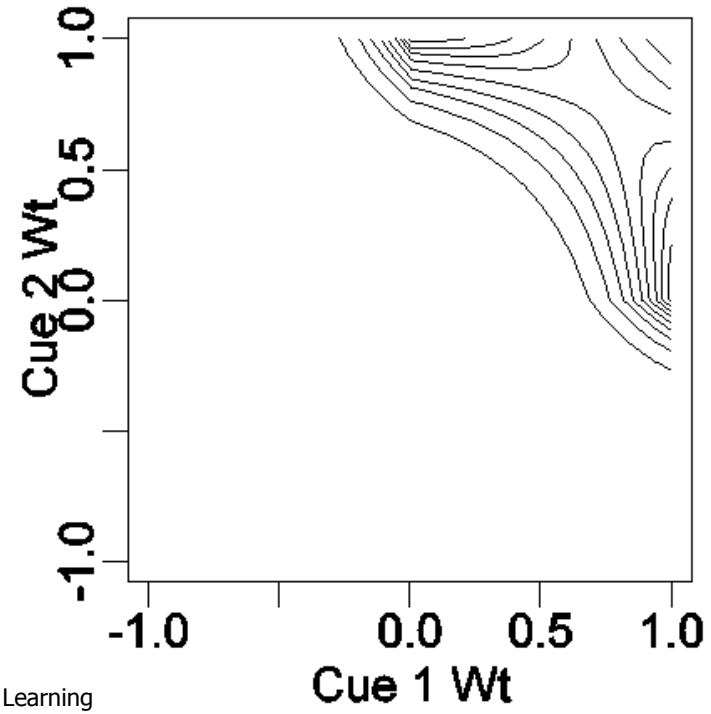
Training: BackwardBlocking

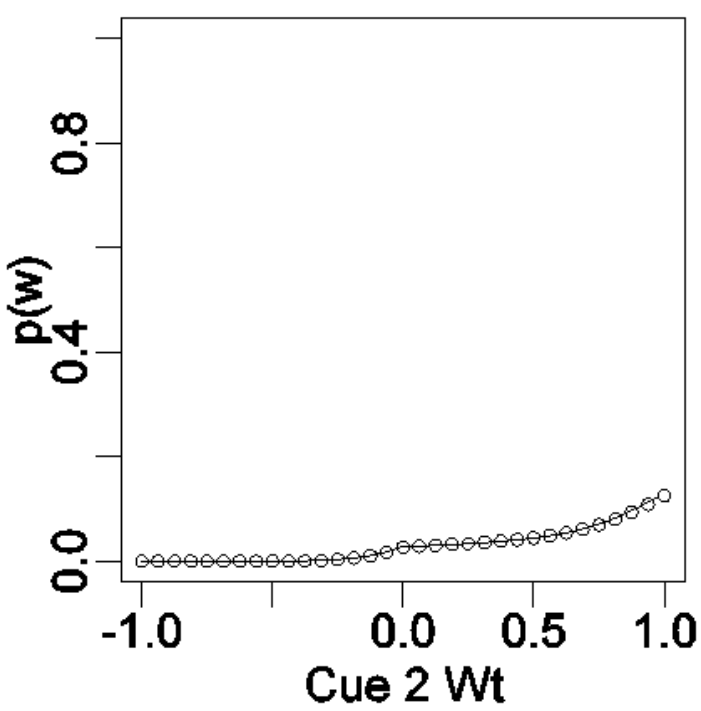
Beginning of Trial 8

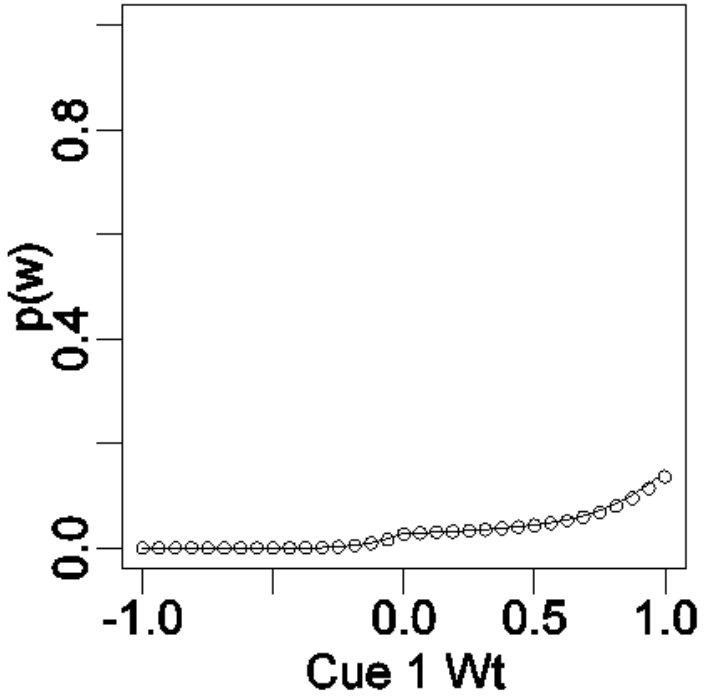
Current Uncertainty: 0.776

Probe: 1 0 => EU: 0.739

Probe: 0 1 => EU: 0.739







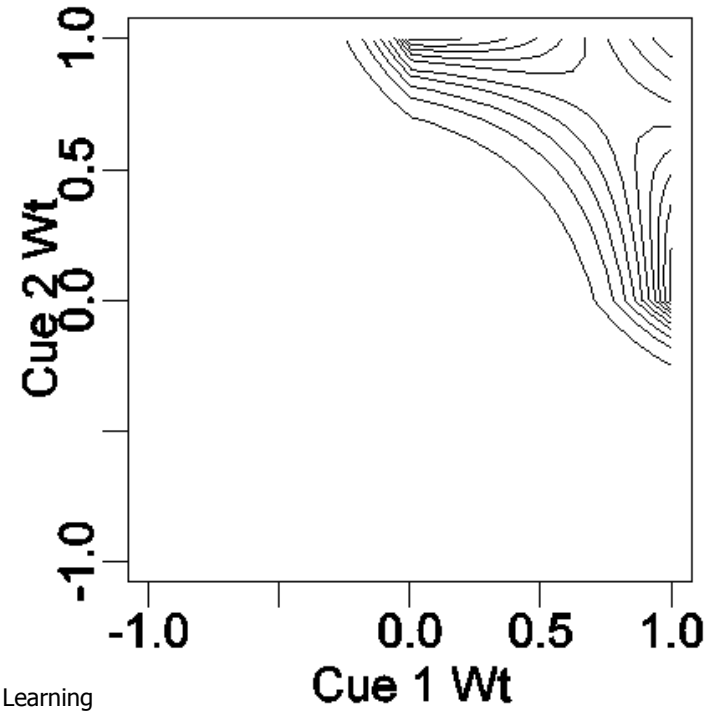
Training: BackwardBlocking

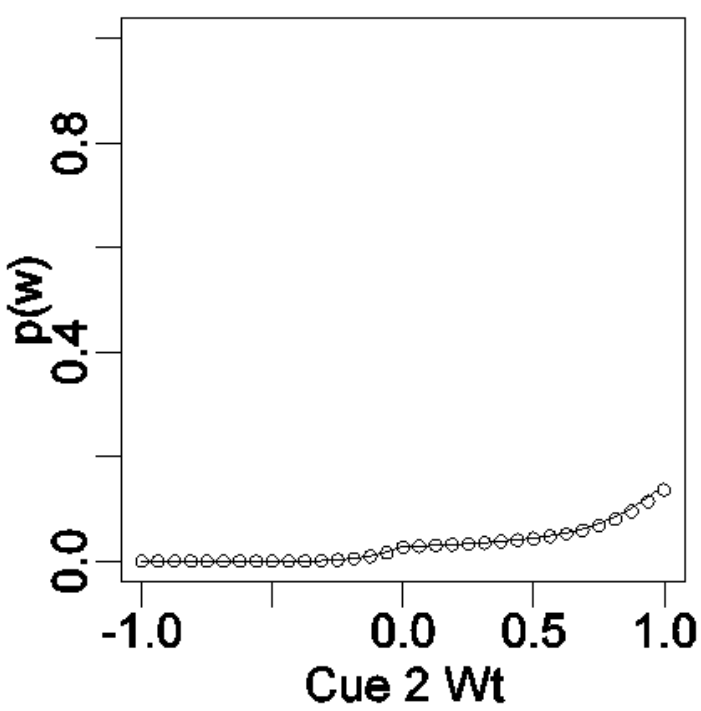
Beginning of Trial 9

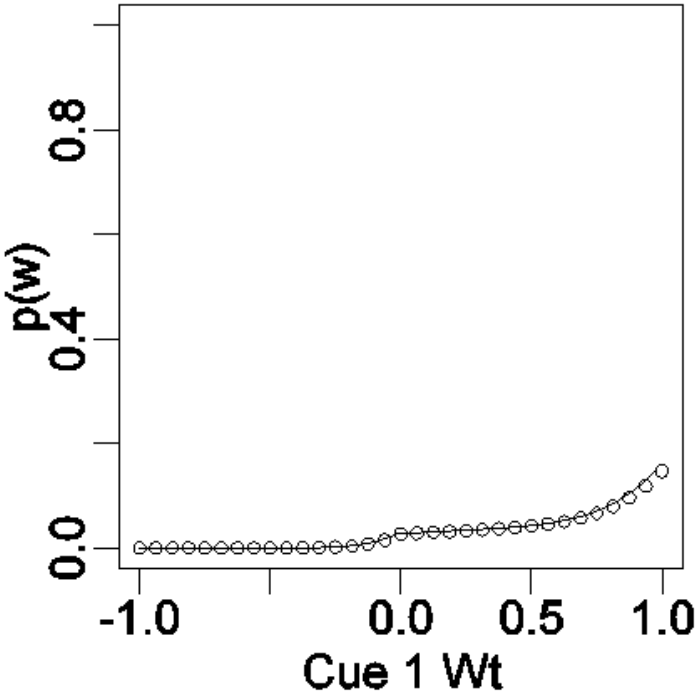
Current Uncertainty: 0.765

Probe: 1 0 => EU: 0.728

Probe: 0 1 => EU: 0.728







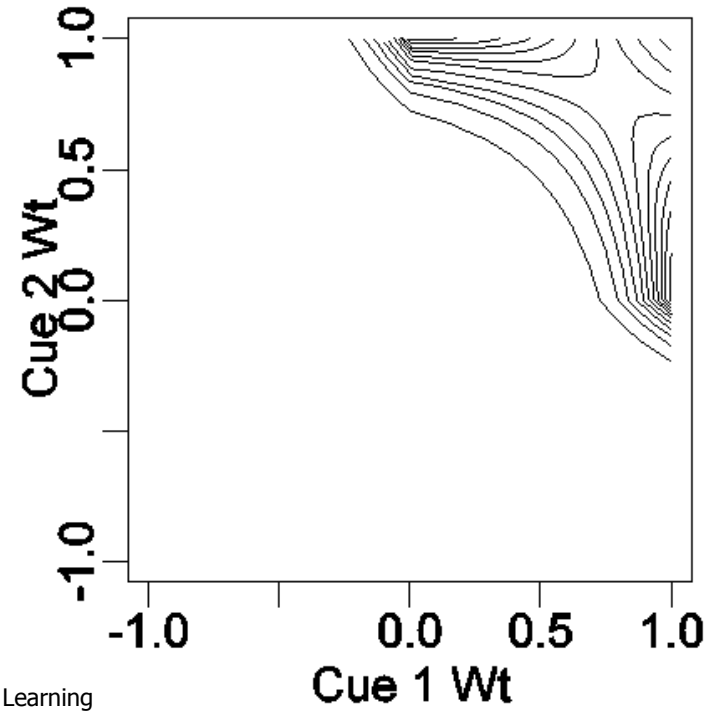
Training: BackwardBlocking

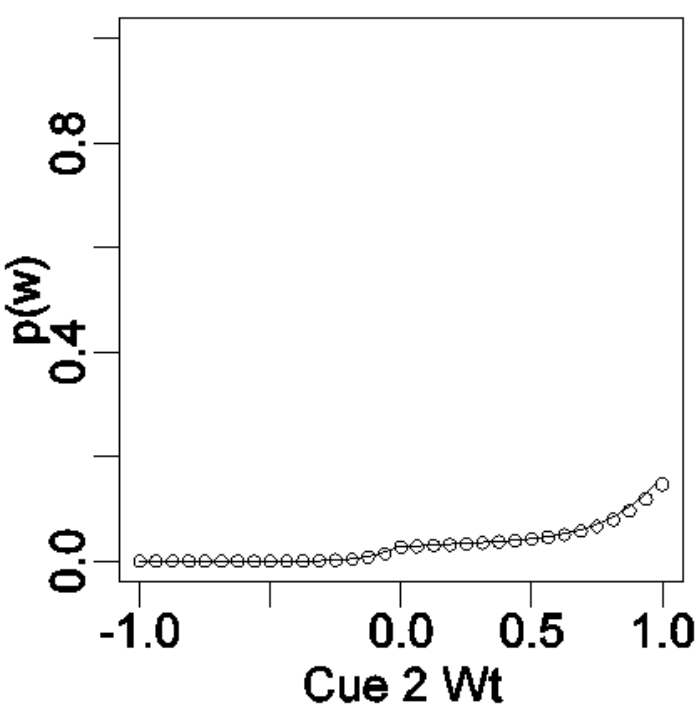
Beginning of Trial 10

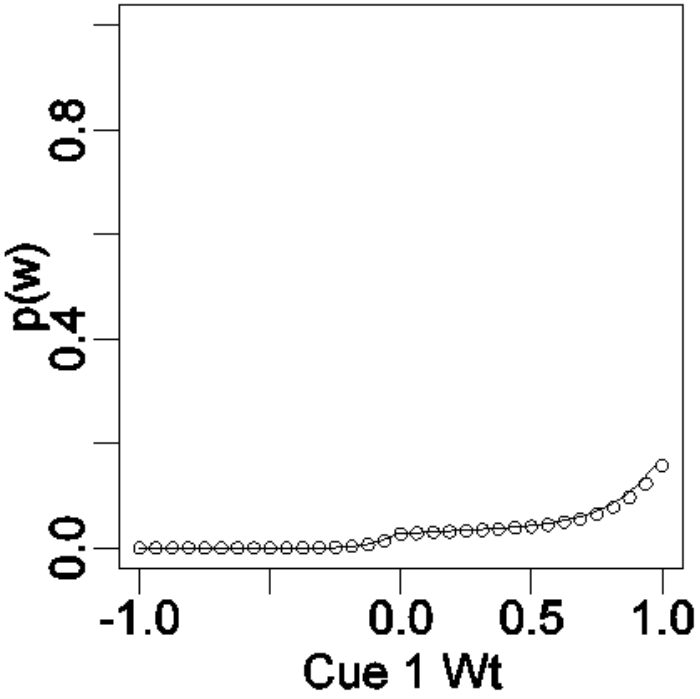
Current Uncertainty: 0.755

Probe: 1 0 => EU: 0.718

Probe: 0 1 => EU: 0.718







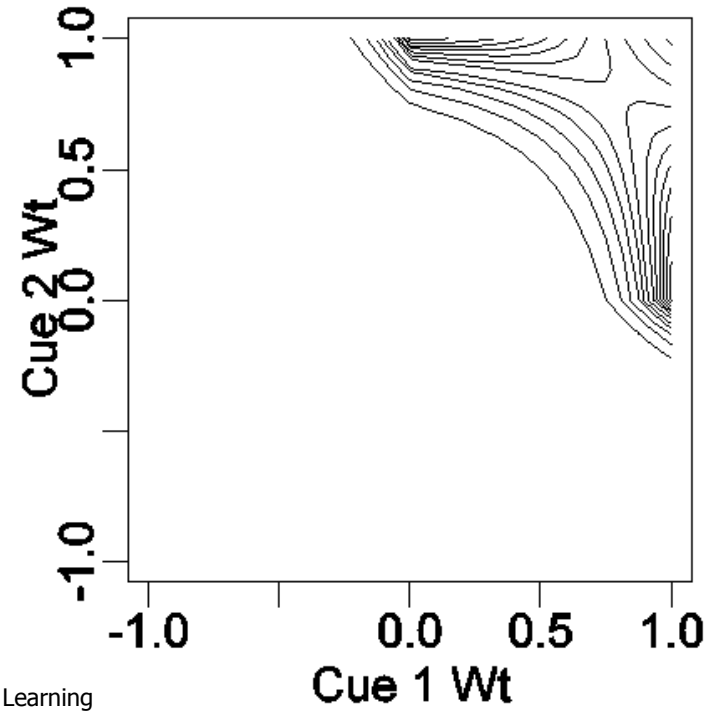
Training: BackwardBlocking

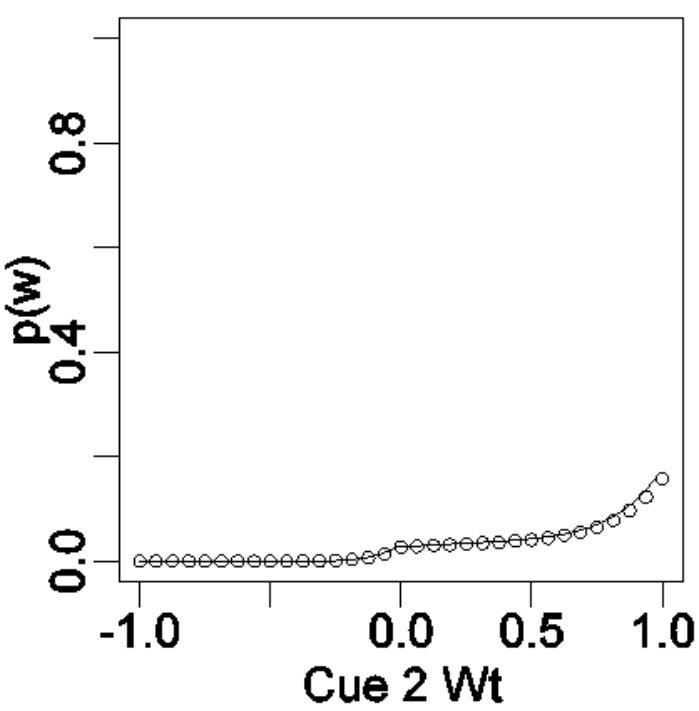
Beginning of Trial 11

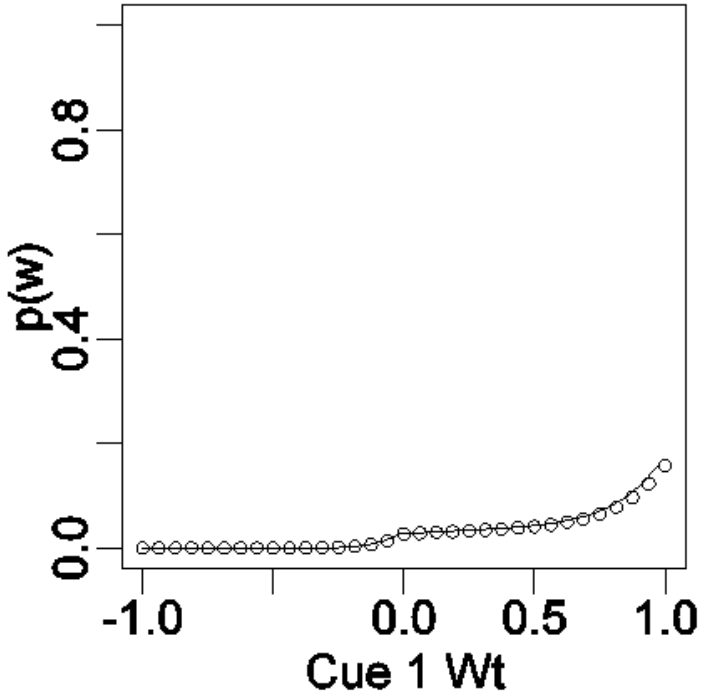
Current Uncertainty: 0.745

Probe: 1 0 => EU: 0.708

Probe: 0 1 => EU: 0.708







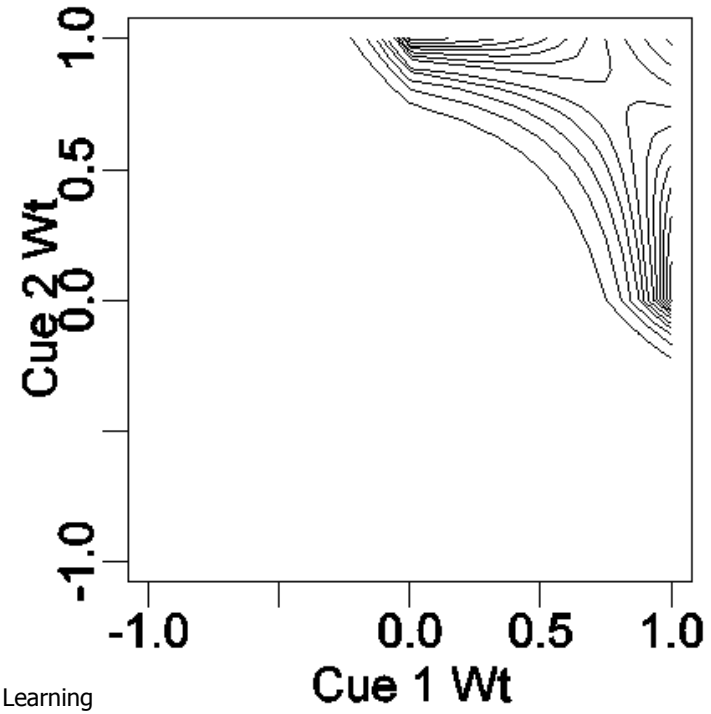
Training: BackwardBlocking

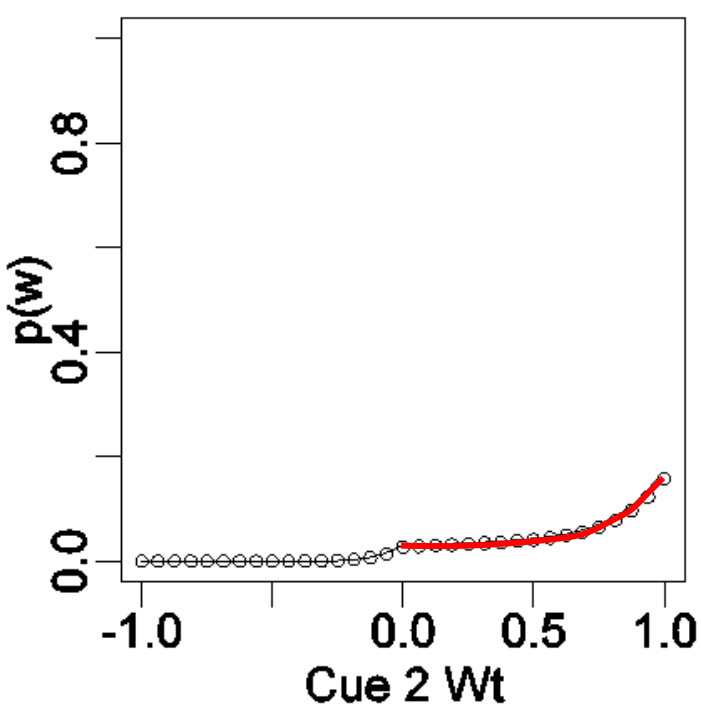
Beginning of Trial 11

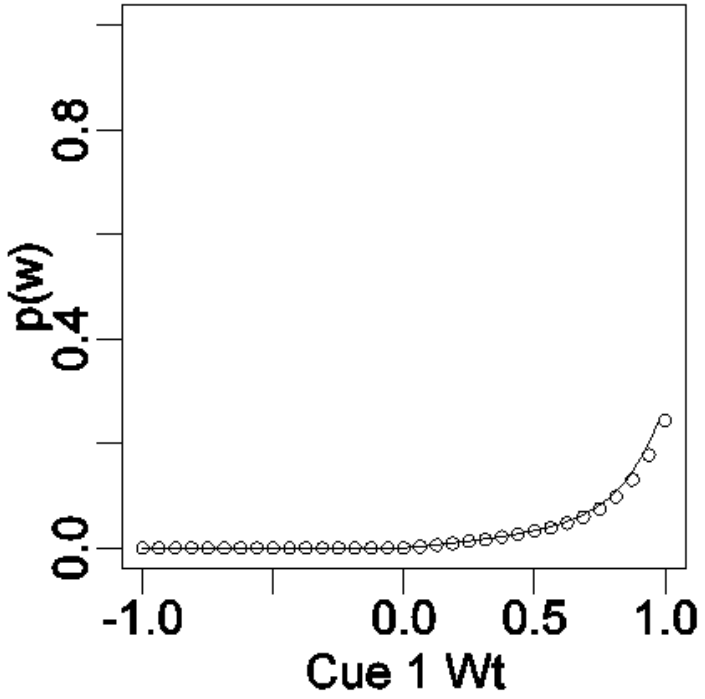
Current Uncertainty: 0.745

Probe: 1 0 => EU: 0.708

Probe: 0 1 => EU: 0.708







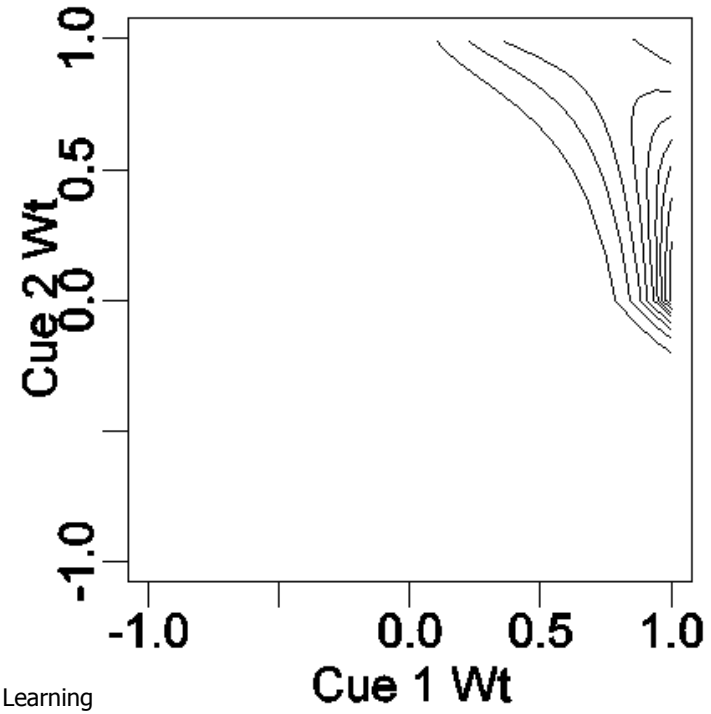
Training: BackwardBlocking

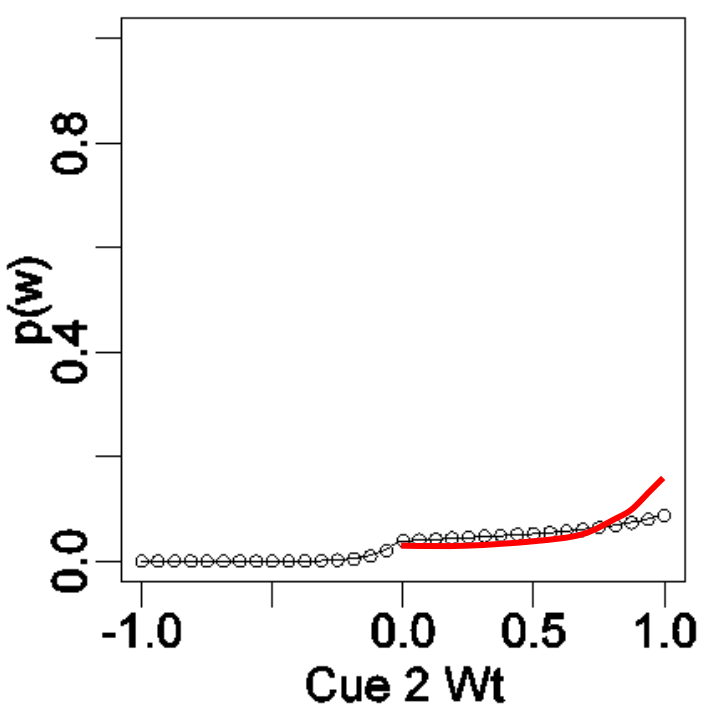
Beginning of Trial 12

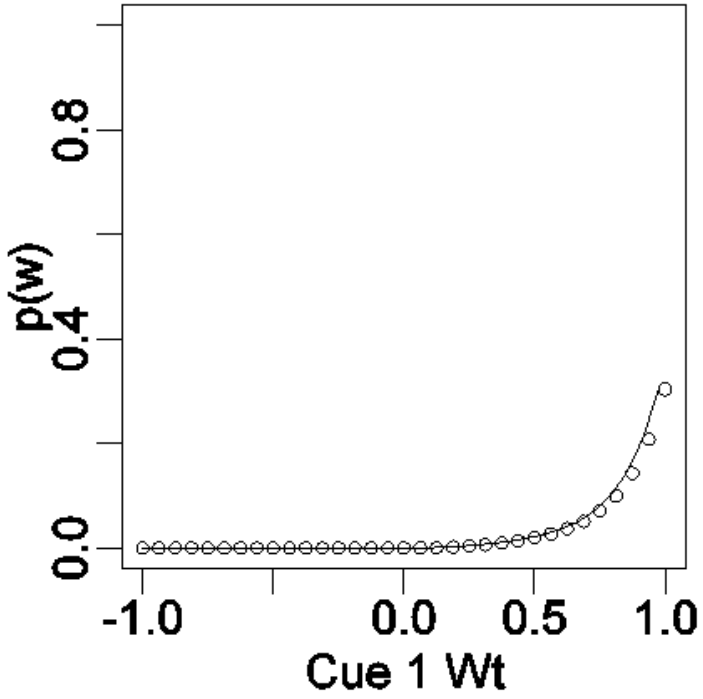
Current Uncertainty: 0.715

Probe: 1 0 => EU: 0.694

Probe: 0 1 => EU: 0.679







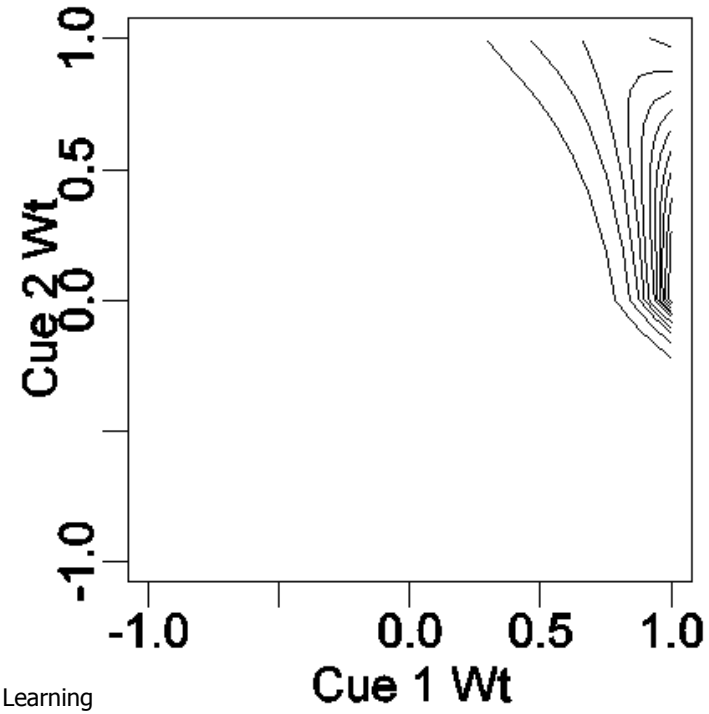
Training: BackwardBlocking

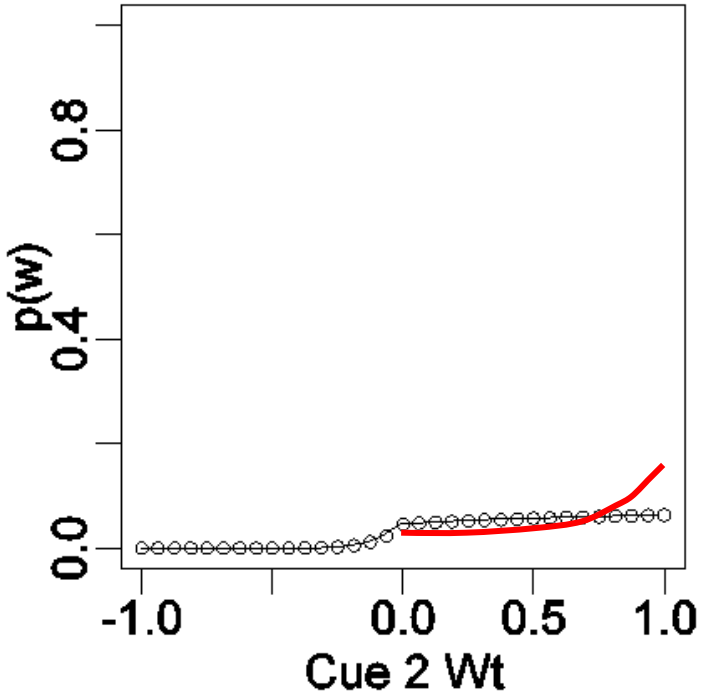
Beginning of Trial 13

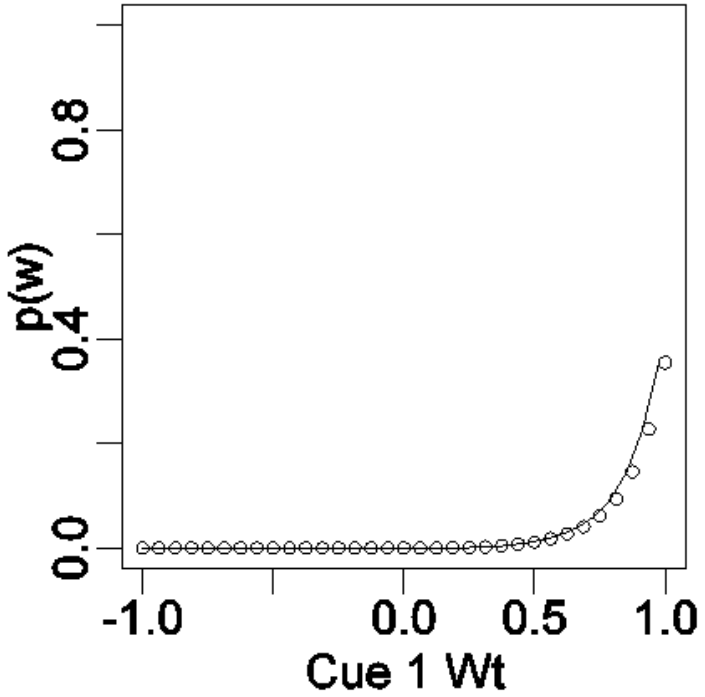
Current Uncertainty: 0.689

Probe: 1 0 => EU: 0.674

Probe: 0 1 => EU: 0.655







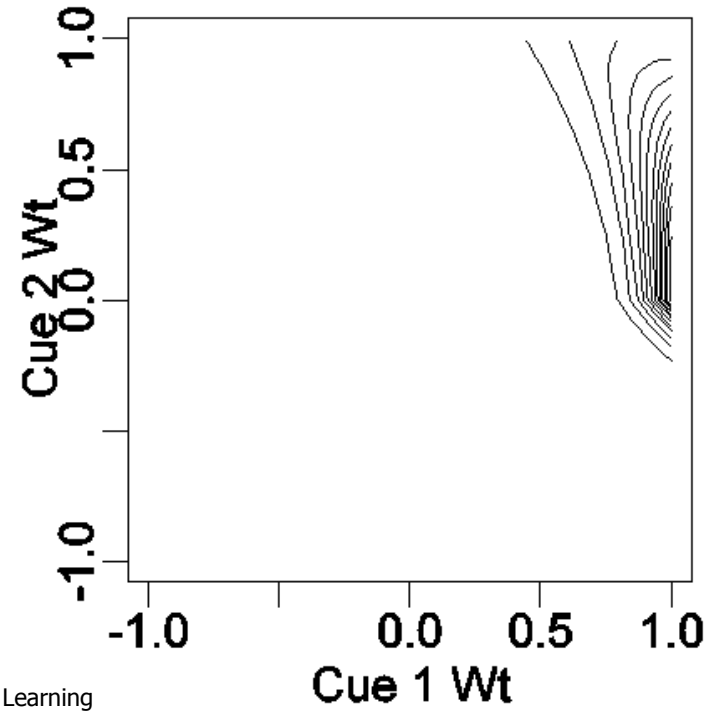
Training: BackwardBlocking

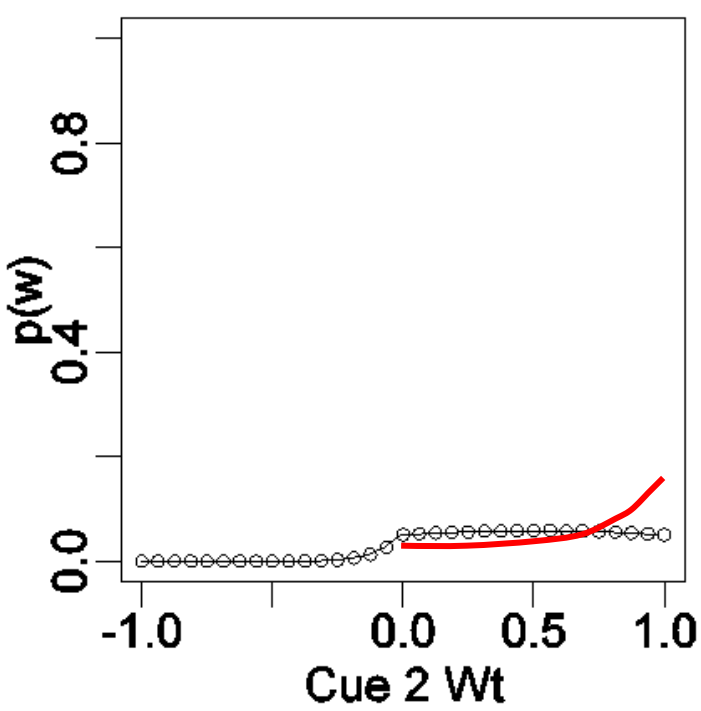
Beginning of Trial 14

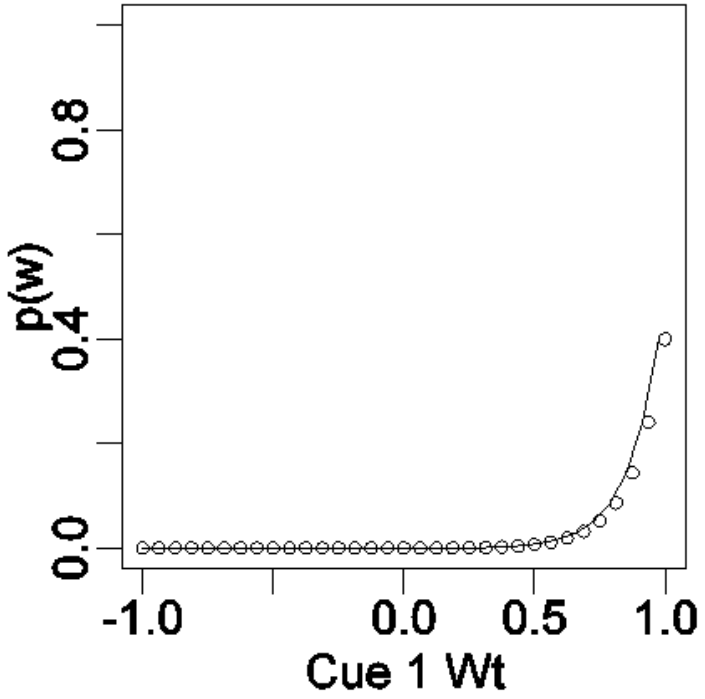
Current Uncertainty: 0.668

Probe: 1 0 => EU: 0.655

Probe: 0 1 => EU: 0.634







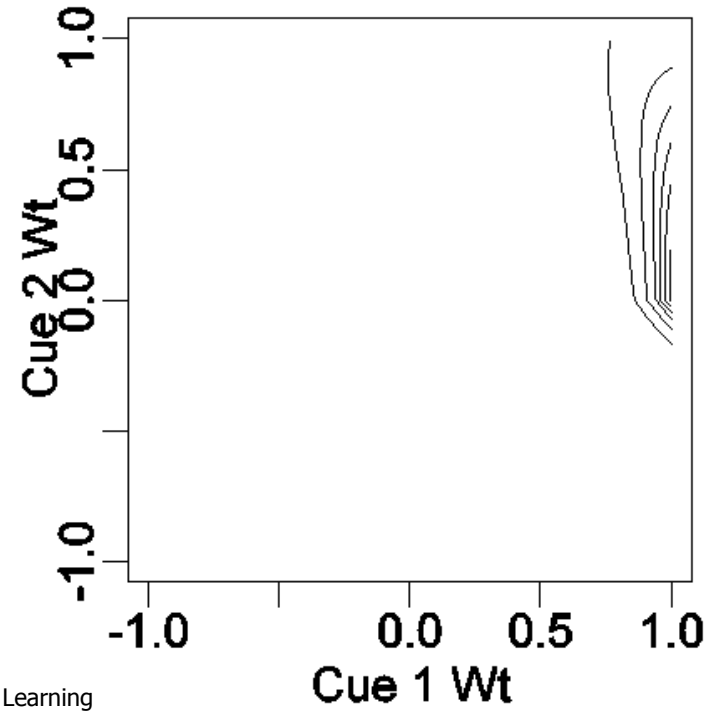
Training: BackwardBlocking

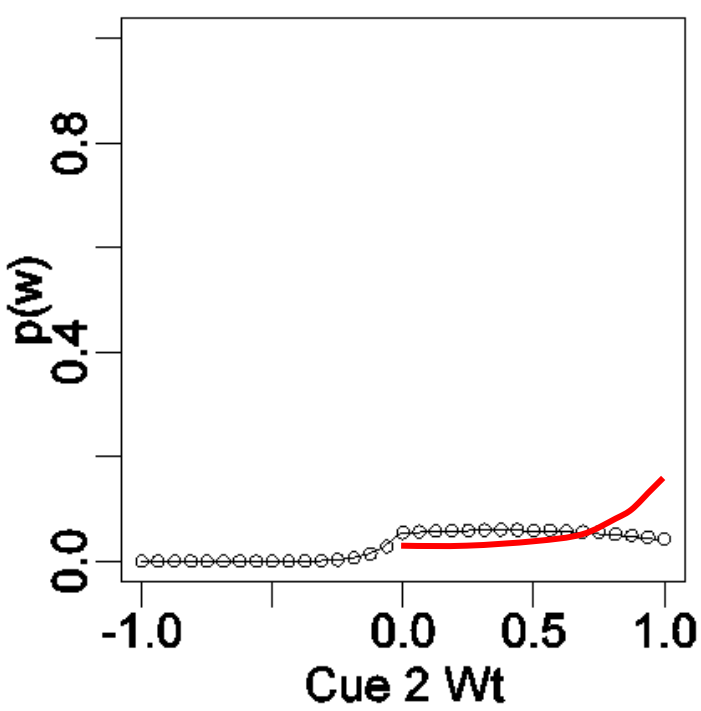
Beginning of Trial 15

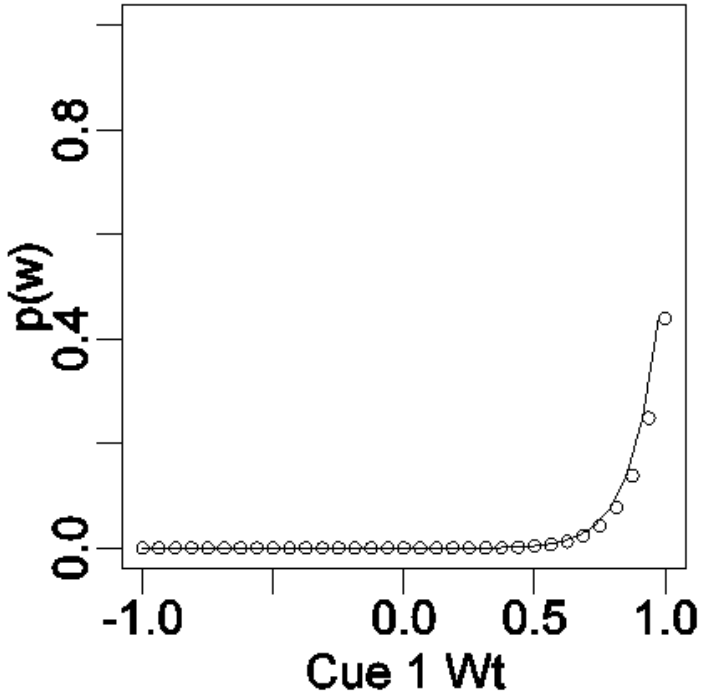
Current Uncertainty: 0.649

Probe: 1 0 => EU: 0.638

Probe: 0 1 => EU: 0.616







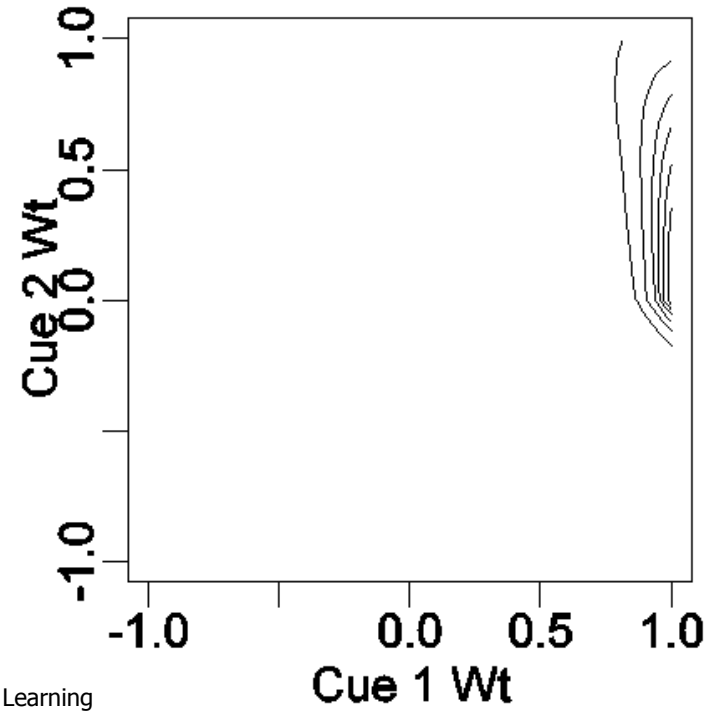
Training: BackwardBlocking

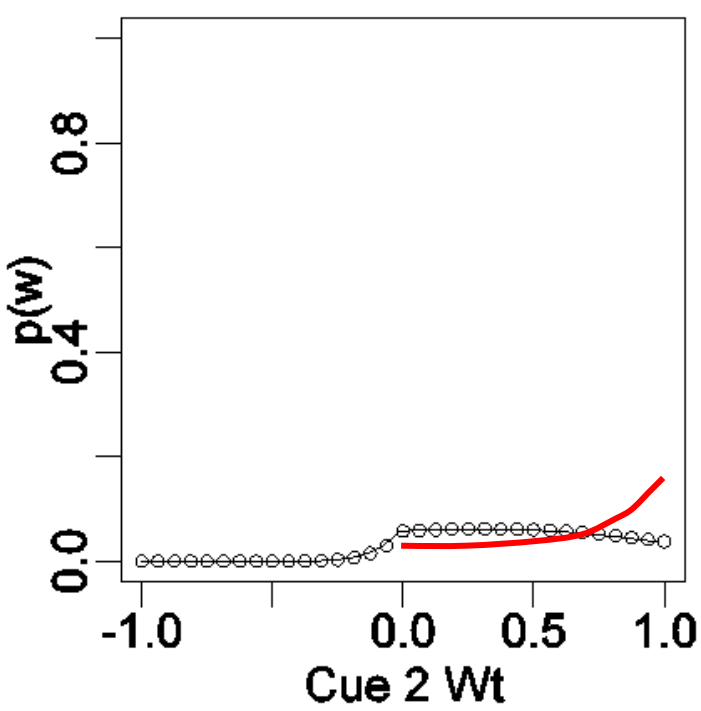
Beginning of Trial 16

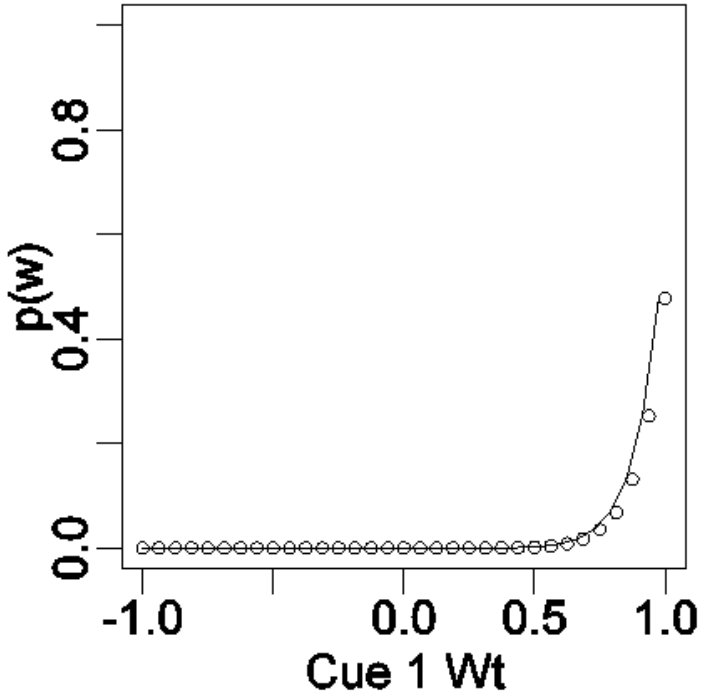
Current Uncertainty: 0.633

Probe: 1 0 => EU: 0.623

Probe: 0 1 => EU: 0.6







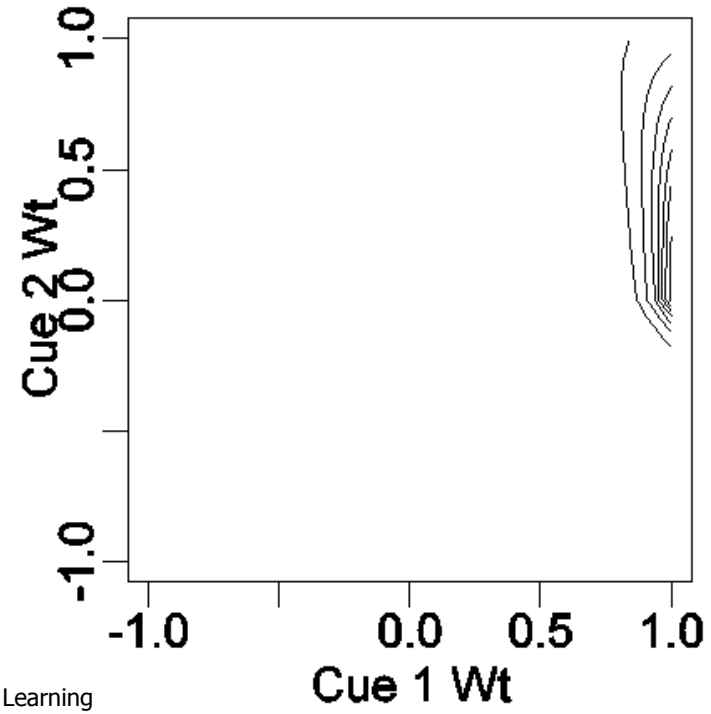
Training: BackwardBlocking

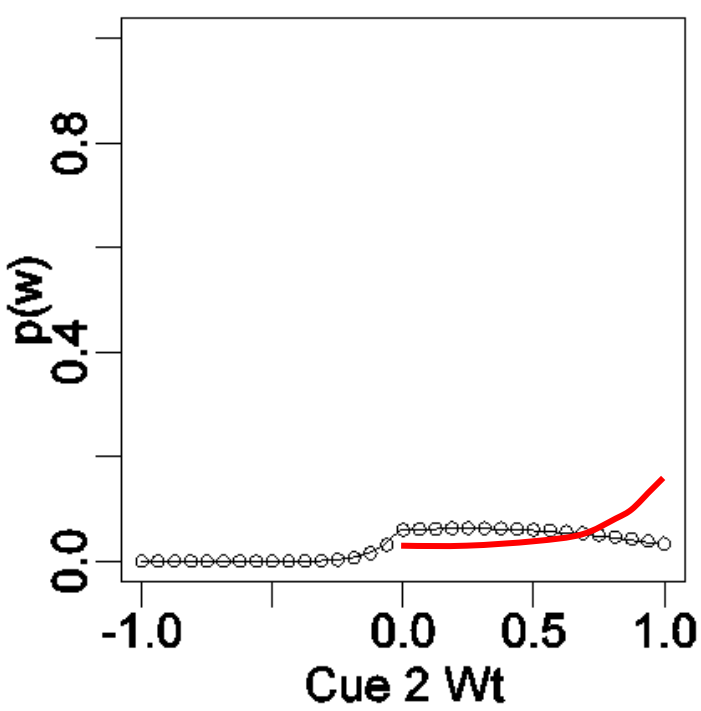
Beginning of Trial 17

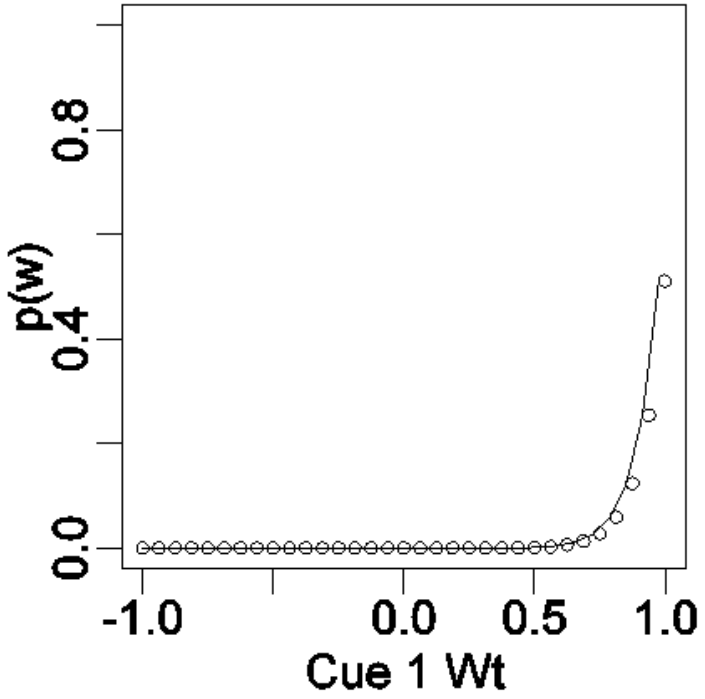
Current Uncertainty: 0.619

Probe: 1 0 => EU: 0.61

Probe: 0 1 => EU: 0.586







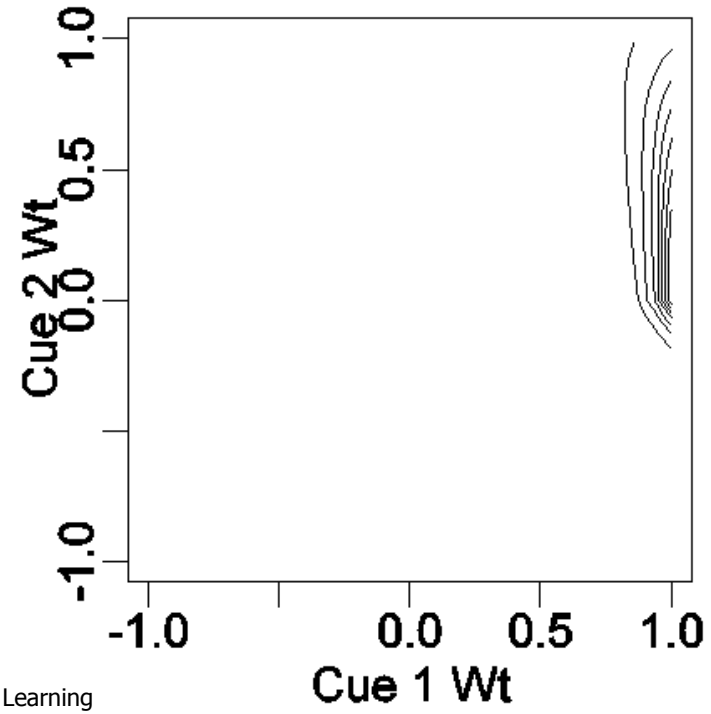
Training: BackwardBlocking

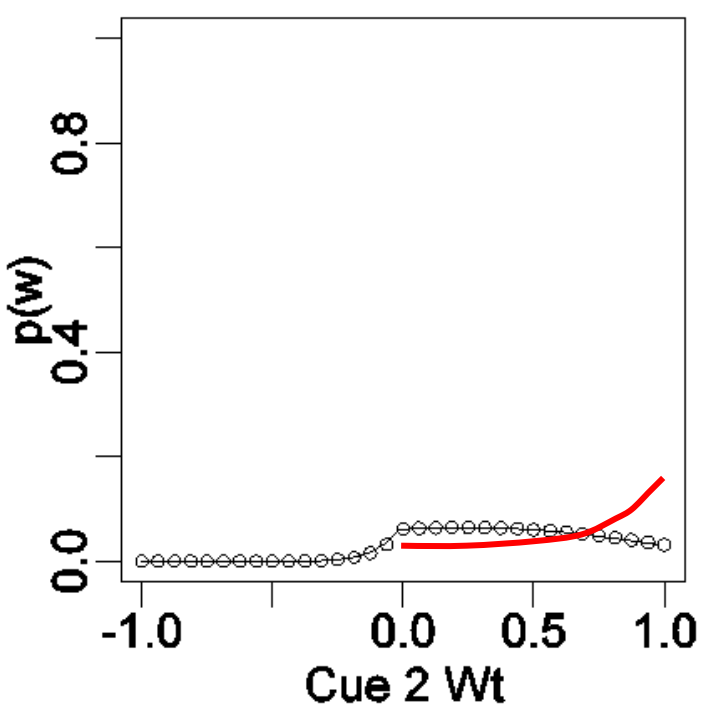
Beginning of Trial 18

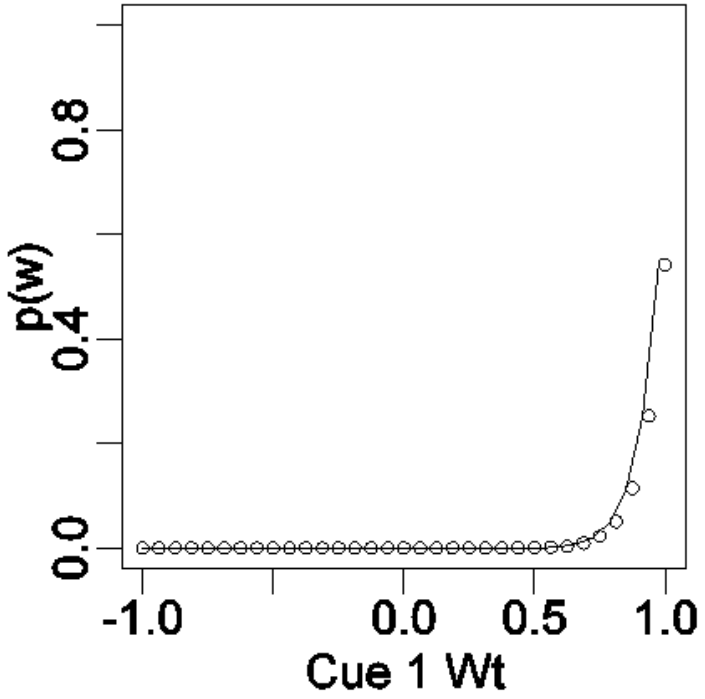
Current Uncertainty: 0.606

Probe: 1 0 => EU: 0.598

Probe: 0 1 => EU: 0.574







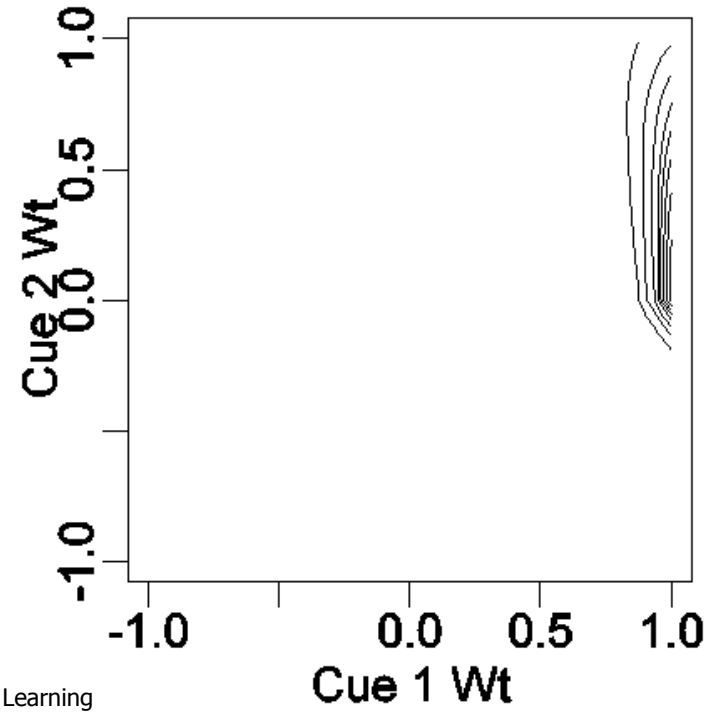
Training: BackwardBlocking

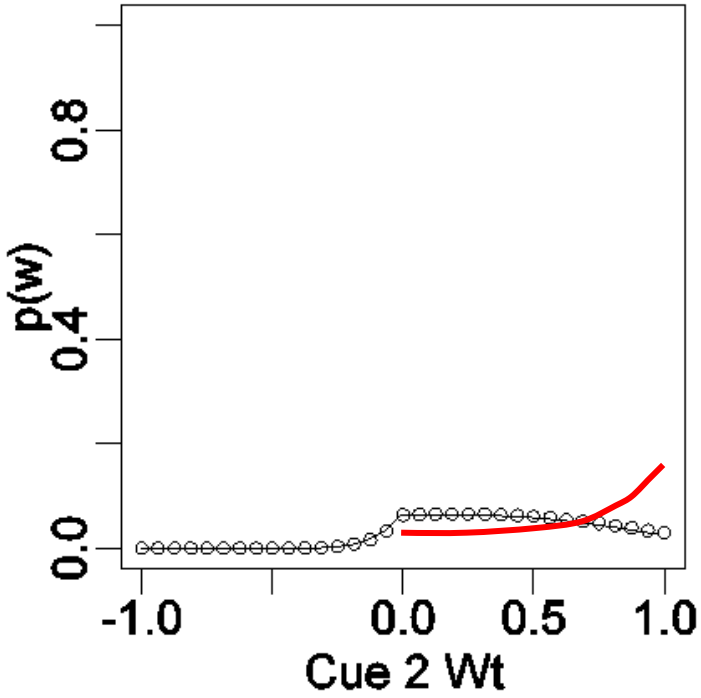
Beginning of Trial 19

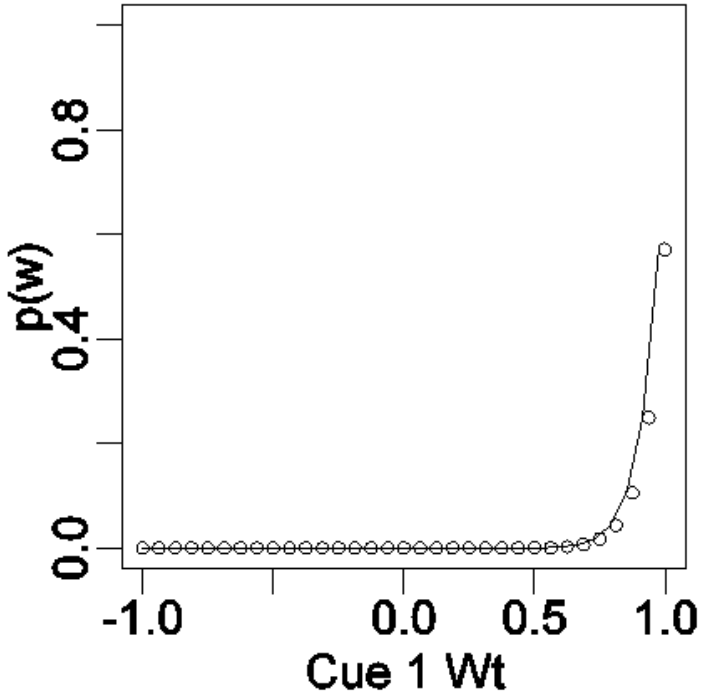
Current Uncertainty: 0.594

Probe: 1 0 => EU: 0.587

Probe: 0 1 => EU: 0.563







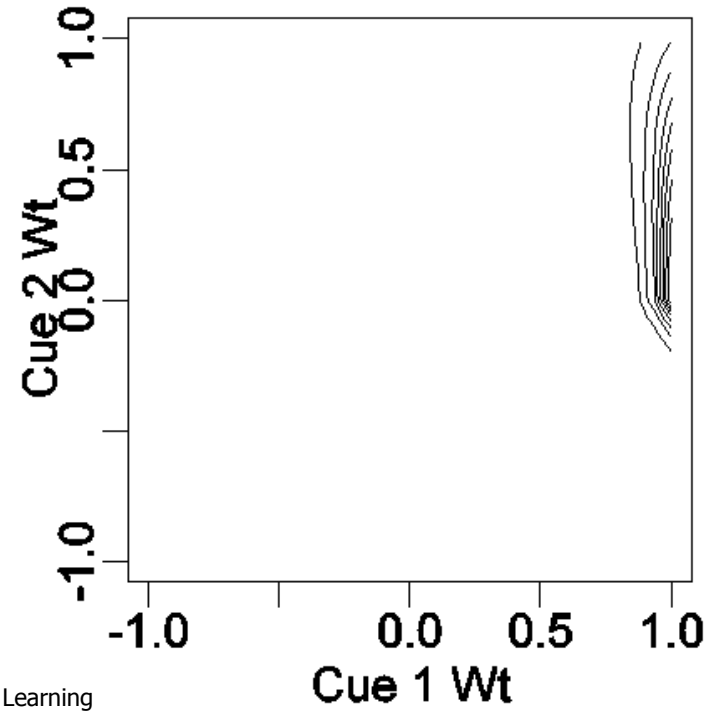
Training: BackwardBlocking

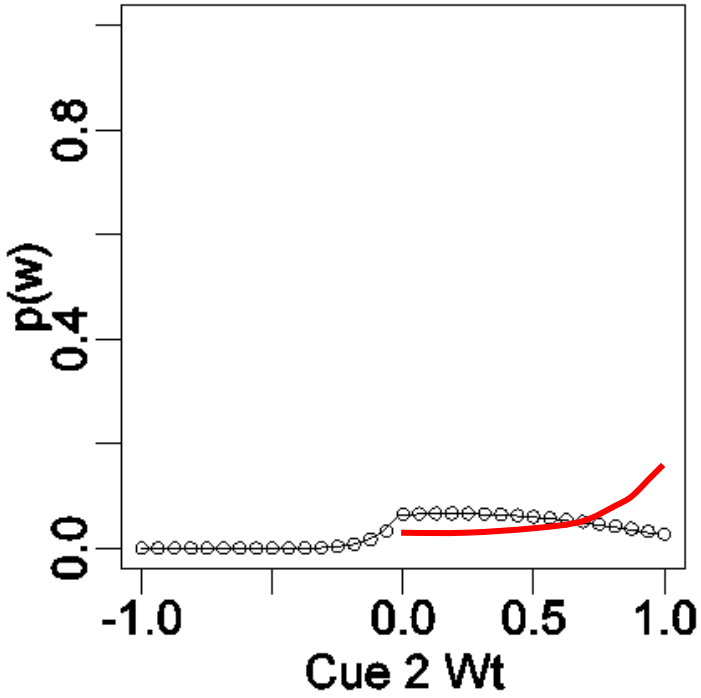
Beginning of Trial 20

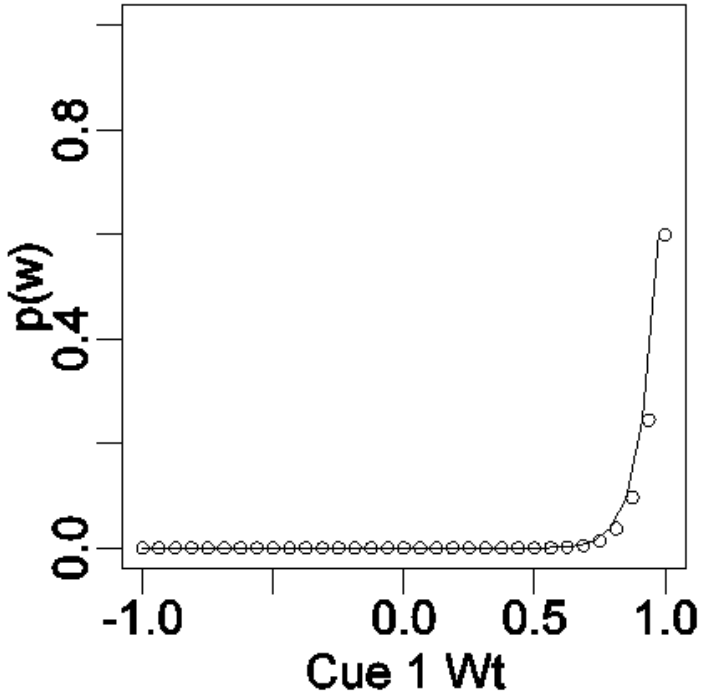
Current Uncertainty: 0.584

Probe: 1 0 => EU: 0.577

Probe: 0 1 => EU: 0.552







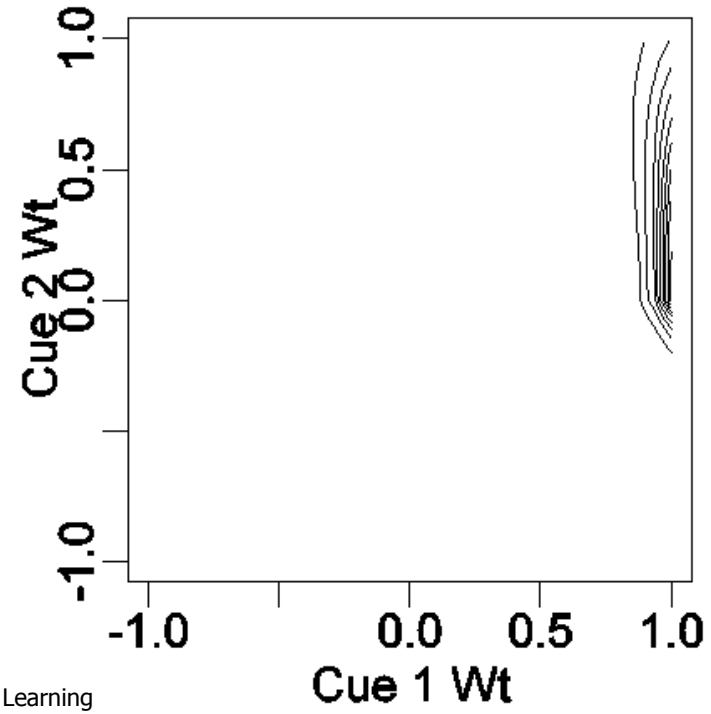
Training: BackwardBlocking

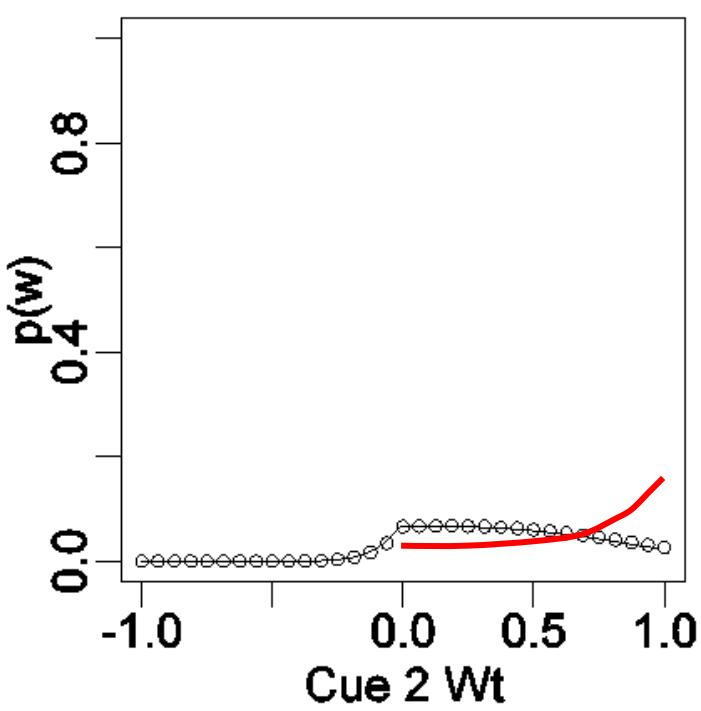
Beginning of Trial 21

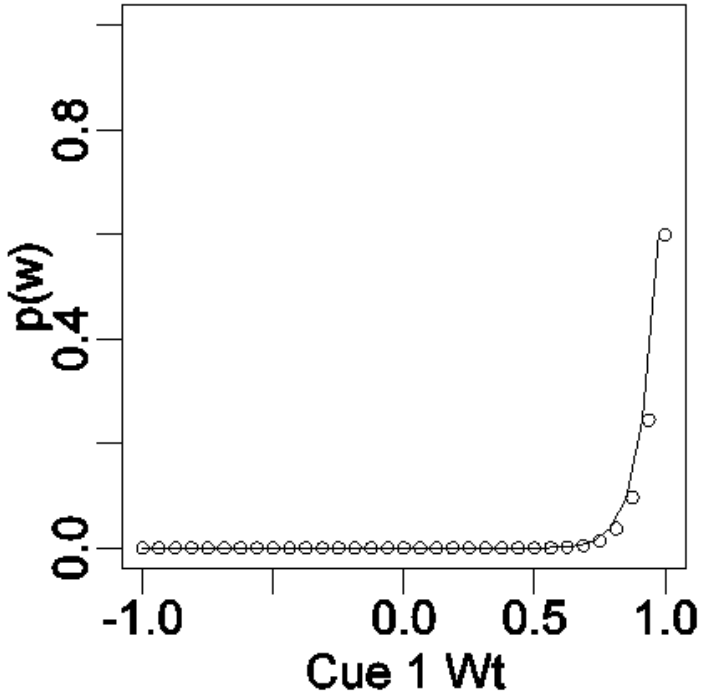
Current Uncertainty: 0.574

Probe: 1 0 => EU: 0.568

Probe: 0 1 => EU: 0.543



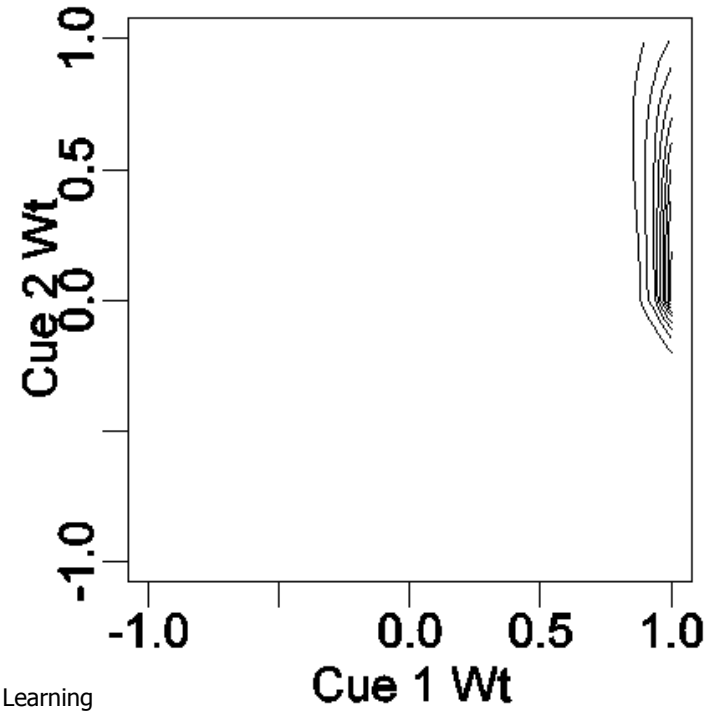


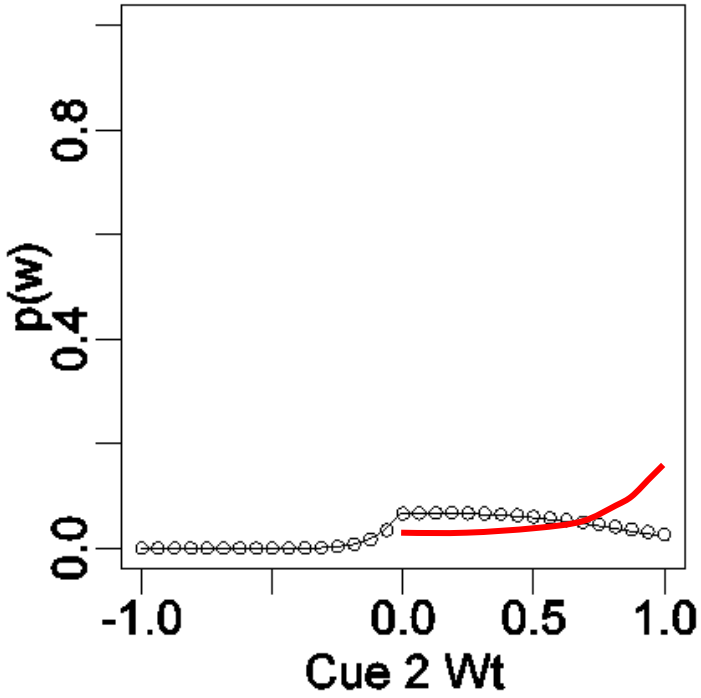


Training: BackwardBlocking

Beginning of Trial 21

Noisy logic gate *does* show backward blocking.





Design of the (rest of the) Talk:
2 x 2 x 3 factorial

Bayesian models of associative learning:

- Kalman filter.
- Noisy-logic gate.

Goals for active learning:

- Minimize expected uncertainty.
- Maximize expected maximum believability.

Training structures:

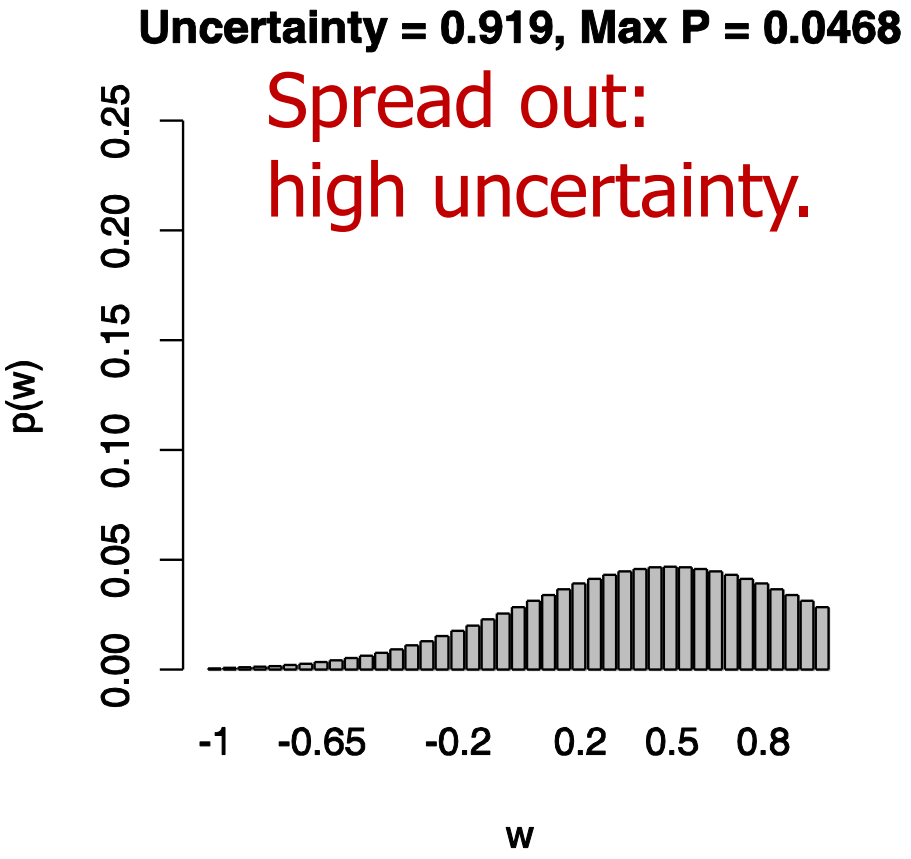
- Backward blocking.
- Blocking and reduced overshadowing.
- Ambiguous cue.

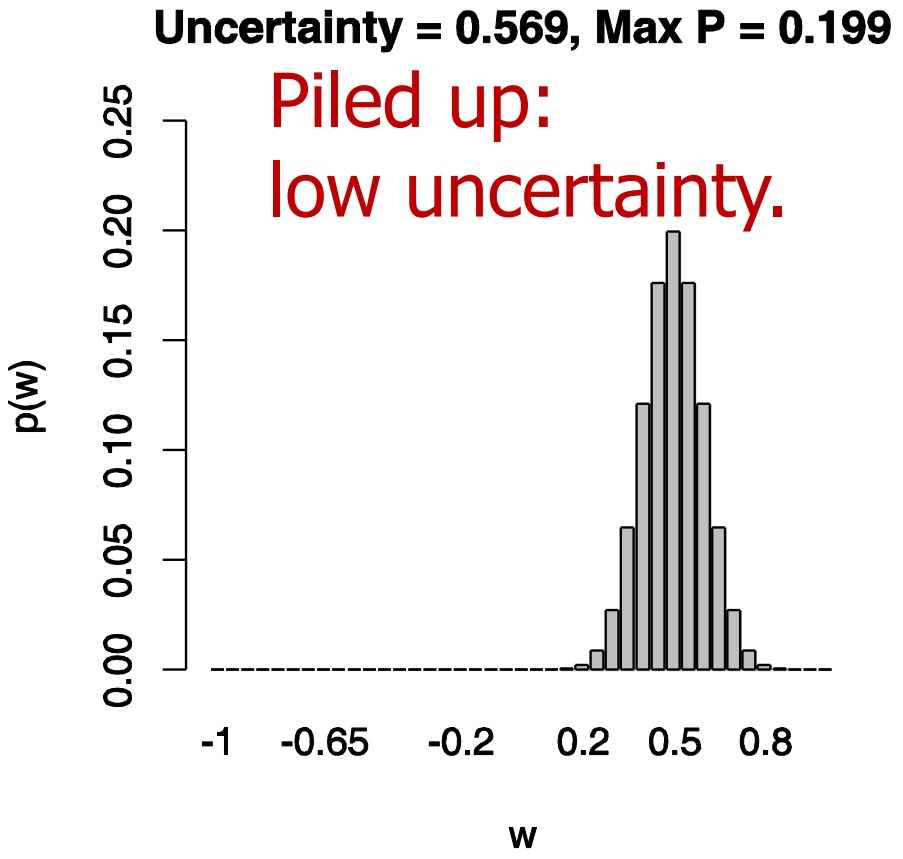
Active Learning after Backward Blocking

Phase	Frequency	Cue 1	Cue 2	Outcome
I	10	1 	1 	1 
II	10	1 	0	1 

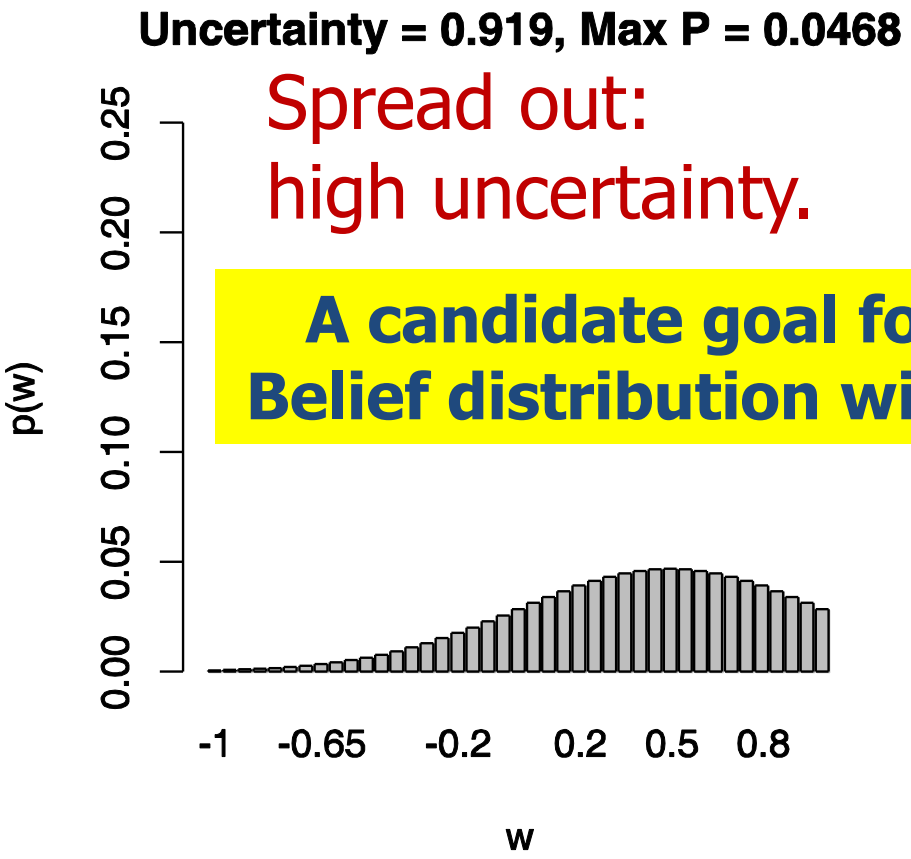
After being trained on backward blocking, your goal is to better determine which cues produce or prevent nausea. Which probe below do you prefer to test?

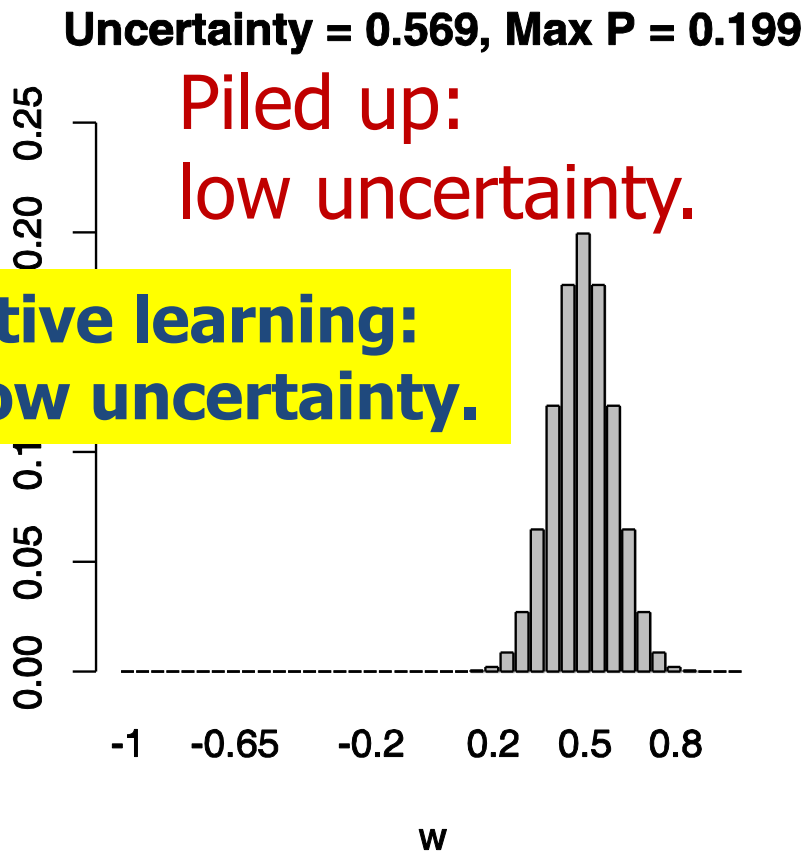
Uncertainty of a Belief Distribution
(a.k.a. entropy)





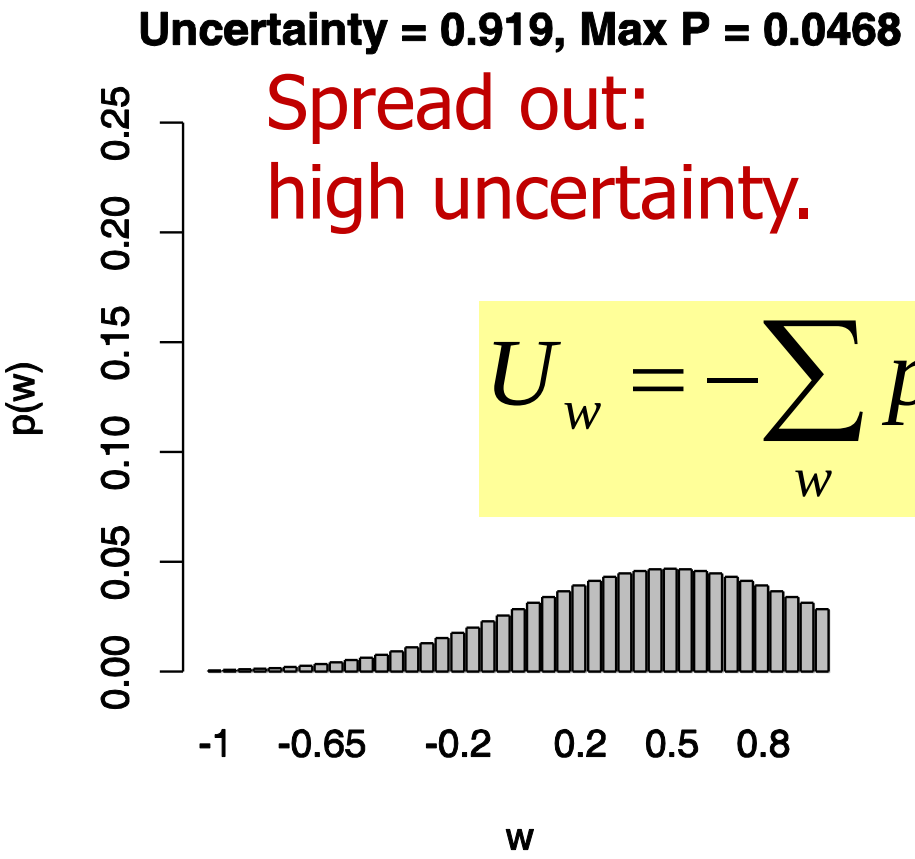
Uncertainty of a Belief Distribution
(a.k.a. entropy)

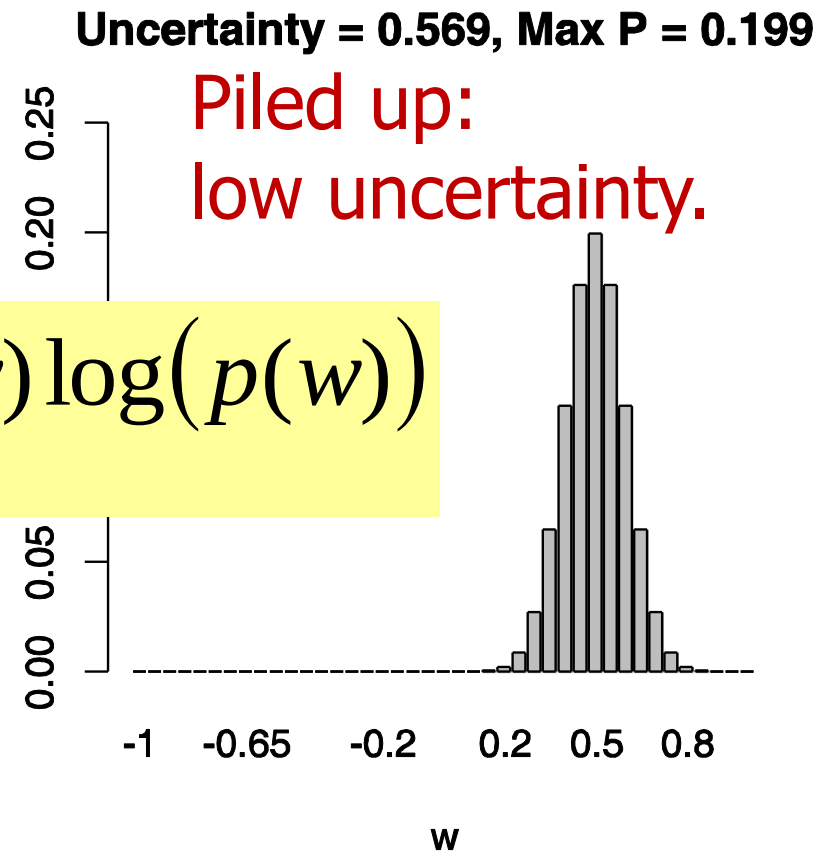




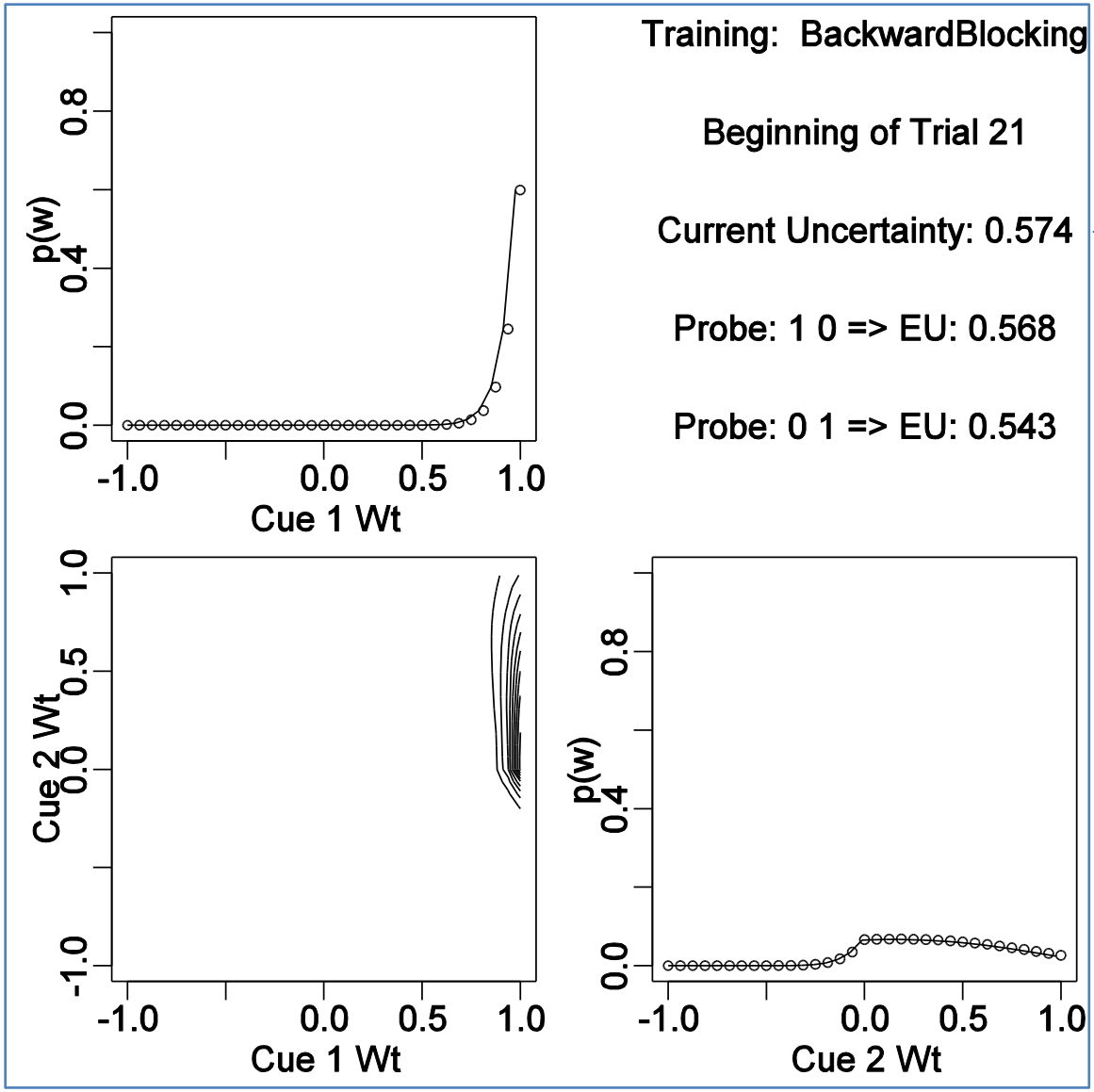
A candidate goal for active learning:
Belief distribution with low uncertainty.

Uncertainty of a Belief Distribution
(a.k.a. entropy)

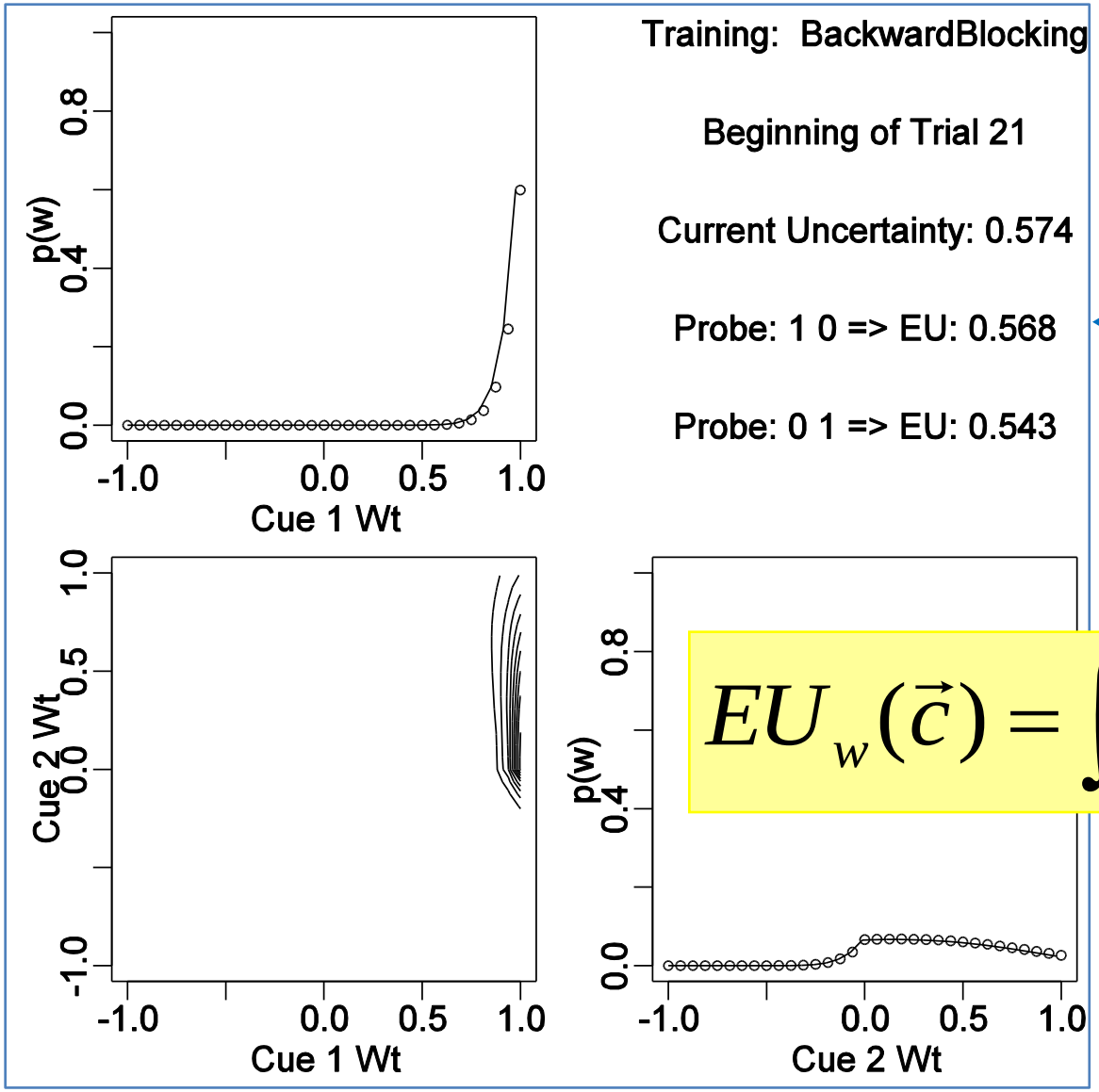




$$U_w = -\sum_w p(w)\log\big(p(w)\big)$$



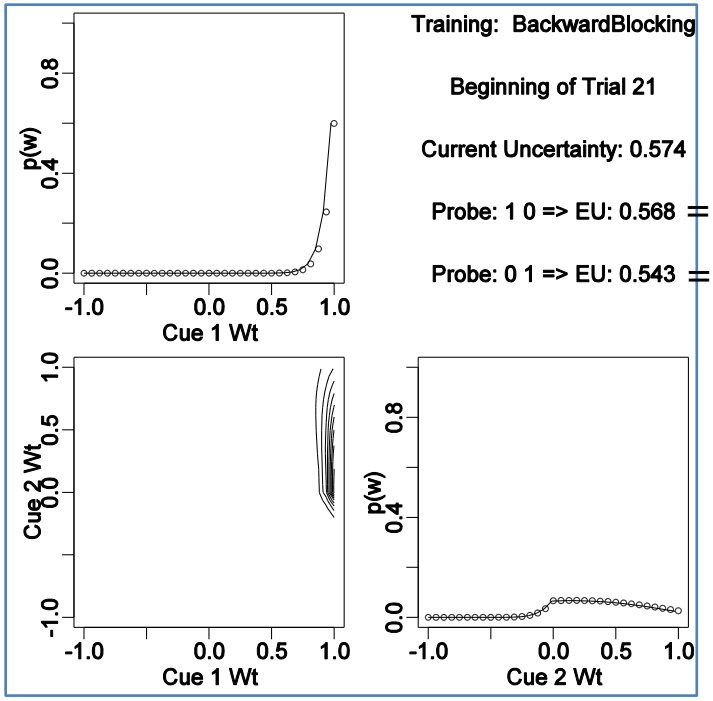
← Current Uncertainty is computed from current belief distribution.



For a candidate probe, we compute the *expected* uncertainty:

$$EU_w(\vec{c}) = \int da \; p(a \mid \vec{c}) \; U_{w|a,\vec{c}}$$

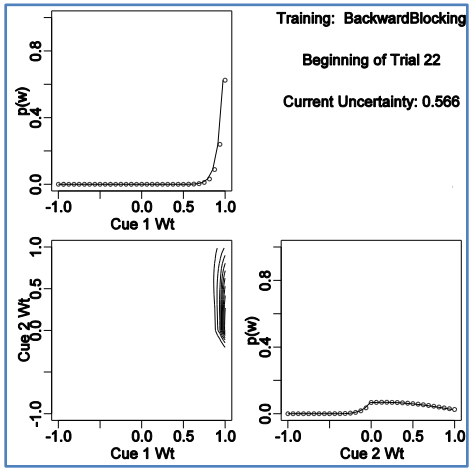
where $p(a \mid \vec{c}) = \int d\vec{w} \; p(a \mid \vec{c}, \vec{w}) \; p(\vec{w})$



Probe: 1 0

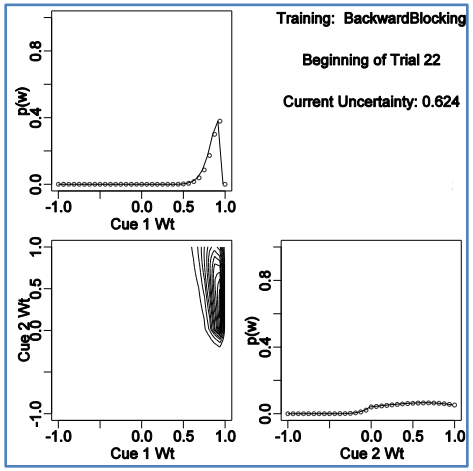
p = .96

outcome = 1



p = .04

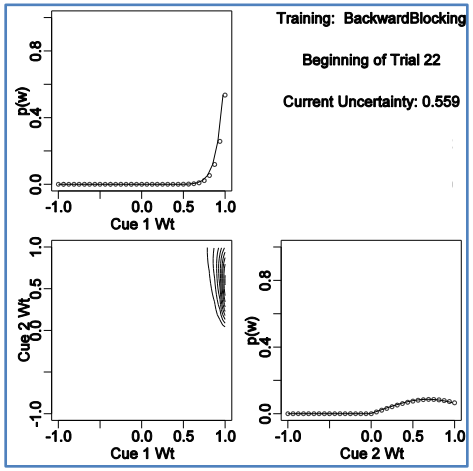
outcome = 0



Probe: 0 1

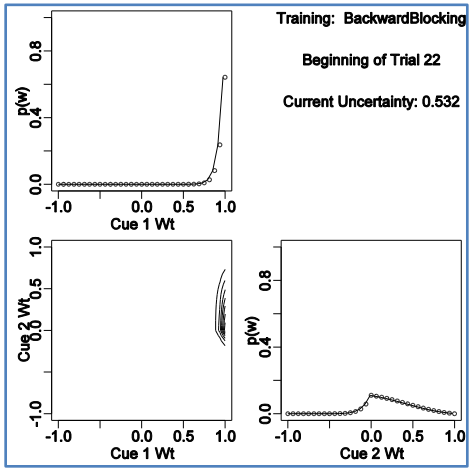
p = .40

outcome = 1



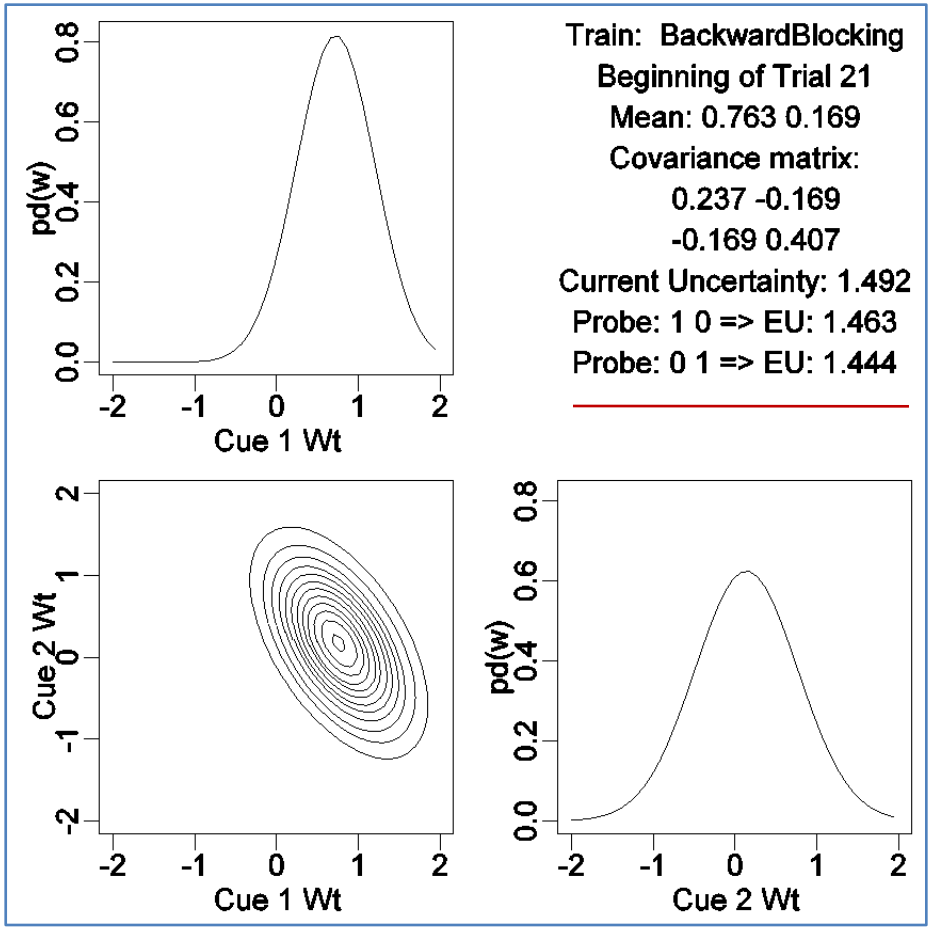
p = .60

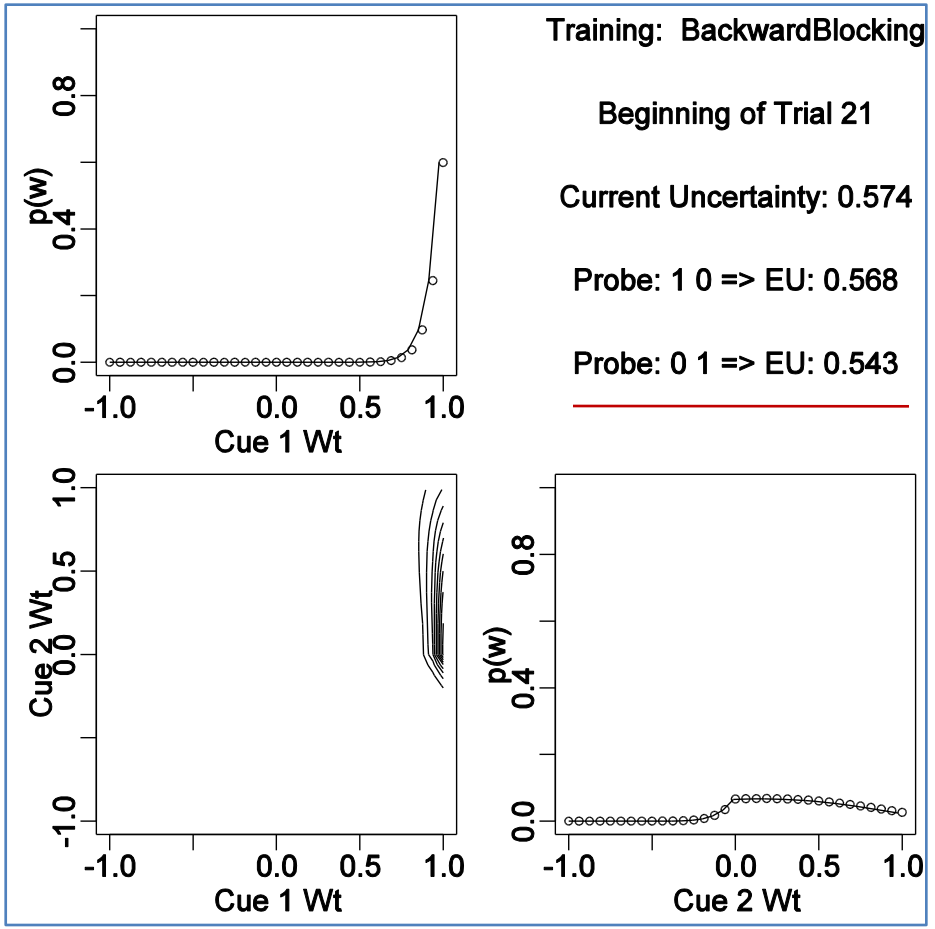
outcome = 0



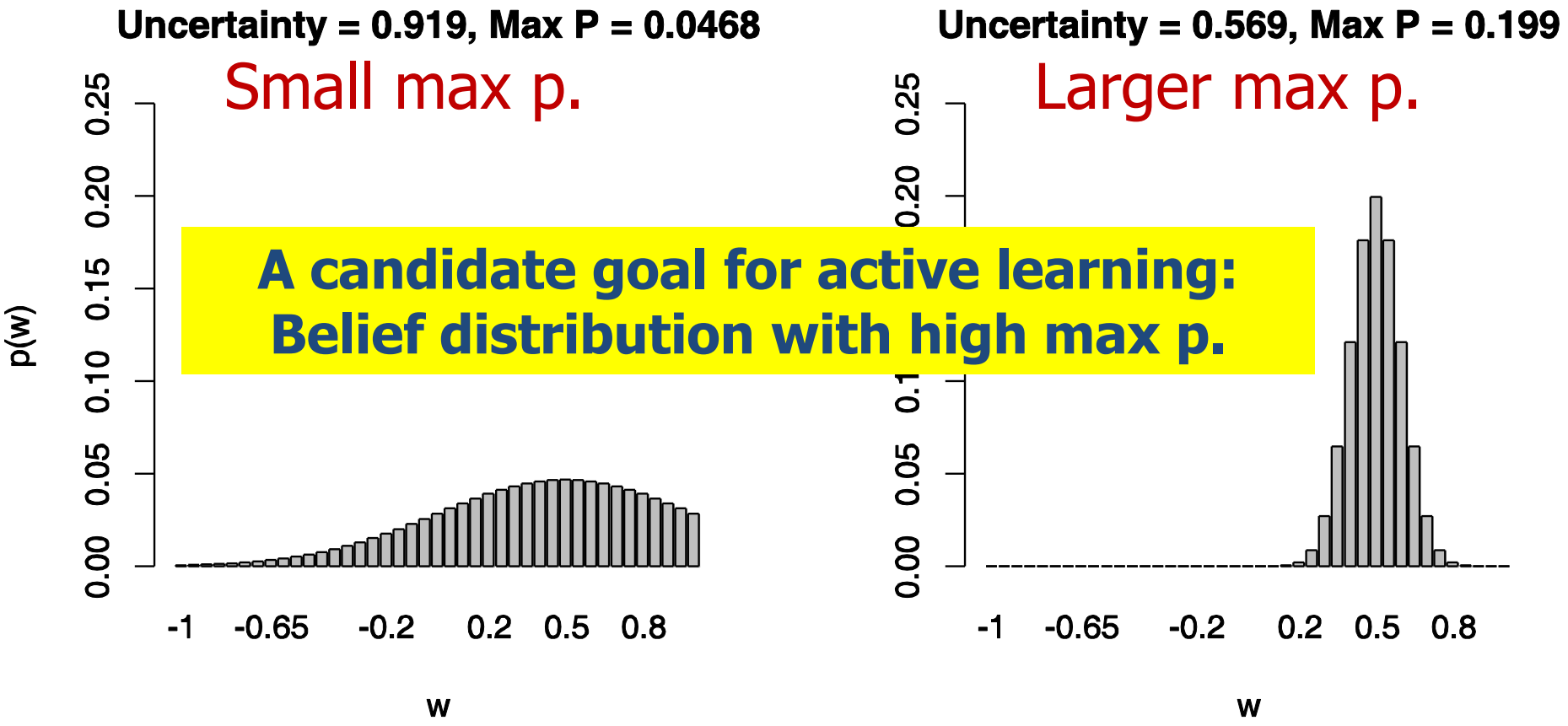
Active Learning after Backward Blocking

Both Kalman filter and noisy-logic gate match human intuition when minimizing Expected Uncertainty.

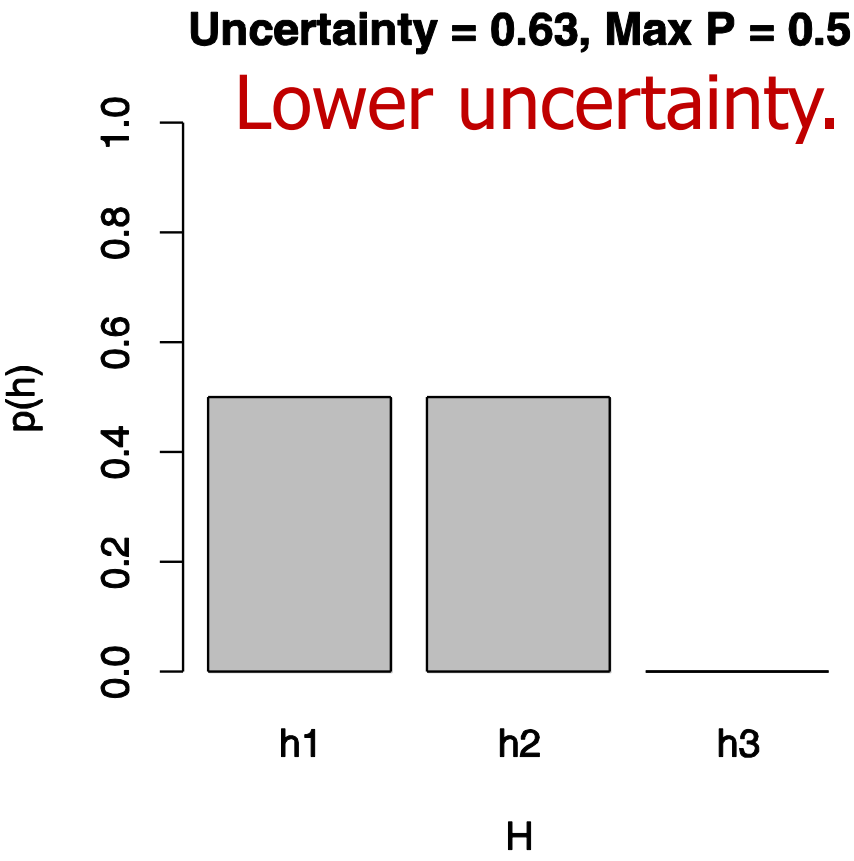


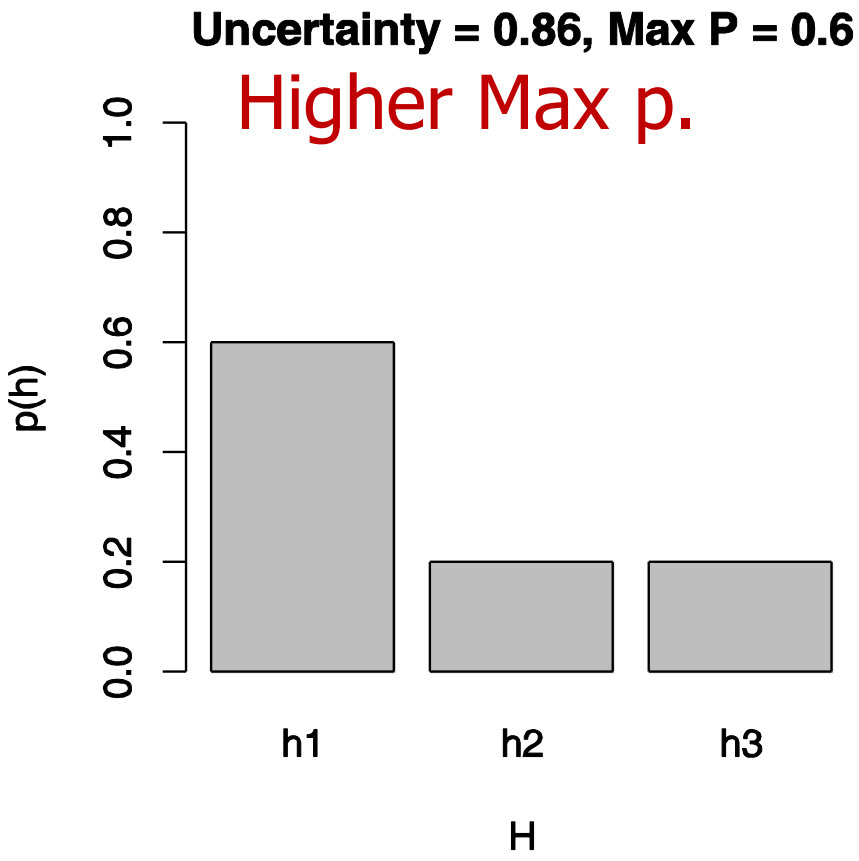


Maximum Probability of a Belief Distribution

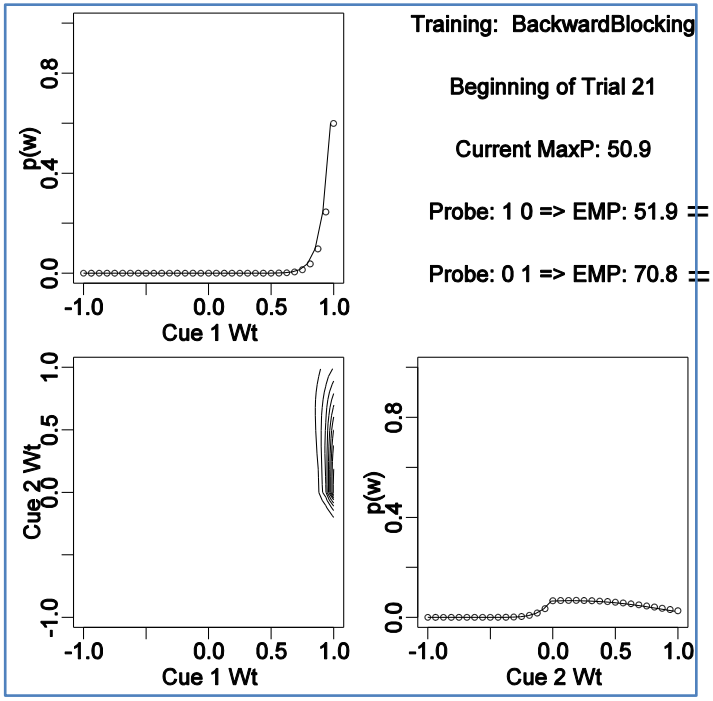


Max P and Minimal Entropy do not always correspond:





For a candidate probe, we compute the *expected* max p:



Probe: 1 0

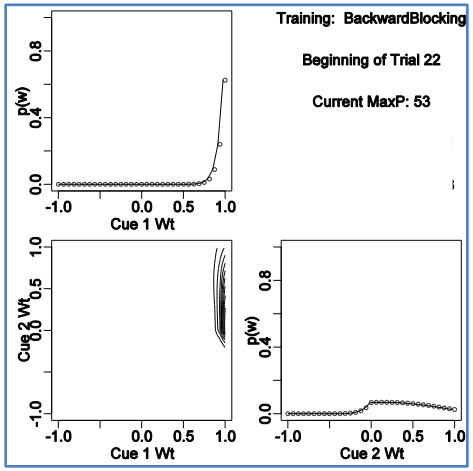
p = .96 → outcome = 1

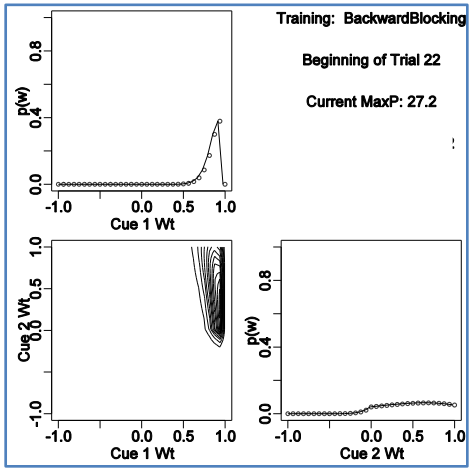
p = .04 → outcome = 0

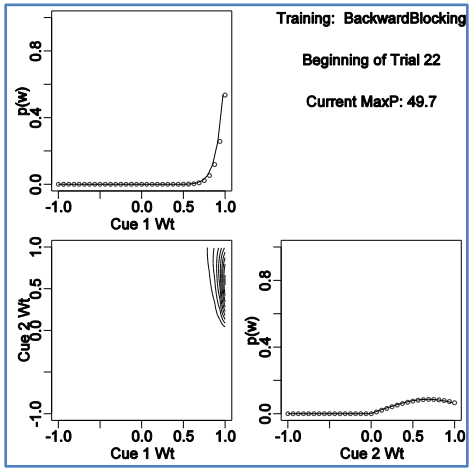
Probe: 0 1

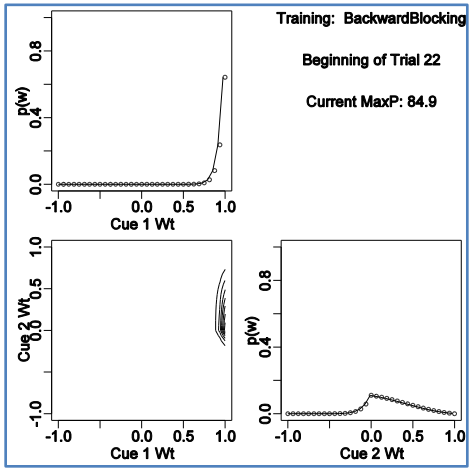
p = .40 → outcome = 1

p = .60 → outcome = 0



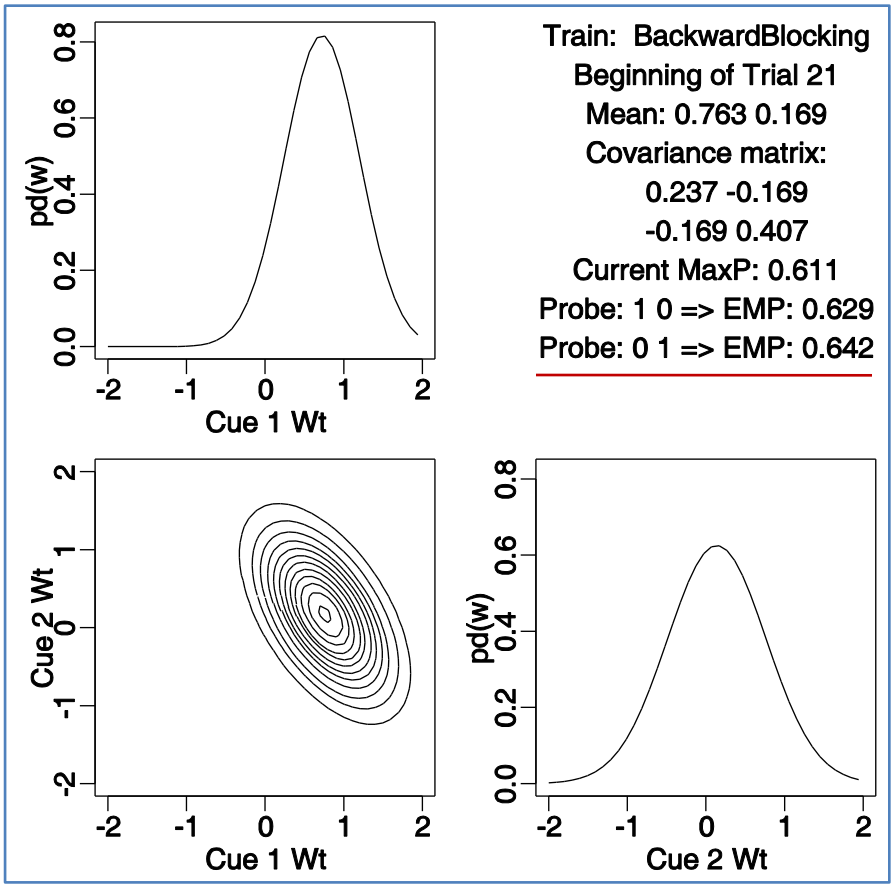


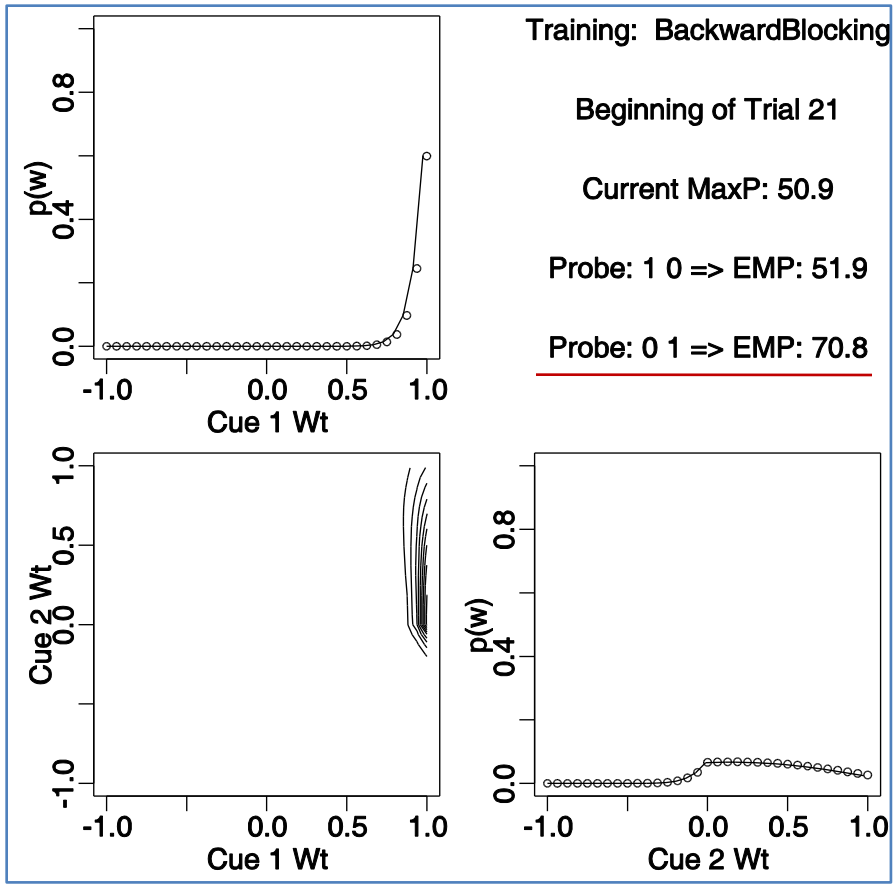




Active Learning after Backward Blocking

Both Kalman filter and noisy-logic gate match human intuition when maximizing Expected Max P:





Design of the (rest of the) Talk:
2 x 2 x 3 factorial

Bayesian models of associative learning:

- Kalman filter.
- Noisy-logic gate.

Goals for active learning:

- Minimize expected uncertainty.
- Maximize expected maximum believability.

Training structures:

- Backward blocking.
- Blocking and reduced overshadowing.
- Ambiguous cue.

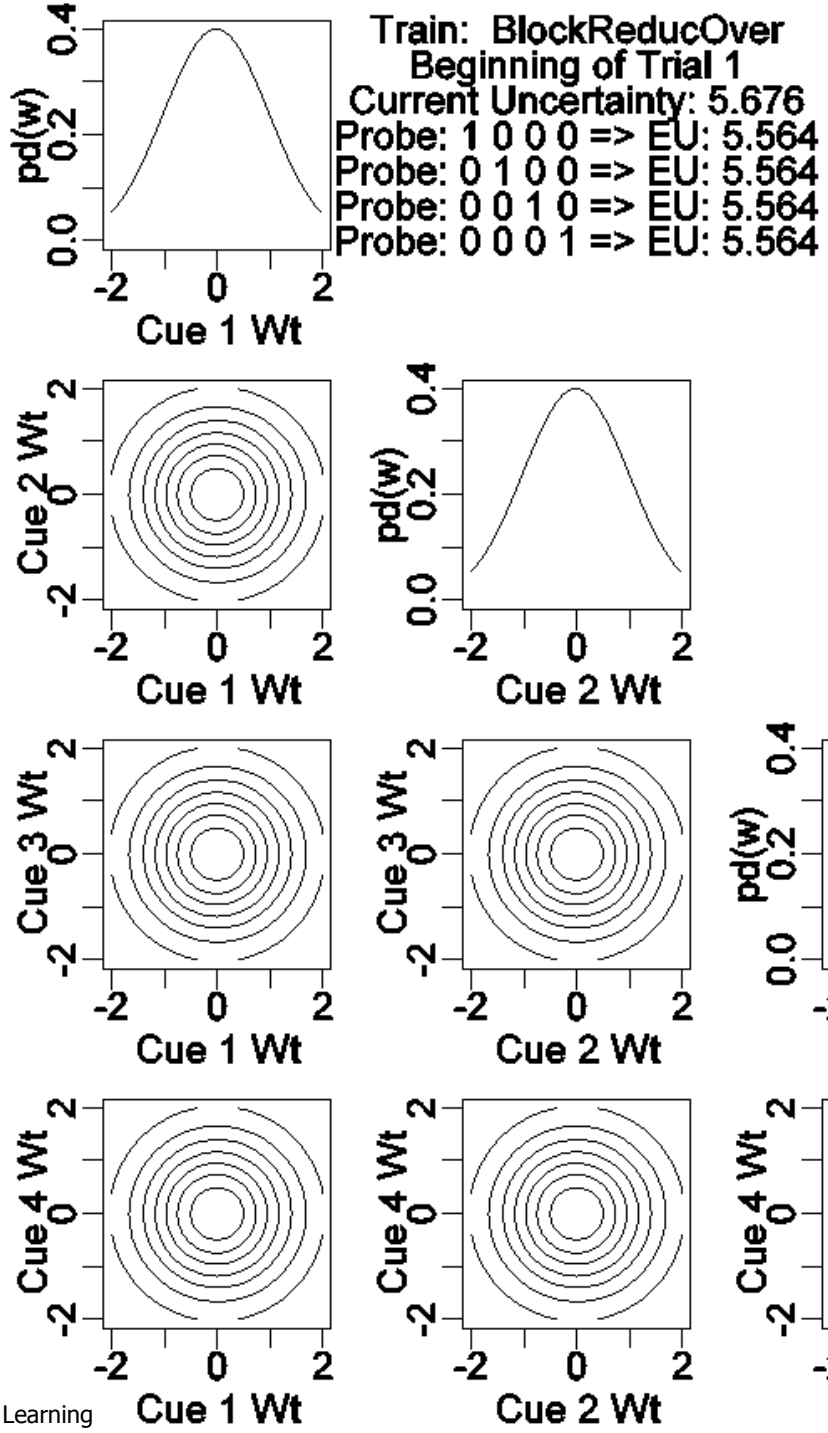
Blocking and Reduced Overshadowing

Phase	Freq.	Cue 1	Cue 2	Cue 3	Cue 4	Outc.
I	6	1	0	0	0	1
		0	0	1	0	0
II	6	1	1	0	0	1
		0	0	1	1	1

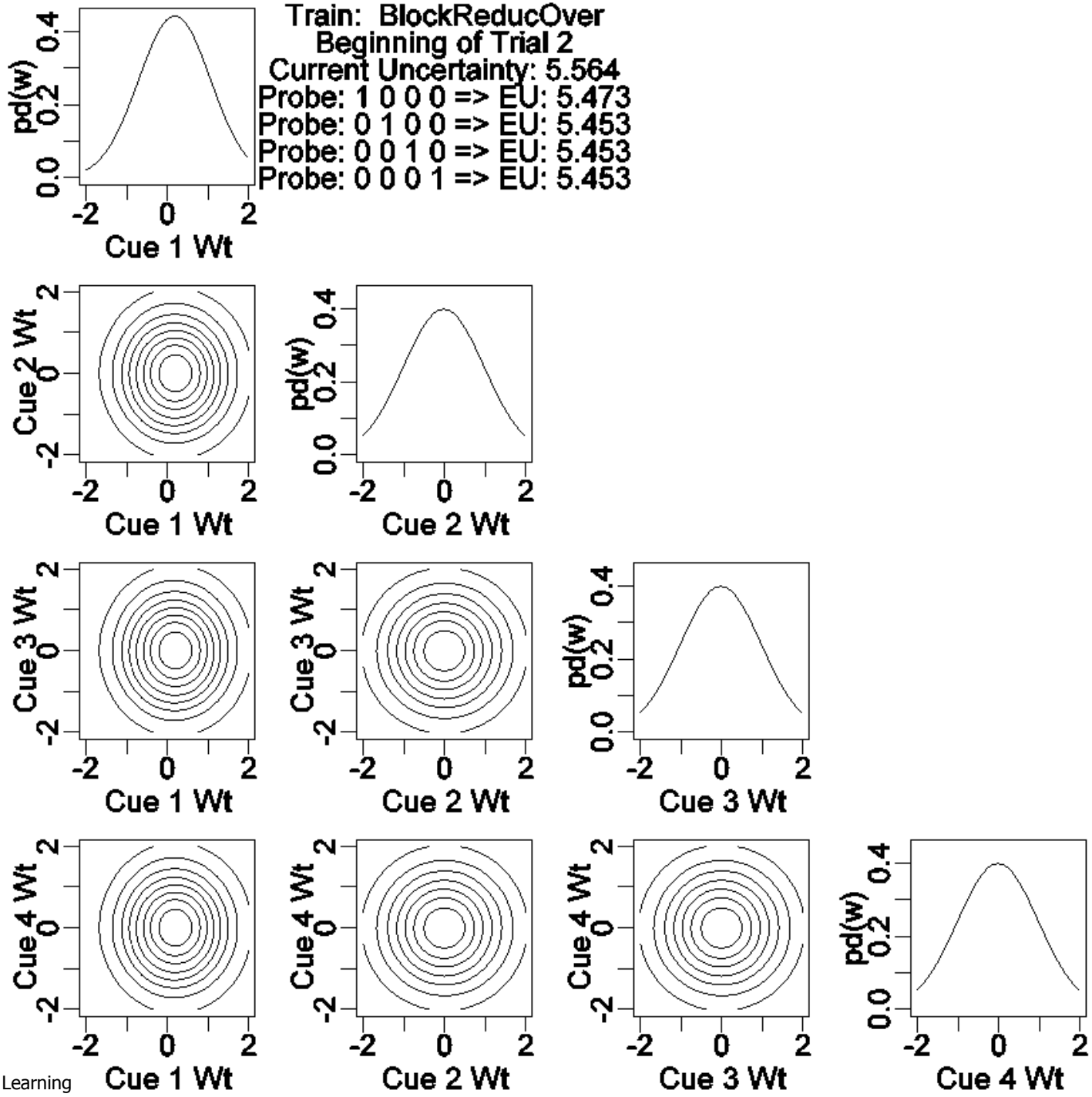
Active Learning
after Blocking and Reduced Overshadowing

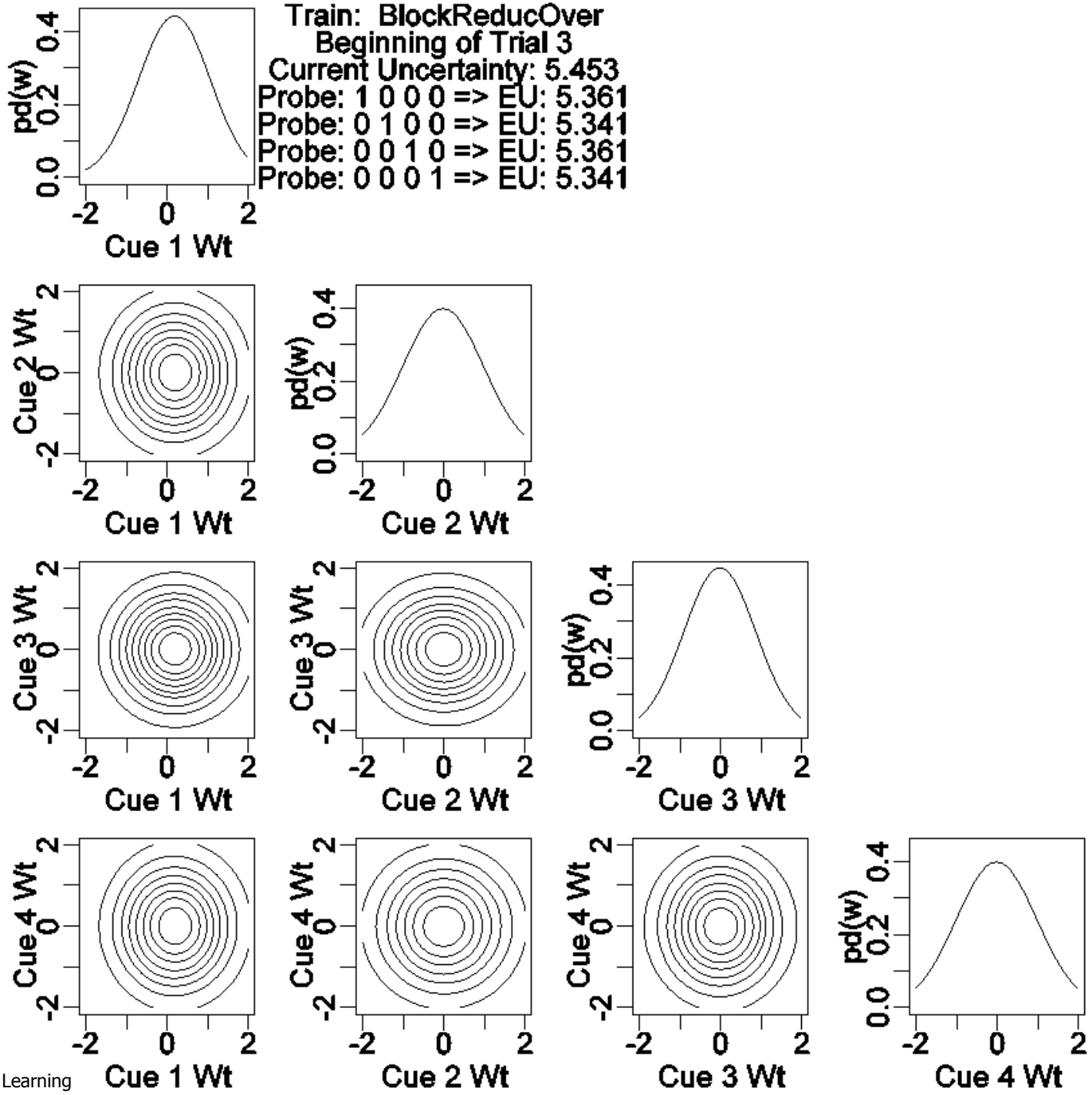
Phase	Freq.	Cue 1	Cue 2	Cue 3	Cue 4	Outc.
I	6	1	0	0	0	1
		0	0	1	0	0
II	6	1	1	0	0	1
		0	0	1	1	1

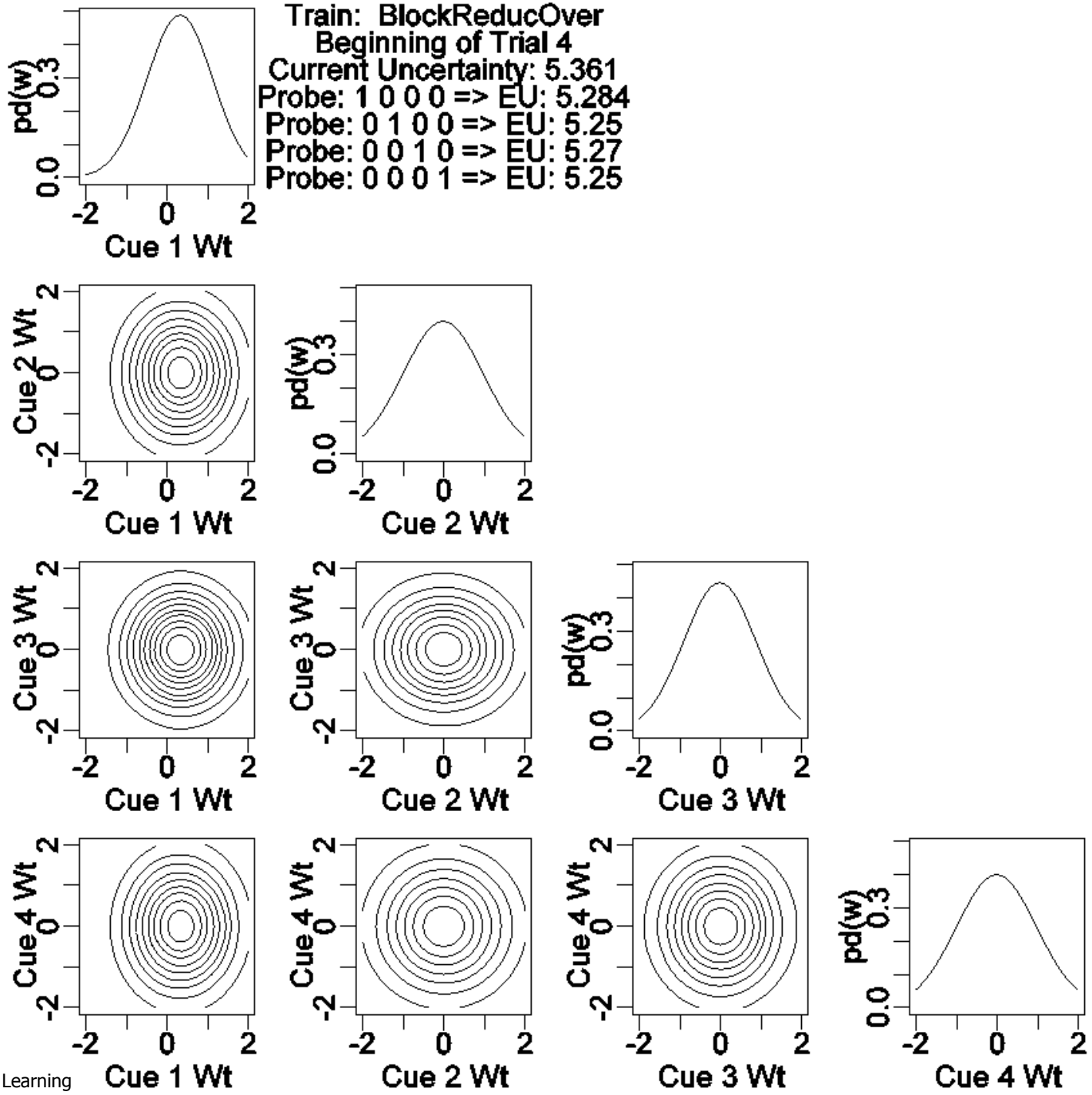
After being trained, your goal is to better determine which cues produce or prevent nausea. Which probe below do you prefer to test?

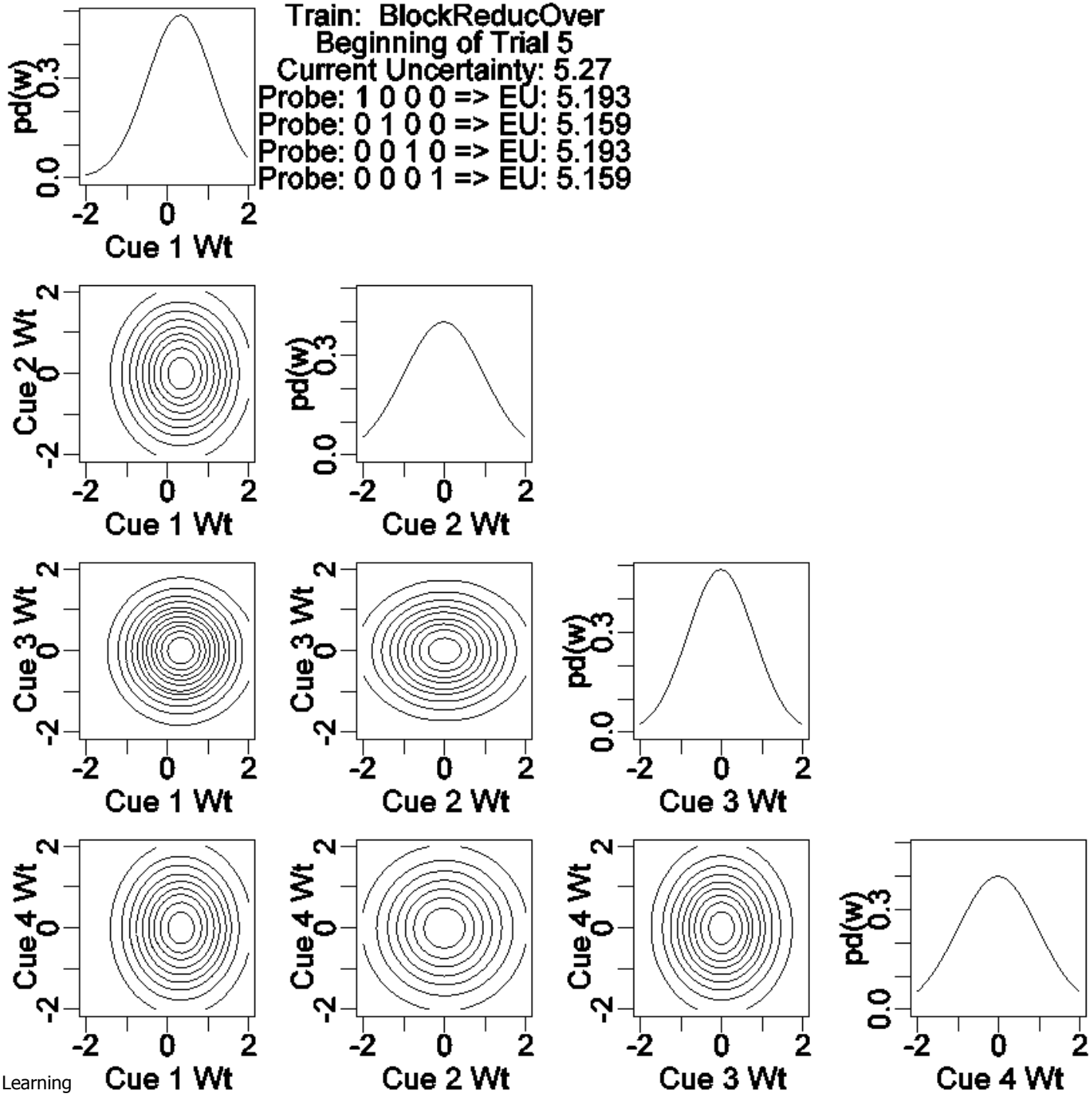


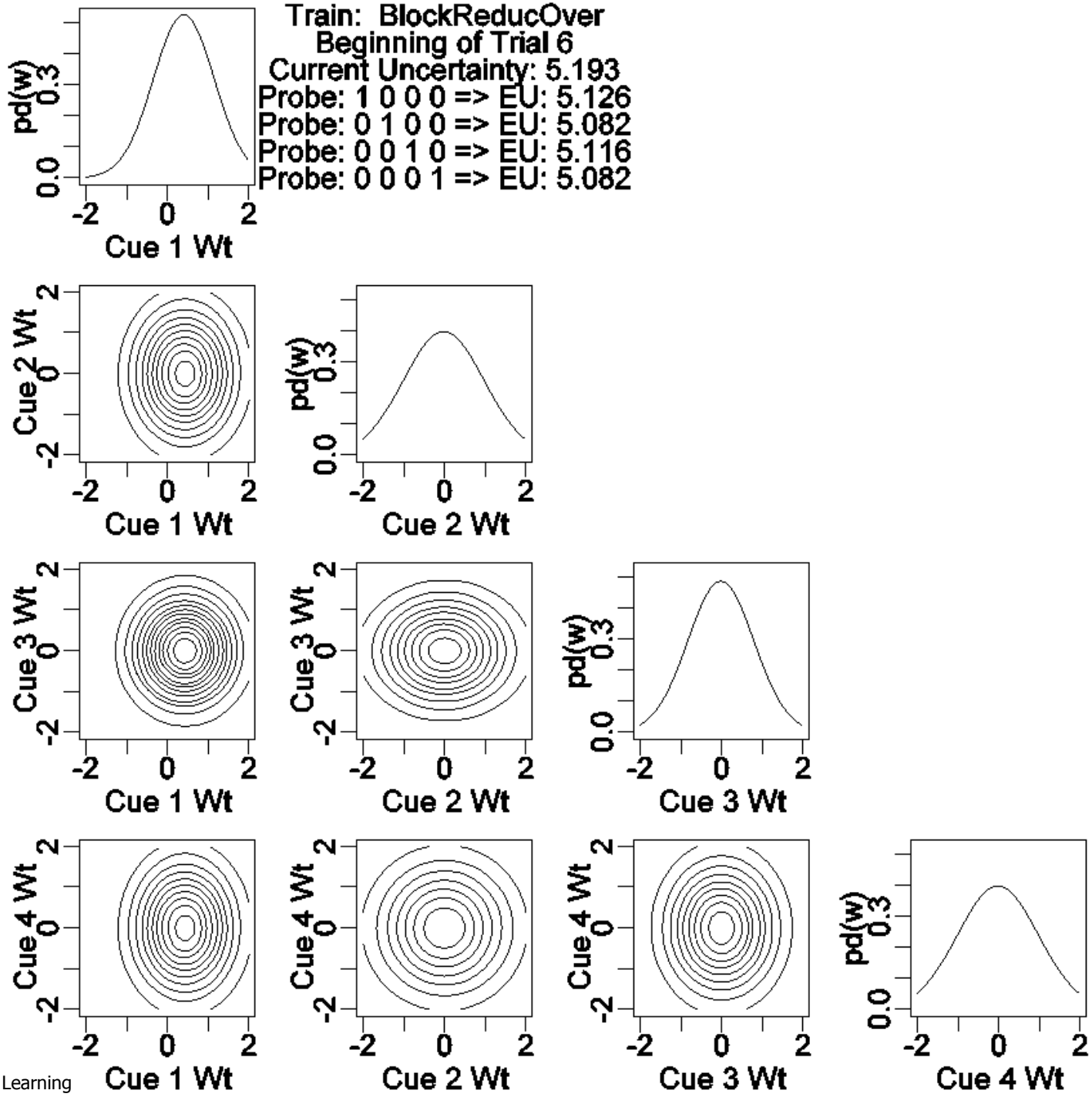
Kalman filter predictions for Expected Uncertainty…

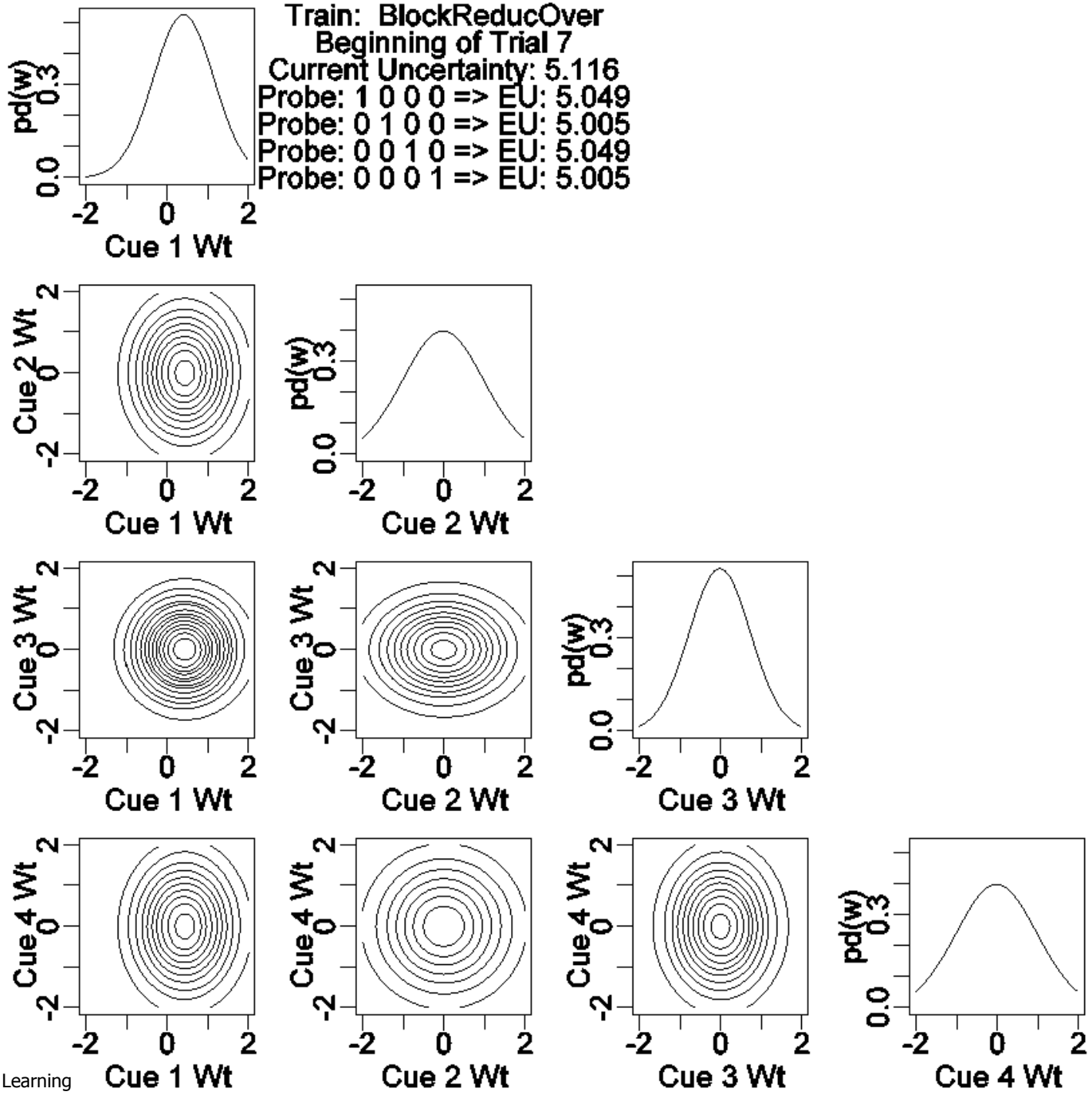

Train: BlockReducOver
Beginning of Trial 2
Current Uncertainty: 5.564
Probe: 1 0 0 0 => EU: 5.473
Probe: 0 1 0 0 => EU: 5.453
Probe: 0 0 1 0 => EU: 5.453
Probe: 0 0 0 1 => EU: 5.453
pd(w)
0.0
0.2
0.4
-2
0
2
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt

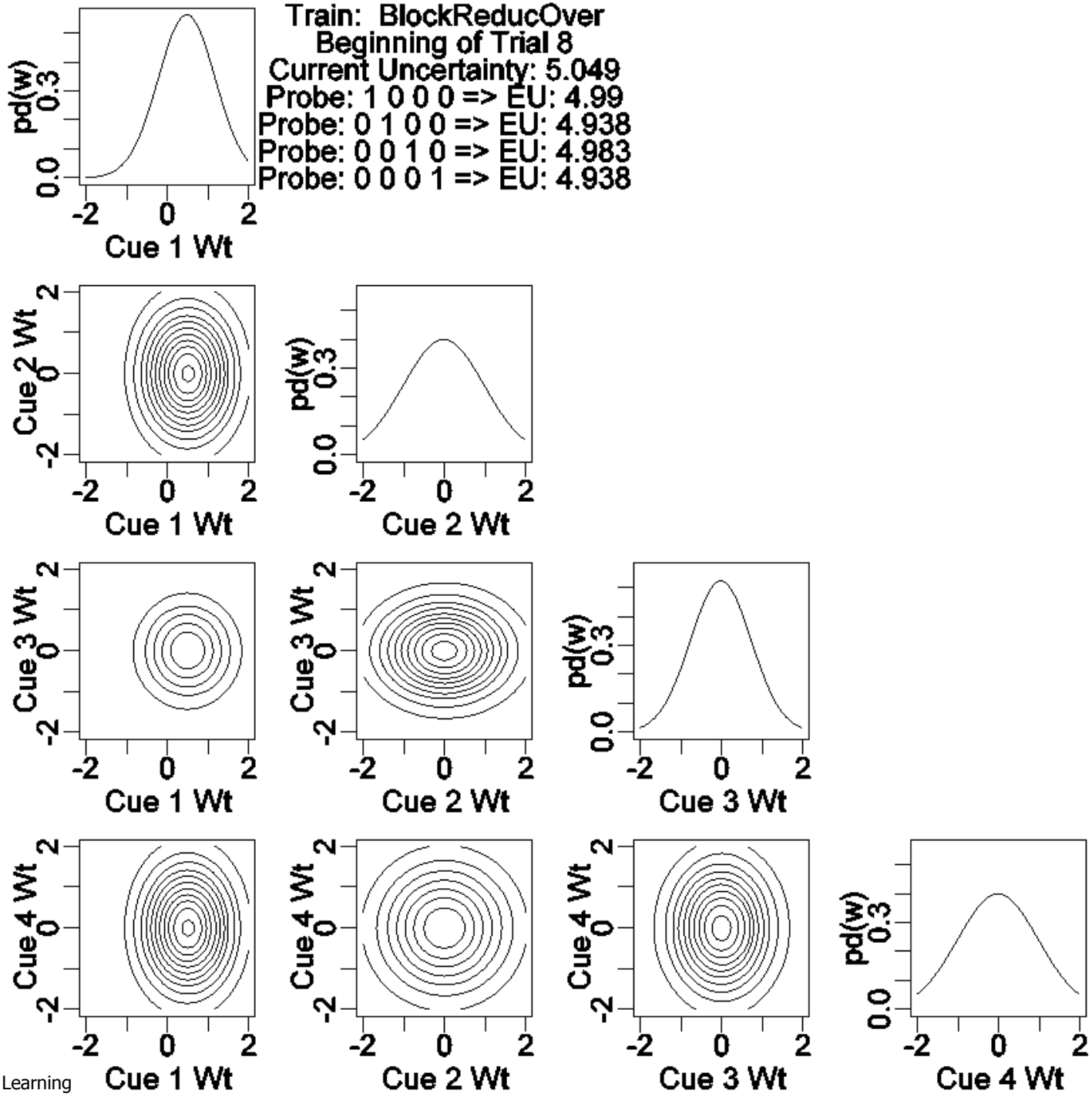

Train: BlockReducOver
Beginning of Trial 3
Current Uncertainty: 5.453
Probe: 1 0 0 0 => EU: 5.361
Probe: 0 1 0 0 => EU: 5.341
Probe: 0 0 1 0 => EU: 5.361
Probe: 0 0 0 1 => EU: 5.341
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt

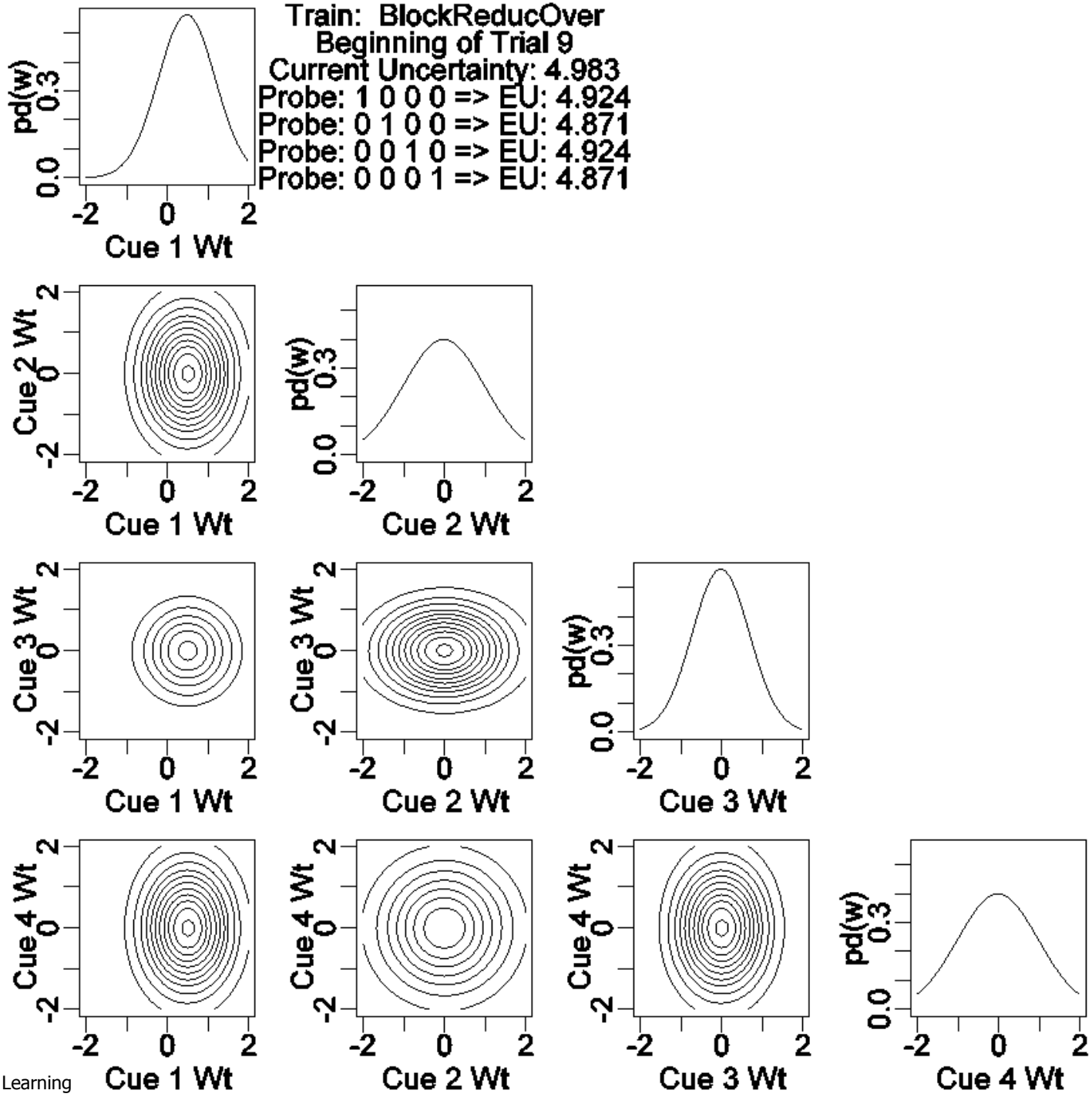

Train: BlockReducOver
Beginning of Trial 4
Current Uncertainty: 5.361
Probe: 1 0 0 0 => EU: 5.284
Probe: 0 1 0 0 => EU: 5.25
Probe: 0 0 1 0 => EU: 5.27
Probe: 0 0 0 1 => EU: 5.25
pd(w)
0.0
0.3
-2
0
2
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt

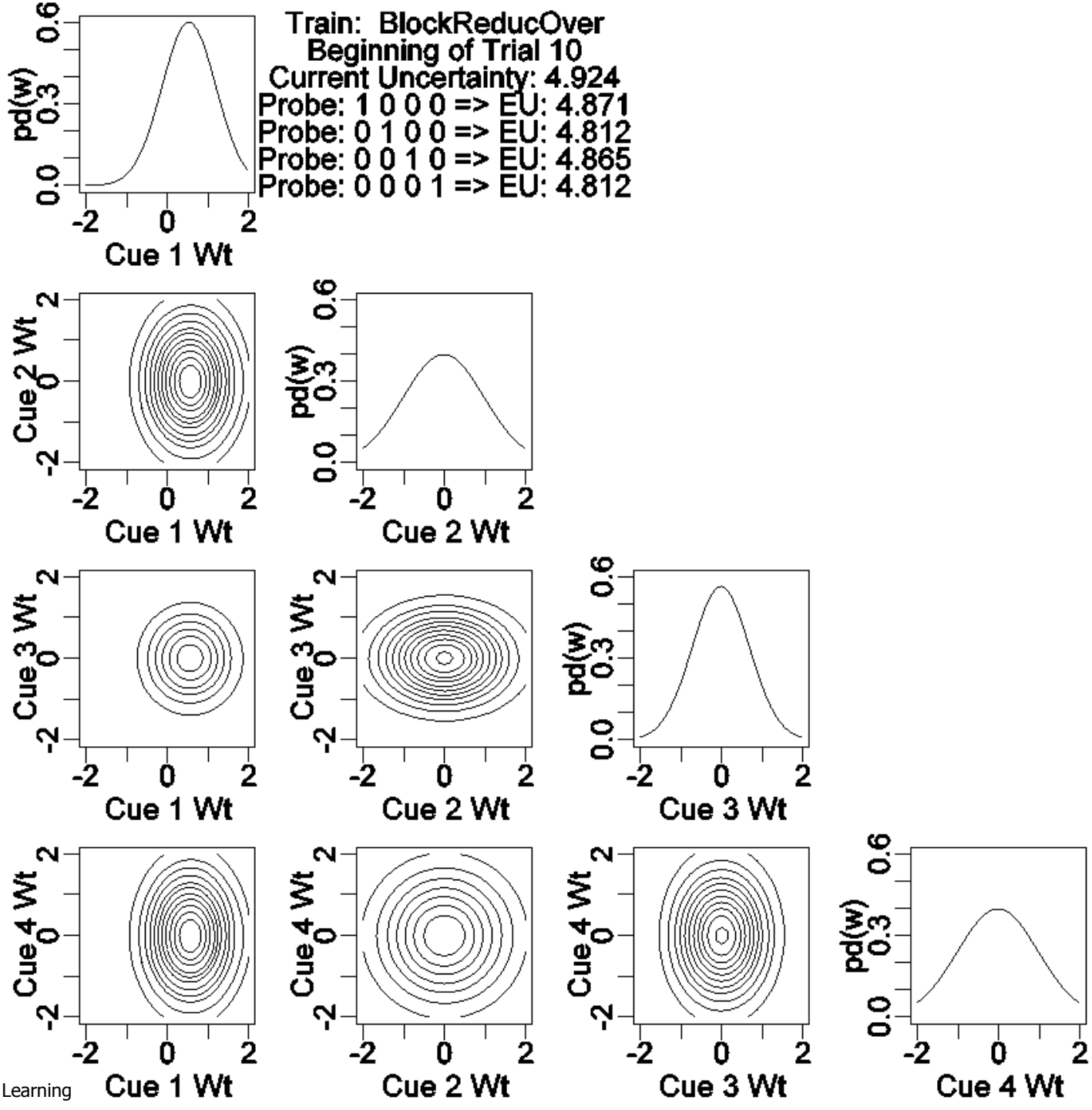

Train: BlockReducOver
Beginning of Trial 5
Current Uncertainty: 5.27
Probe: 1 0 0 0 => EU: 5.193
Probe: 0 1 0 0 => EU: 5.159
Probe: 0 0 1 0 => EU: 5.193
Probe: 0 0 0 1 => EU: 5.159
pd(w)
0.0
0.3
-2
0
2
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt

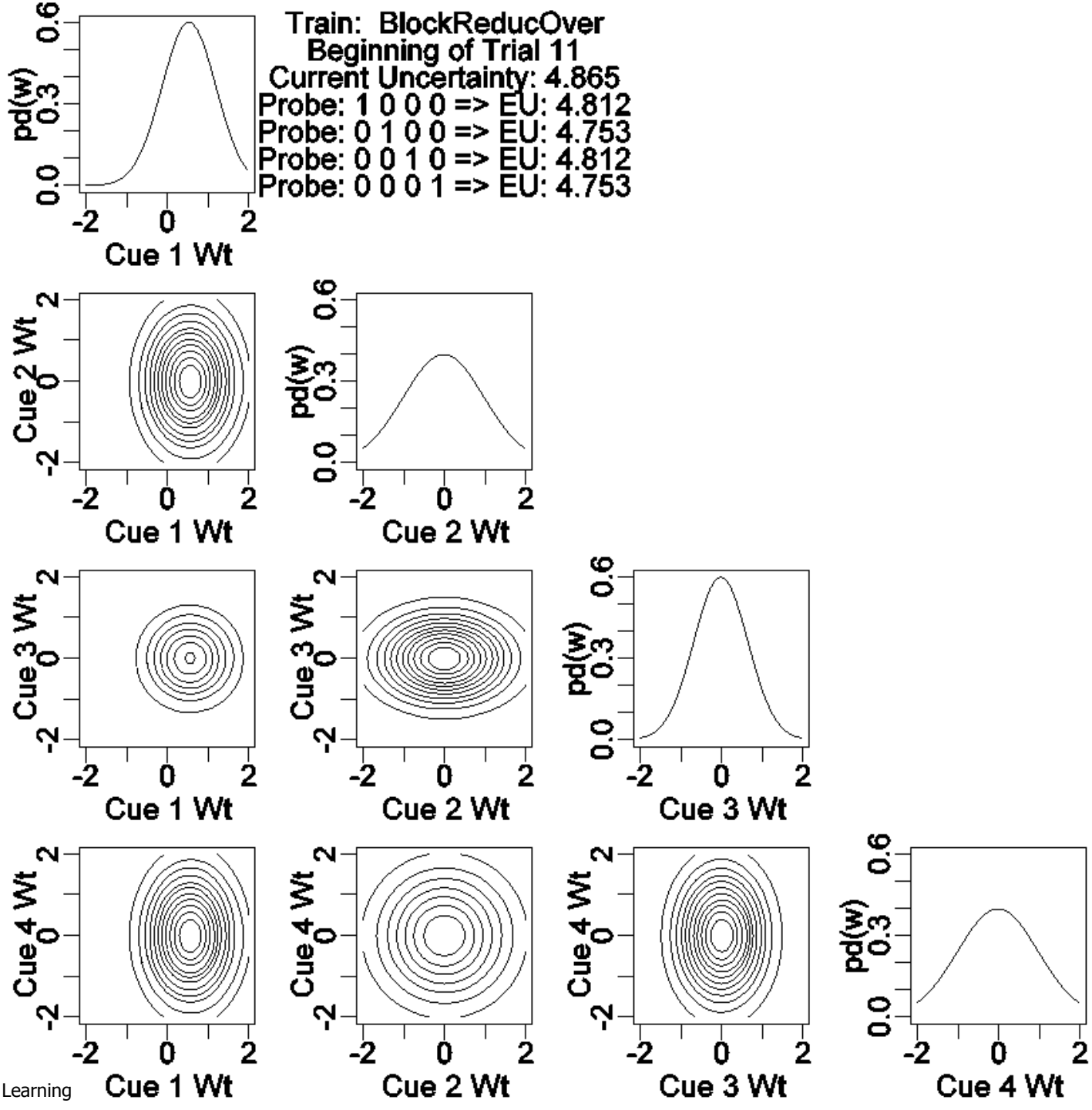

Train: BlockReducOver
Beginning of Trial 6
Current Uncertainty: 5.193
Probe: 1 0 0 0 => EU: 5.126
Probe: 0 1 0 0 => EU: 5.082
Probe: 0 0 1 0 => EU: 5.116
Probe: 0 0 0 1 => EU: 5.082
pd(w)
0.0
0.3
-2
0
2
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt

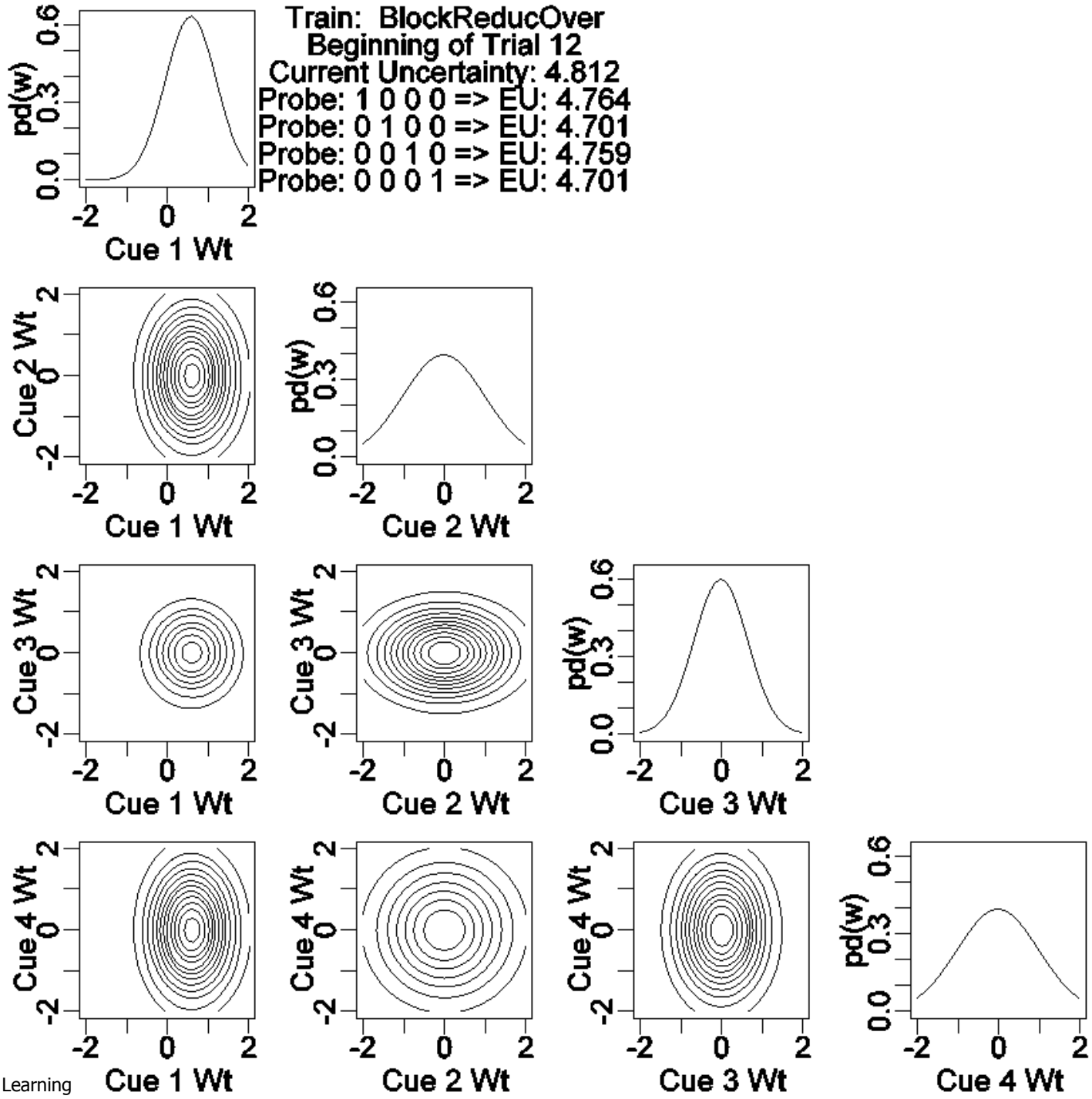

Train: BlockReducOver
Beginning of Trial 7
Current Uncertainty: 5.116
Probe: 1 0 0 0 => EU: 5.049
Probe: 0 1 0 0 => EU: 5.005
Probe: 0 0 1 0 => EU: 5.049
Probe: 0 0 0 1 => EU: 5.005
pd(w)
0.0
0.3
-2
0
2
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt

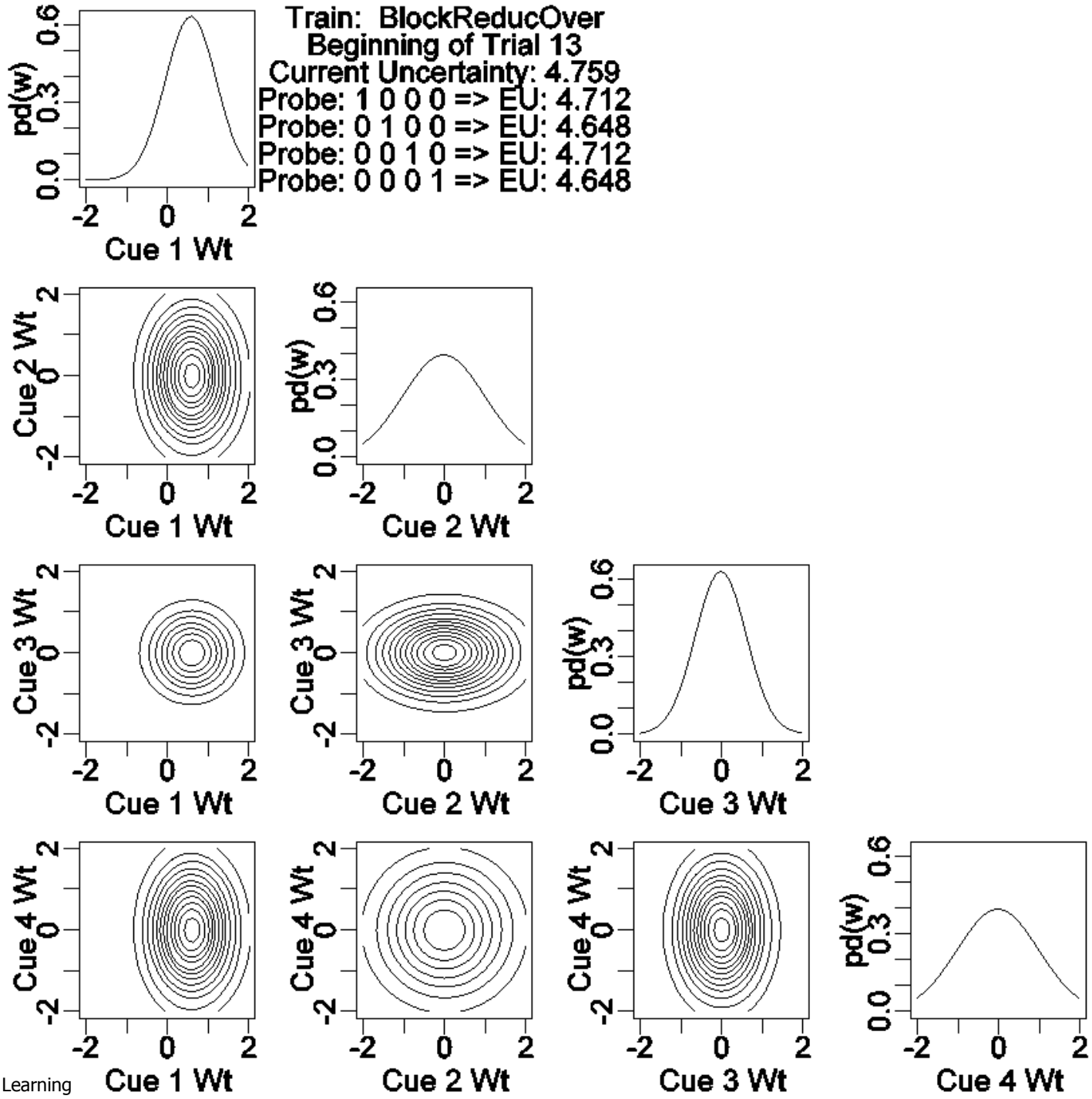

Train: BlockReducOver
Beginning of Trial 8
Current Uncertainty: 5.049
Probe: 1 0 0 0 => EU: 4.99
Probe: 0 1 0 0 => EU: 4.938
Probe: 0 0 1 0 => EU: 4.983
Probe: 0 0 0 1 => EU: 4.938
pd(w)
0.0
0.3
-2
0
2
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt

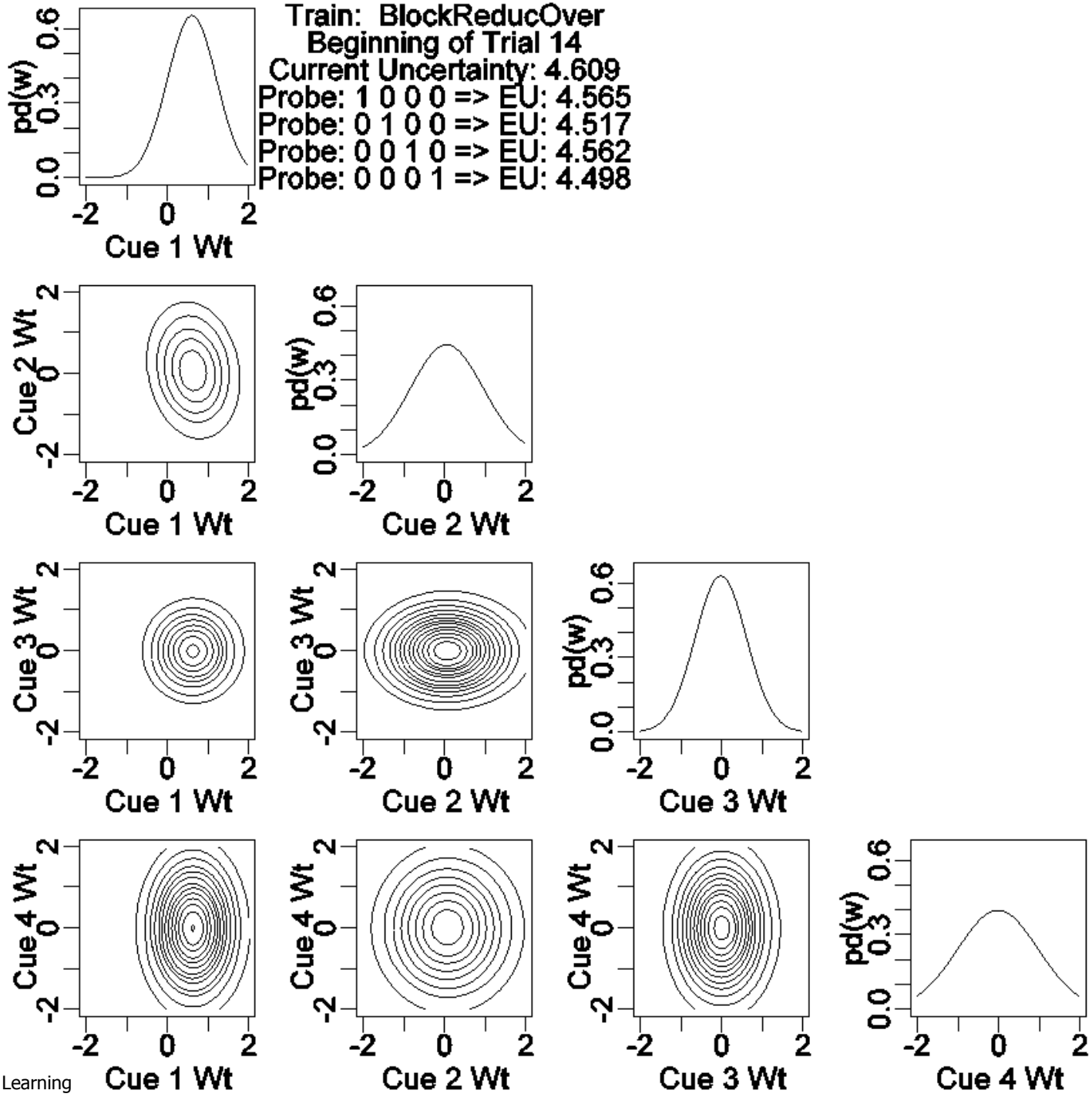

Train: BlockReducOver
Beginning of Trial 9
Current Uncertainty: 4.983
Probe: 1 0 0 0 => EU: 4.924
Probe: 0 1 0 0 => EU: 4.871
Probe: 0 0 1 0 => EU: 4.924
Probe: 0 0 0 1 => EU: 4.871
pd(w)
0.0
0.3
-2
0
2
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt

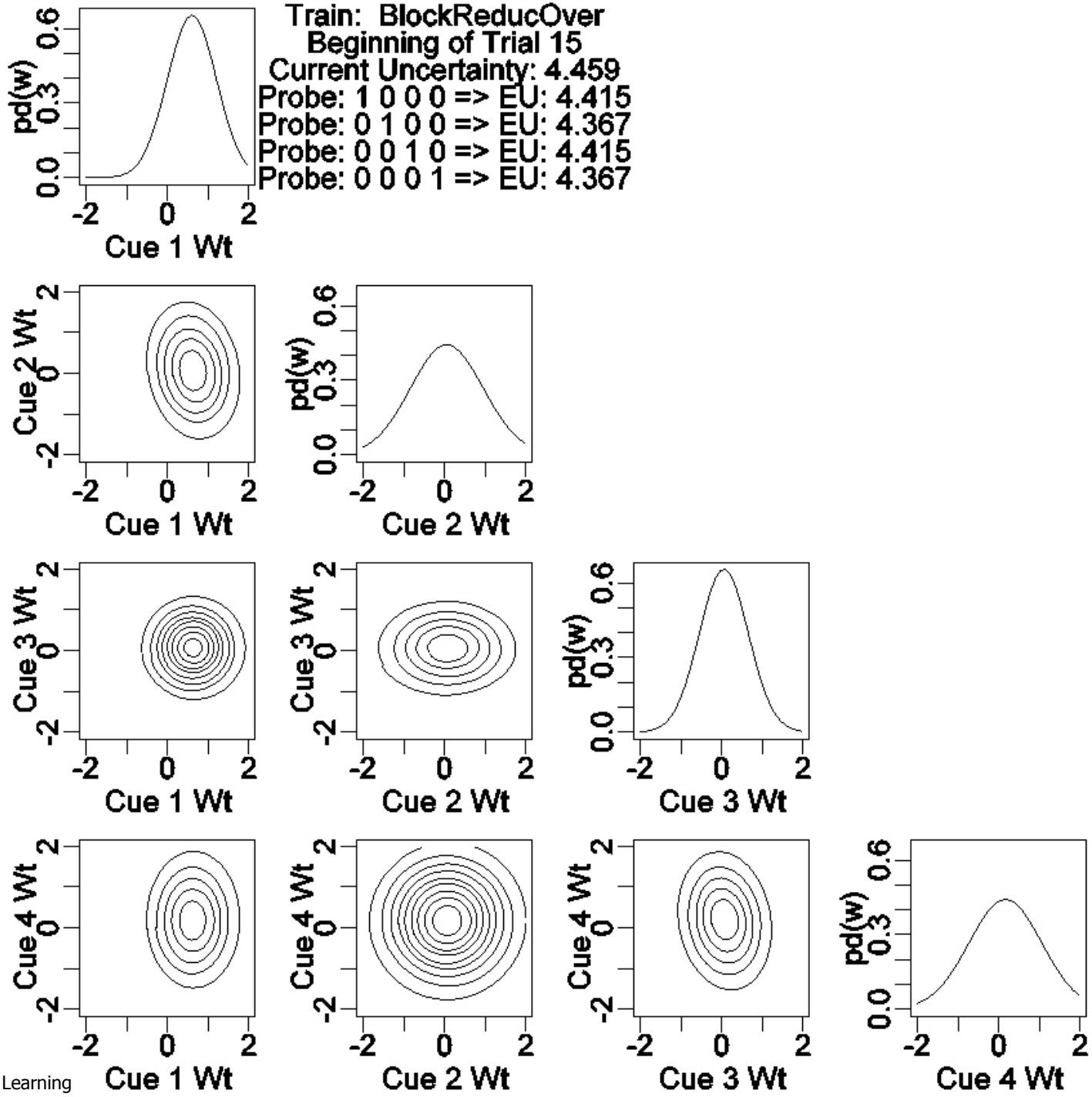

Train: BlockReducOver
Beginning of Trial 10
Current Uncertainty: 4.924
Probe: 1 0 0 0 => EU: 4.871
Probe: 0 1 0 0 => EU: 4.812
Probe: 0 0 1 0 => EU: 4.865
Probe: 0 0 0 1 => EU: 4.812
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt

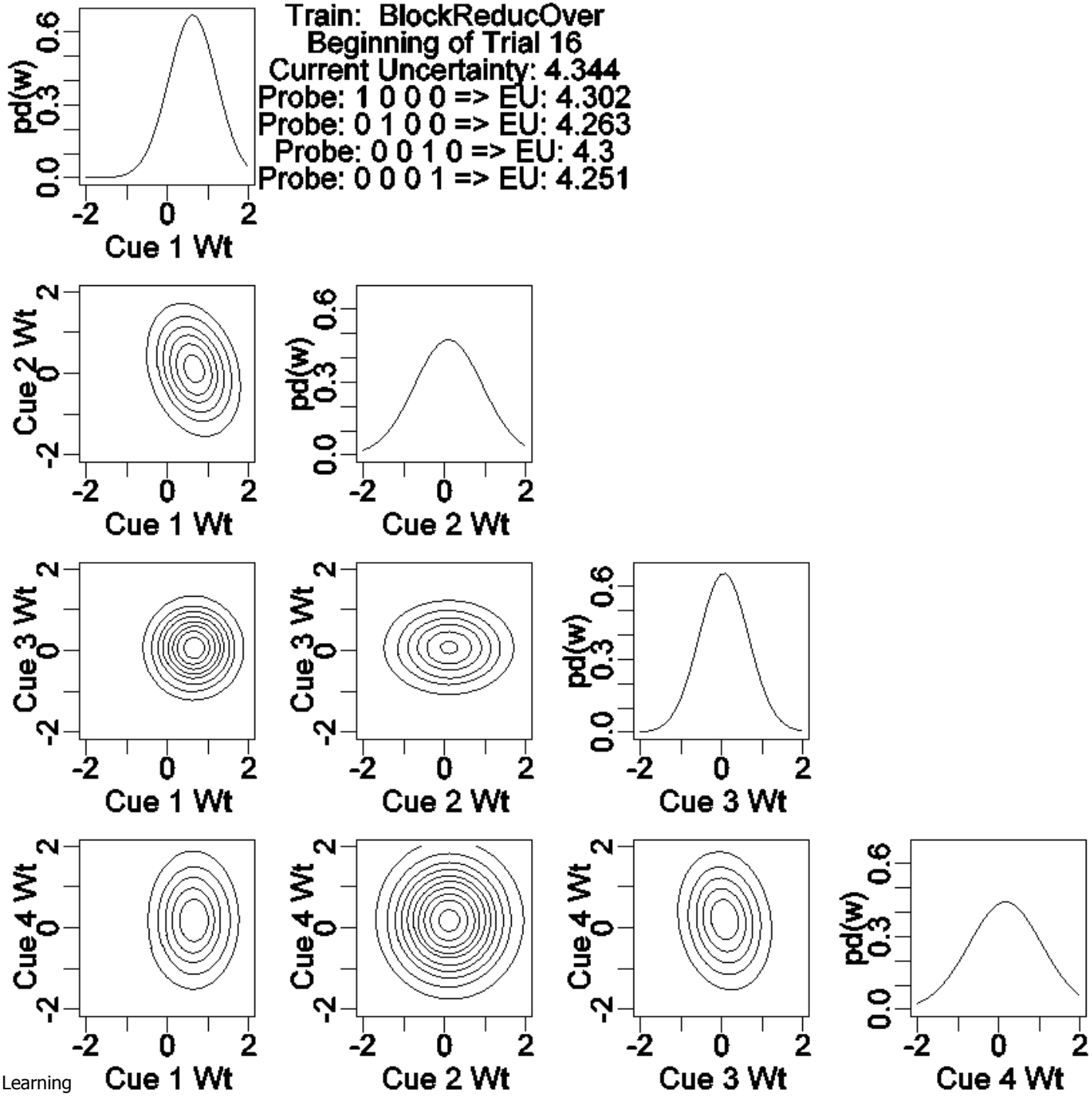

Train: BlockReducOver
Beginning of Trial 11
Current Uncertainty: 4.865
Probe: 1 0 0 0 => EU: 4.812
Probe: 0 1 0 0 => EU: 4.753
Probe: 0 0 1 0 => EU: 4.812
Probe: 0 0 0 1 => EU: 4.753
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt

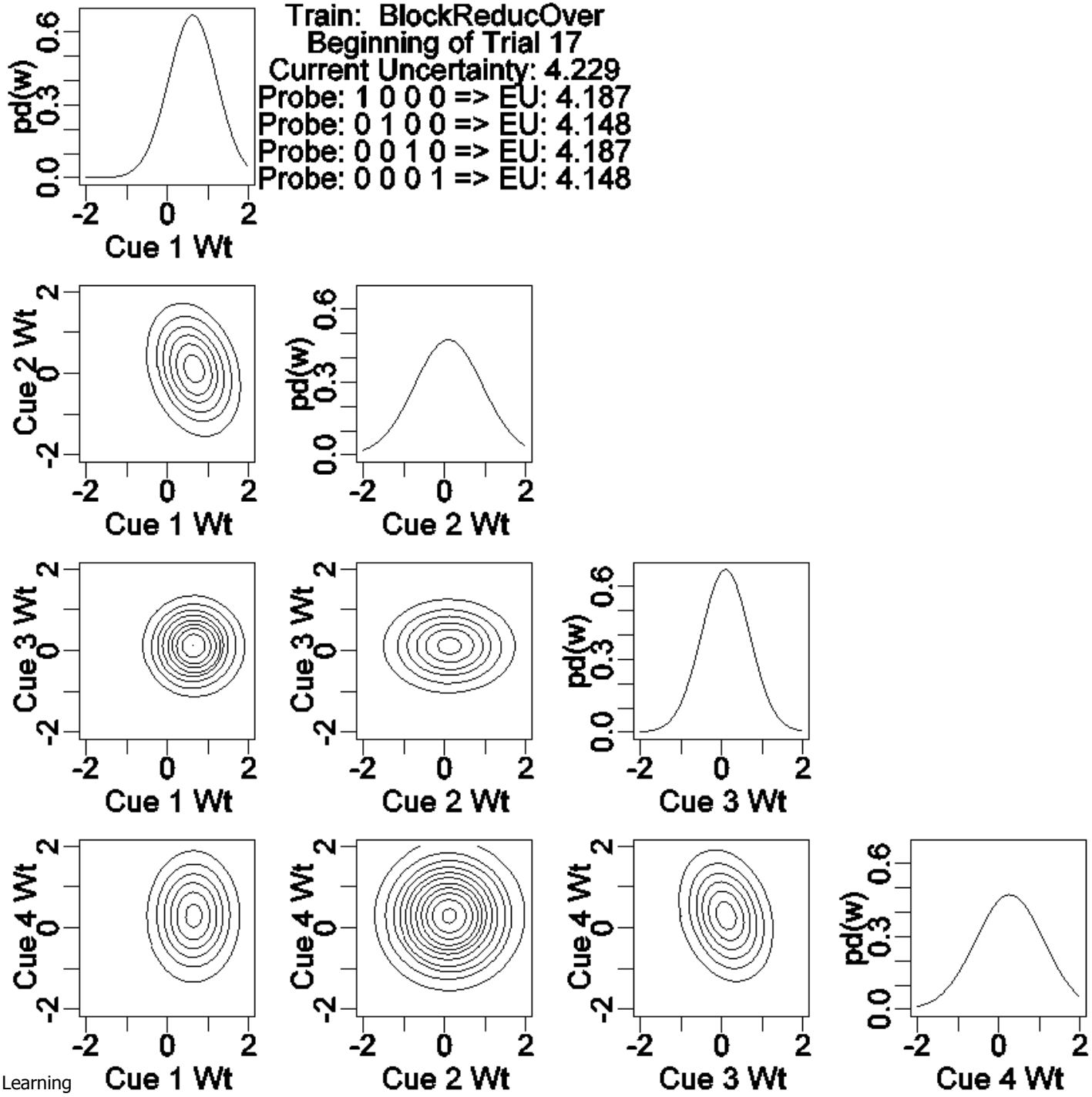

Train: BlockReducOver
Beginning of Trial 12
Current Uncertainty: 4.812
Probe: 1 0 0 0 => EU: 4.764
Probe: 0 1 0 0 => EU: 4.701
Probe: 0 0 1 0 => EU: 4.759
Probe: 0 0 0 1 => EU: 4.701
pd(w)
0.0
0.3
0.6
-2
0
2
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt

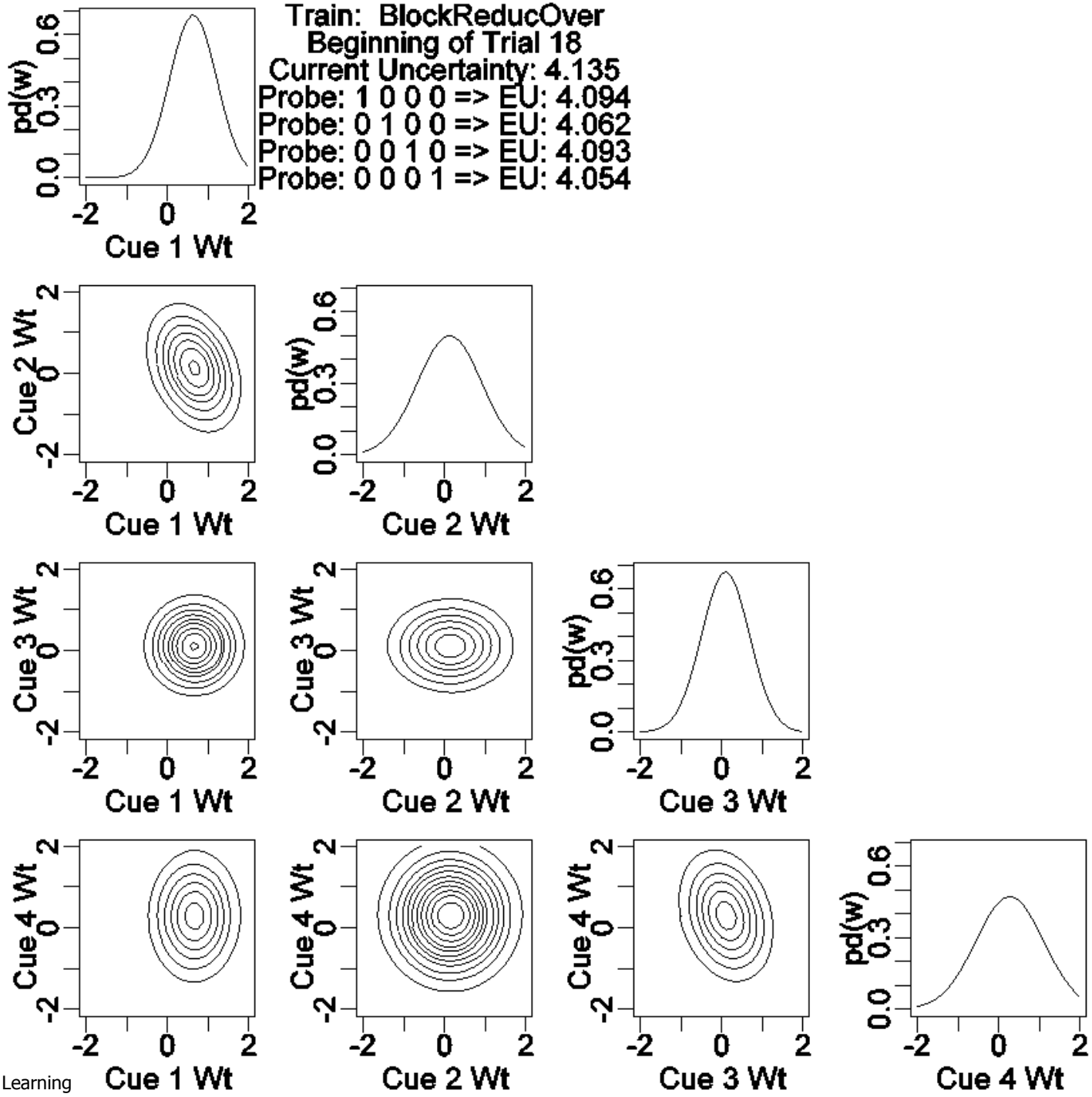

Train: BlockReducOver
Beginning of Trial 13
Current Uncertainty: 4.759
Probe: 1 0 0 0 => EU: 4.712
Probe: 0 1 0 0 => EU: 4.648
Probe: 0 0 1 0 => EU: 4.712
Probe: 0 0 0 1 => EU: 4.648
pd(w)
0.0
0.3
0.6
-2
0
2
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt

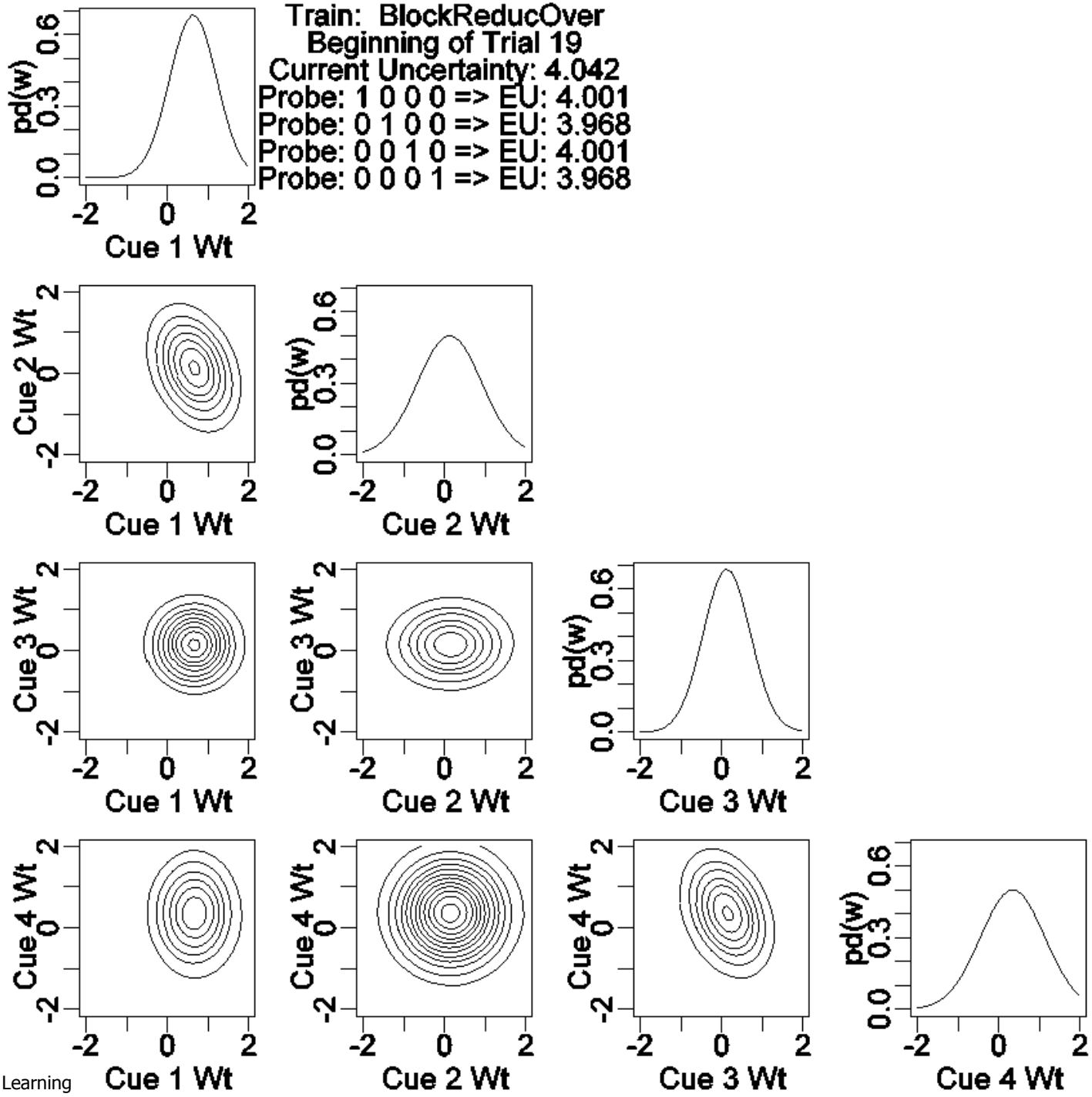

Train: BlockReducOver
Beginning of Trial 14
Current Uncertainty: 4.609
Probe: 1 0 0 0 => EU: 4.565
Probe: 0 1 0 0 => EU: 4.517
Probe: 0 0 1 0 => EU: 4.562
Probe: 0 0 0 1 => EU: 4.498
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt

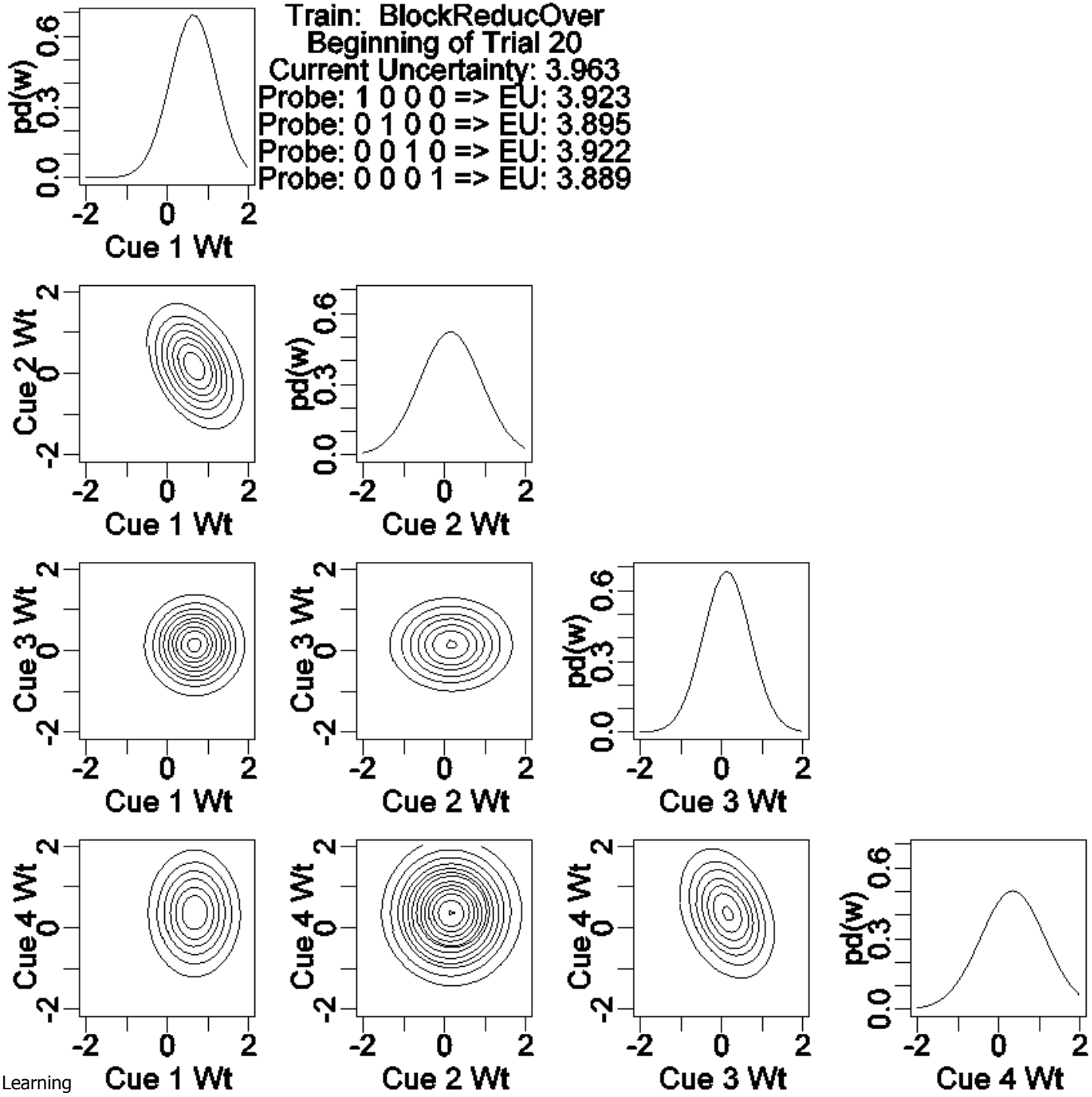

Train: BlockReducOver
Beginning of Trial 15
Current Uncertainty: 4.459
Probe: 1 0 0 0 => EU: 4.415
Probe: 0 1 0 0 => EU: 4.367
Probe: 0 0 1 0 => EU: 4.415
Probe: 0 0 0 1 => EU: 4.367
pd(w)
0.0
0.3
0.6
-2
0
2
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt

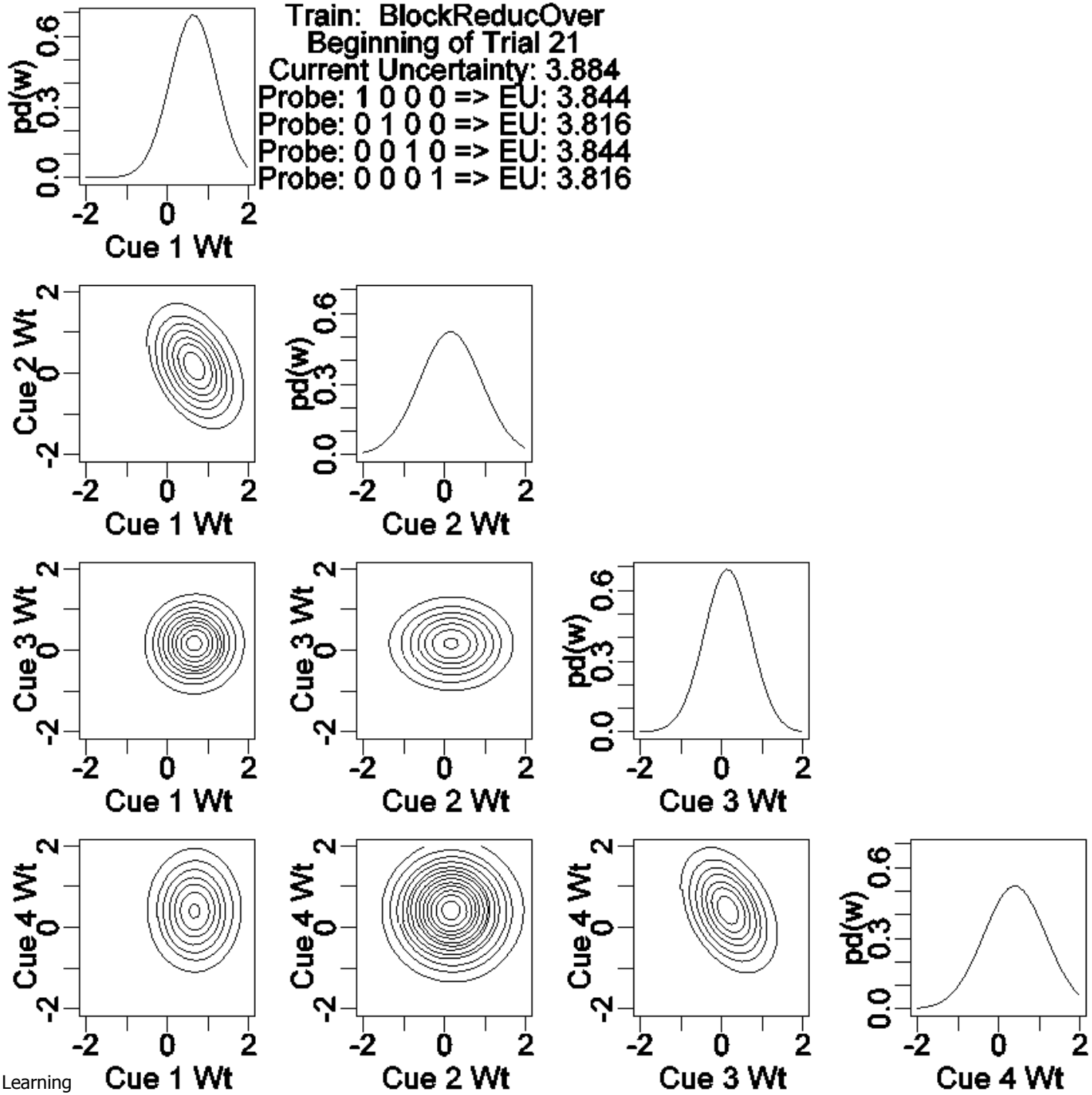

Train: BlockReducOver
Beginning of Trial 16
Current Uncertainty: 4.344
Probe: 1 0 0 0 => EU: 4.302
Probe: 0 1 0 0 => EU: 4.263
Probe: 0 0 1 0 => EU: 4.3
Probe: 0 0 0 1 => EU: 4.251
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt

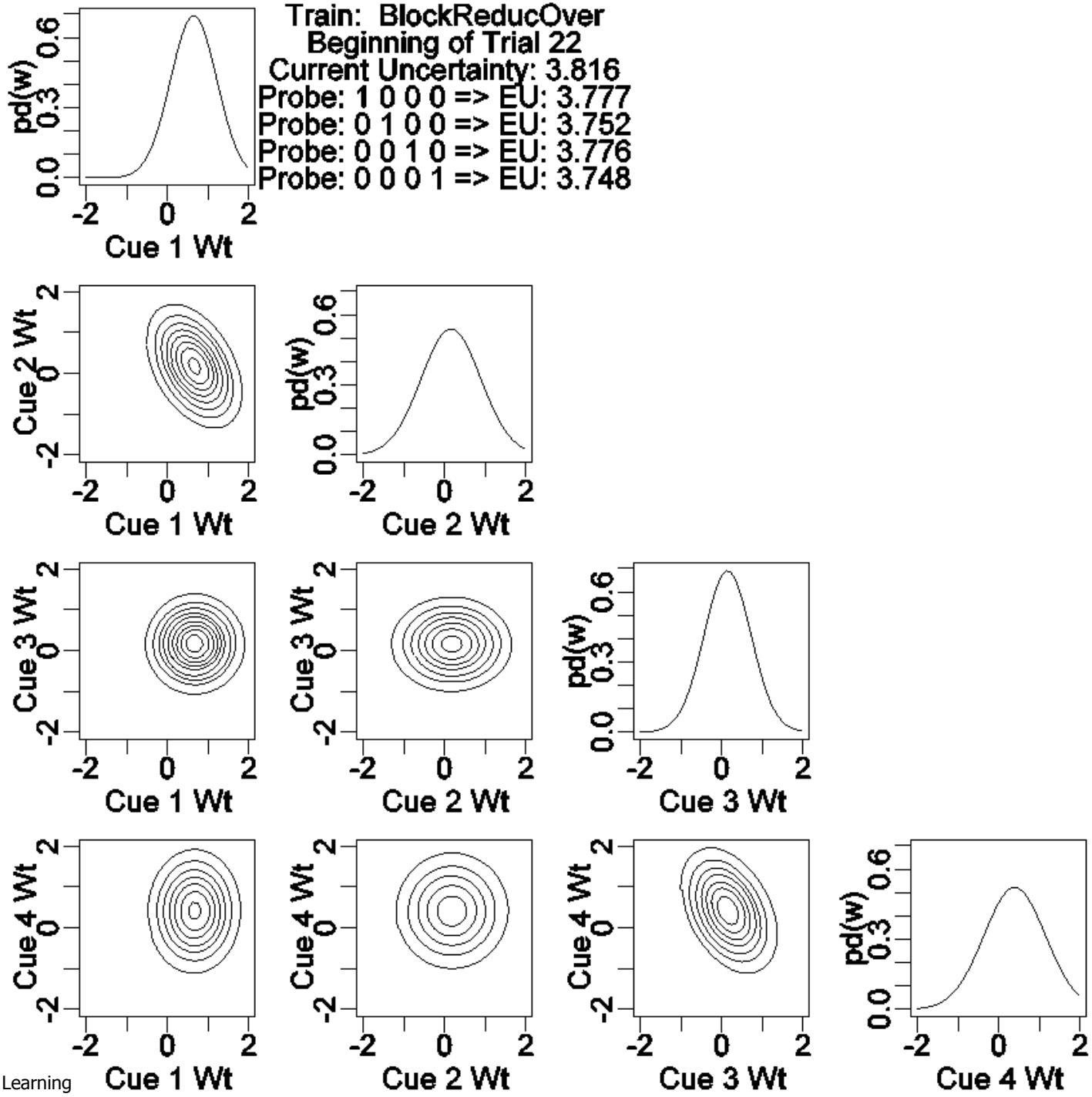

Train: BlockReducOver
Beginning of Trial 17
Current Uncertainty: 4.229
Probe: 1 0 0 0 => EU: 4.187
Probe: 0 1 0 0 => EU: 4.148
Probe: 0 0 1 0 => EU: 4.187
Probe: 0 0 0 1 => EU: 4.148
pd(w)
0.0
0.3
0.6
-2
0
2
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt

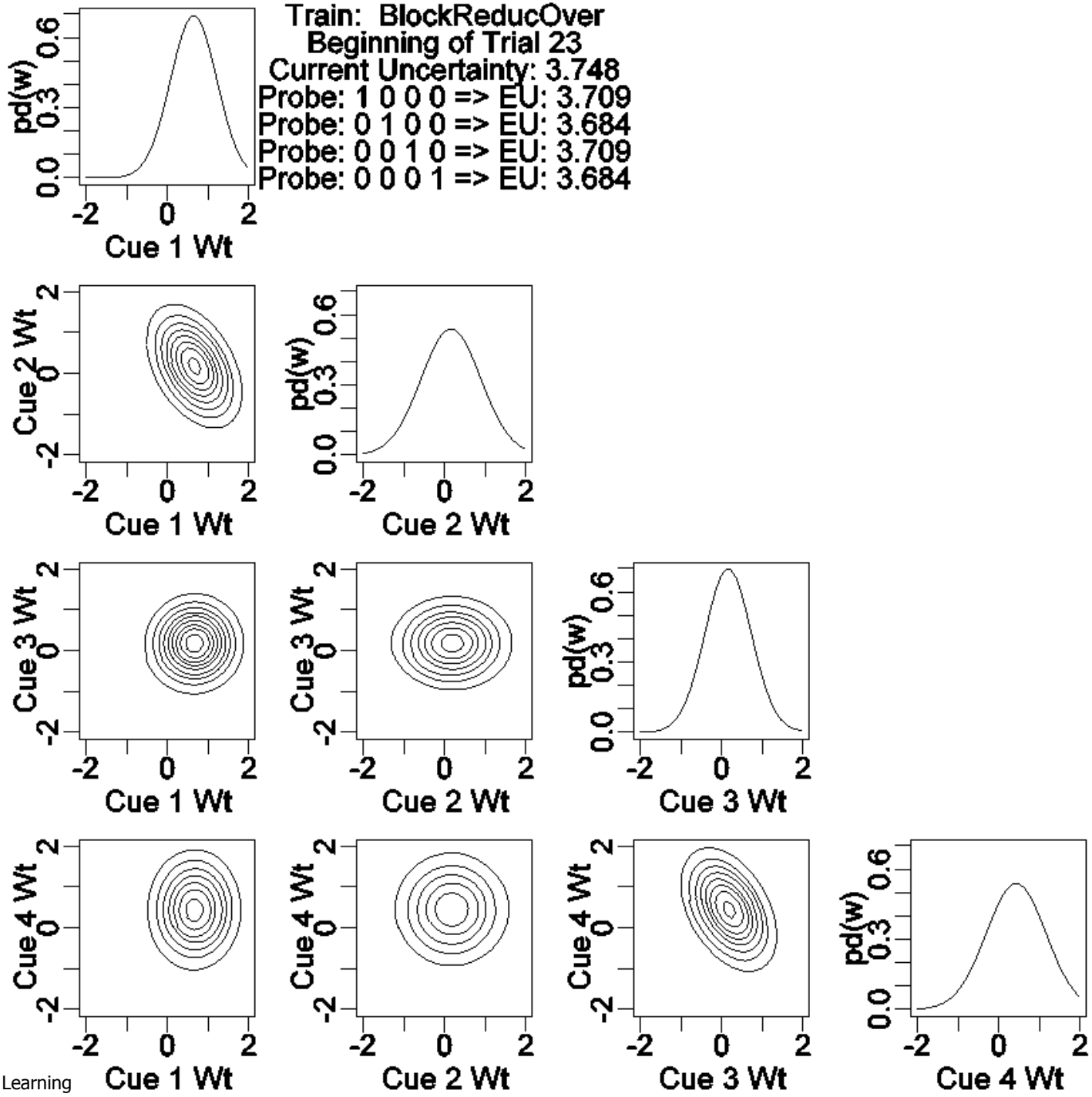

Train: BlockReducOver
Beginning of Trial 18
Current Uncertainty: 4.135
Probe: 1 0 0 0 => EU: 4.094
Probe: 0 1 0 0 => EU: 4.062
Probe: 0 0 1 0 => EU: 4.093
Probe: 0 0 0 1 => EU: 4.054
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt

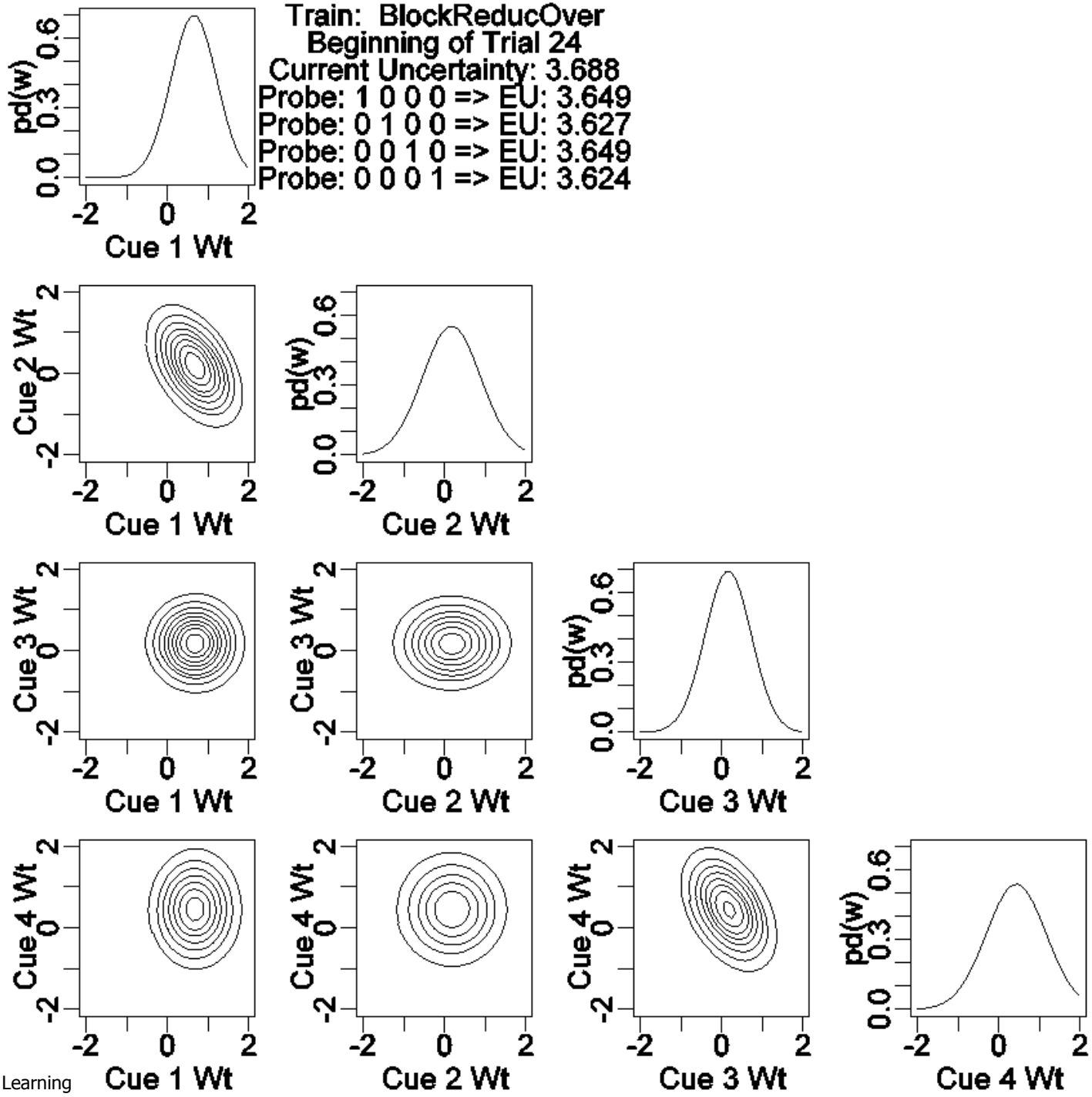

Train: BlockReducOver
Beginning of Trial 19
Current Uncertainty: 4.042
Probe: 1 0 0 0 => EU: 4.001
Probe: 0 1 0 0 => EU: 3.968
Probe: 0 0 1 0 => EU: 4.001
Probe: 0 0 0 1 => EU: 3.968
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt

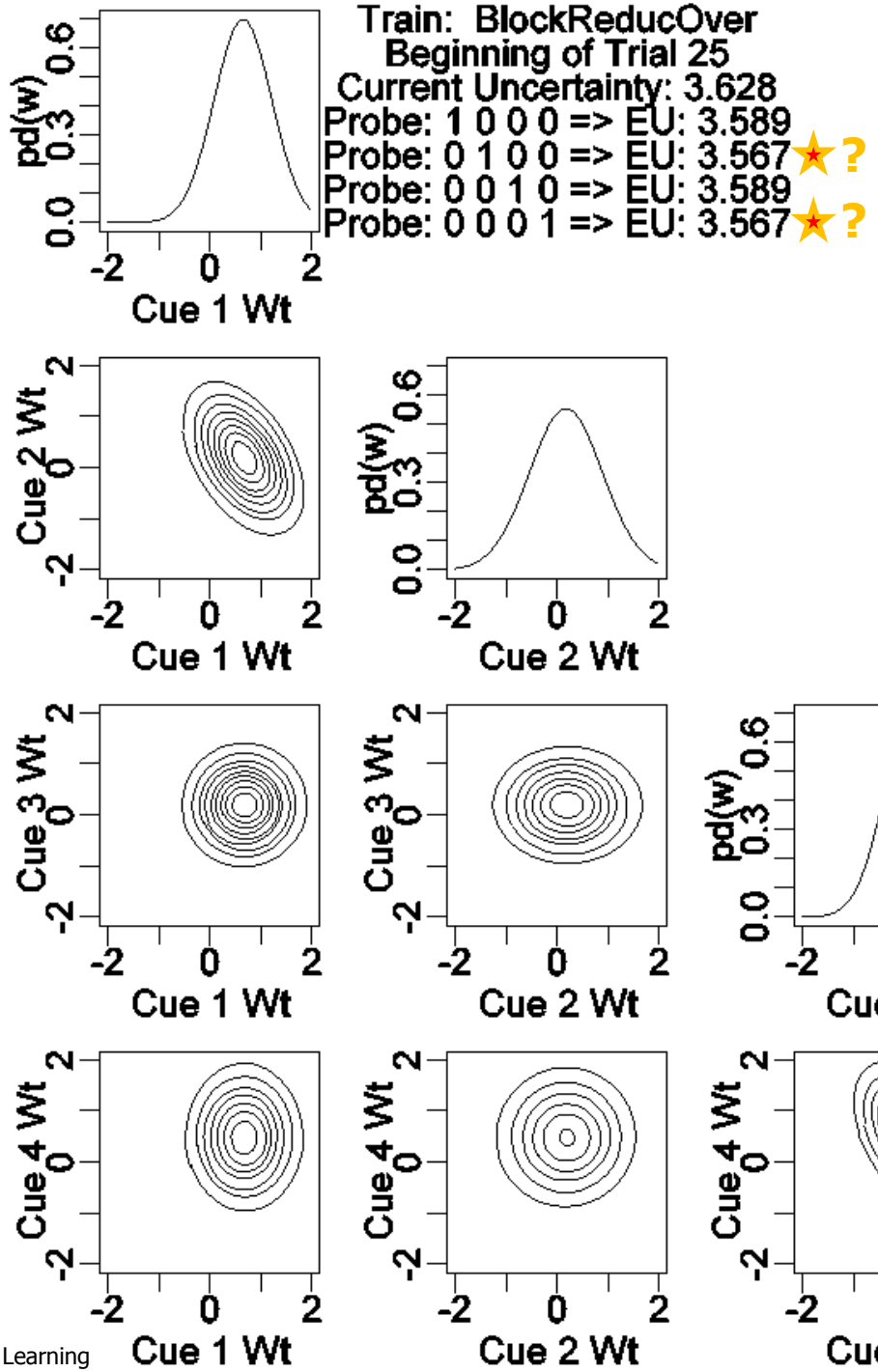

Train: BlockReducOver
Beginning of Trial 20
Current Uncertainty: 3.963
Probe: 1 0 0 0 => EU: 3.923
Probe: 0 1 0 0 => EU: 3.895
Probe: 0 0 1 0 => EU: 3.922
Probe: 0 0 0 1 => EU: 3.889
pd(w)
0.0
0.3
0.6
-2
0
2
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt


Train: BlockReducOver
Beginning of Trial 21
Current Uncertainty: 3.884
Probe: 1 0 0 0 => EU: 3.844
Probe: 0 1 0 0 => EU: 3.816
Probe: 0 0 1 0 => EU: 3.844
Probe: 0 0 0 1 => EU: 3.816
pd(w)
0.0
0.3
0.6
-2
0
2
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt


Train: BlockReducOver
Beginning of Trial 22
Current Uncertainty: 3.816
Probe: 1 0 0 0 => EU: 3.777
Probe: 0 1 0 0 => EU: 3.752
Probe: 0 0 1 0 => EU: 3.776
Probe: 0 0 0 1 => EU: 3.748
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt

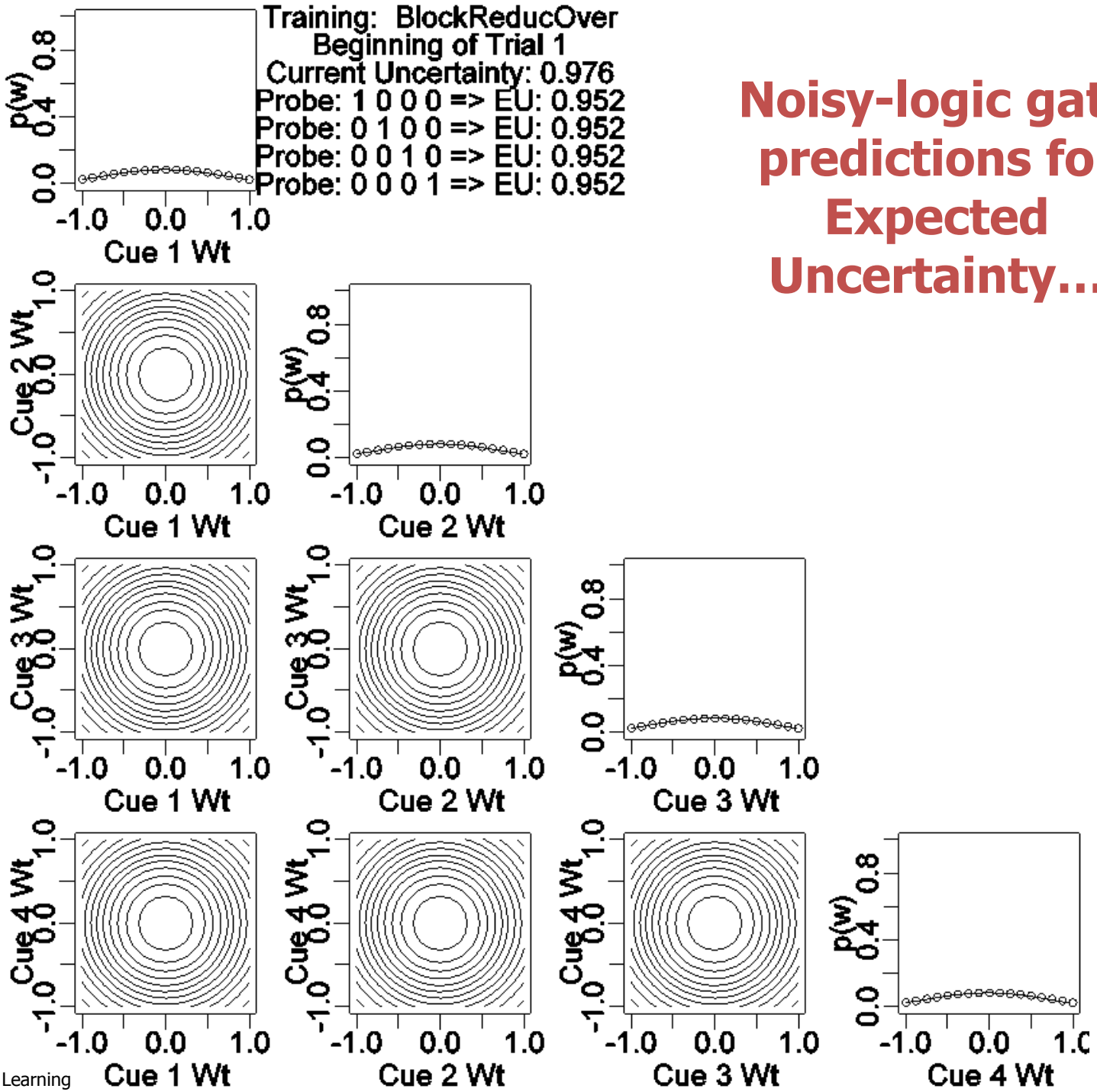

Train: BlockReducOver
Beginning of Trial 23
Current Uncertainty: 3.748
Probe: 1 0 0 0 => EU: 3.709
Probe: 0 1 0 0 => EU: 3.684
Probe: 0 0 1 0 => EU: 3.709
Probe: 0 0 0 1 => EU: 3.684
pd(w)
0.0
0.3
0.6
-2
0
2
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt


Train: BlockReducOver
Beginning of Trial 24
Current Uncertainty: 3.688
Probe: 1 0 0 0 => EU: 3.649
Probe: 0 1 0 0 => EU: 3.627
Probe: 0 0 1 0 => EU: 3.649
Probe: 0 0 0 1 => EU: 3.624
pd(w)
0.0
0.3
0.6
-2
0
2
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt

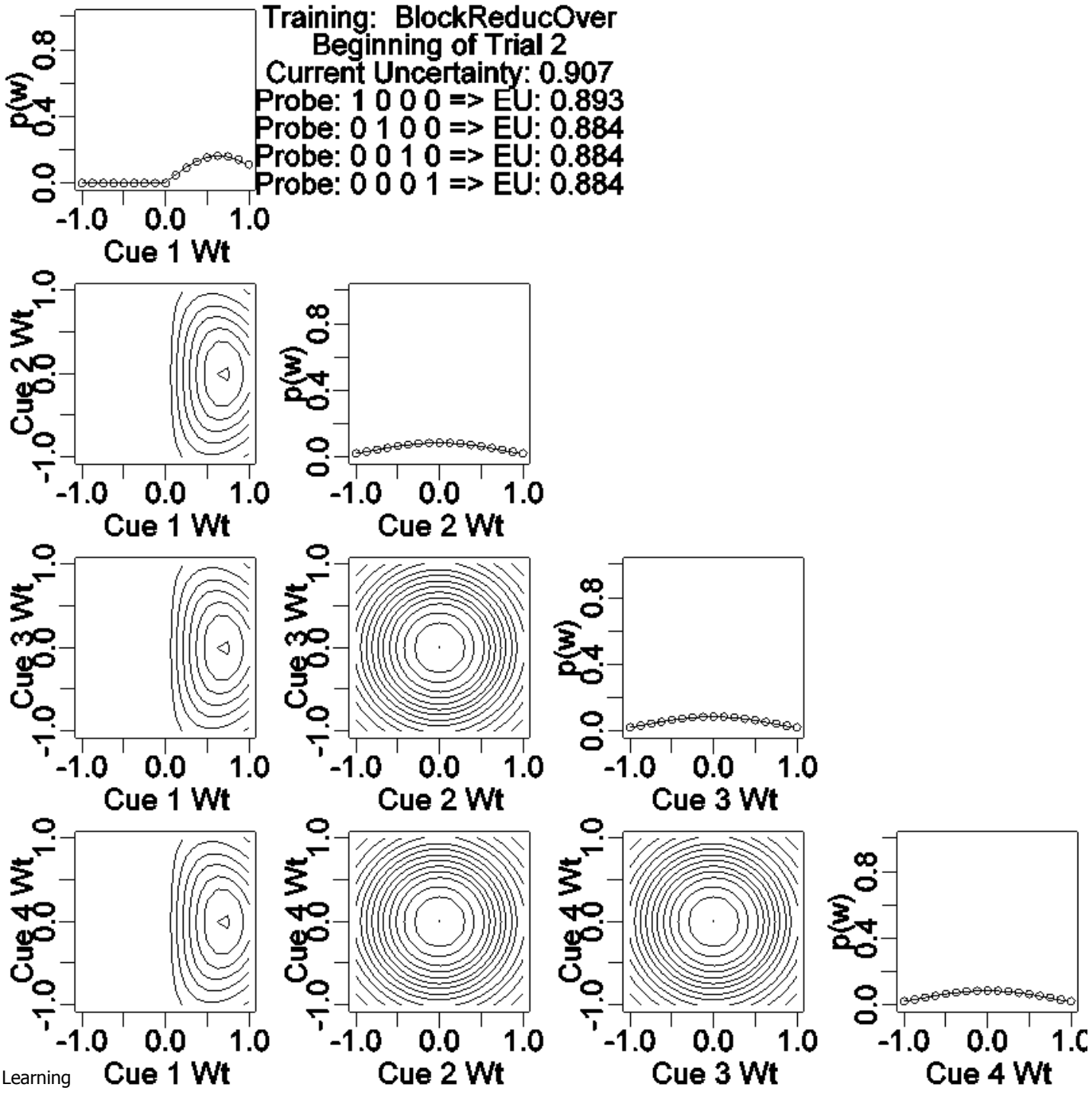


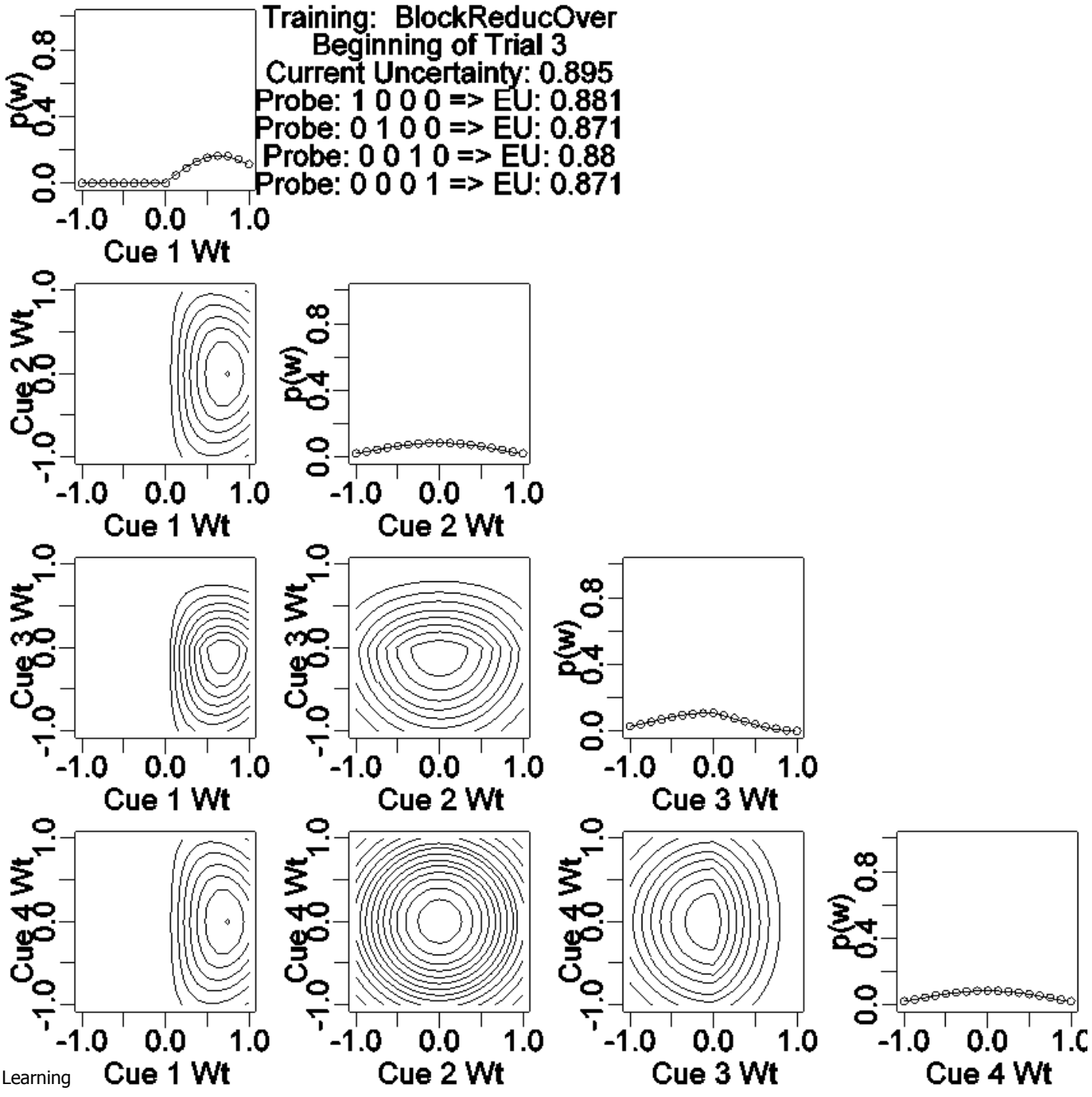
Kalman filter predictions: No preference!

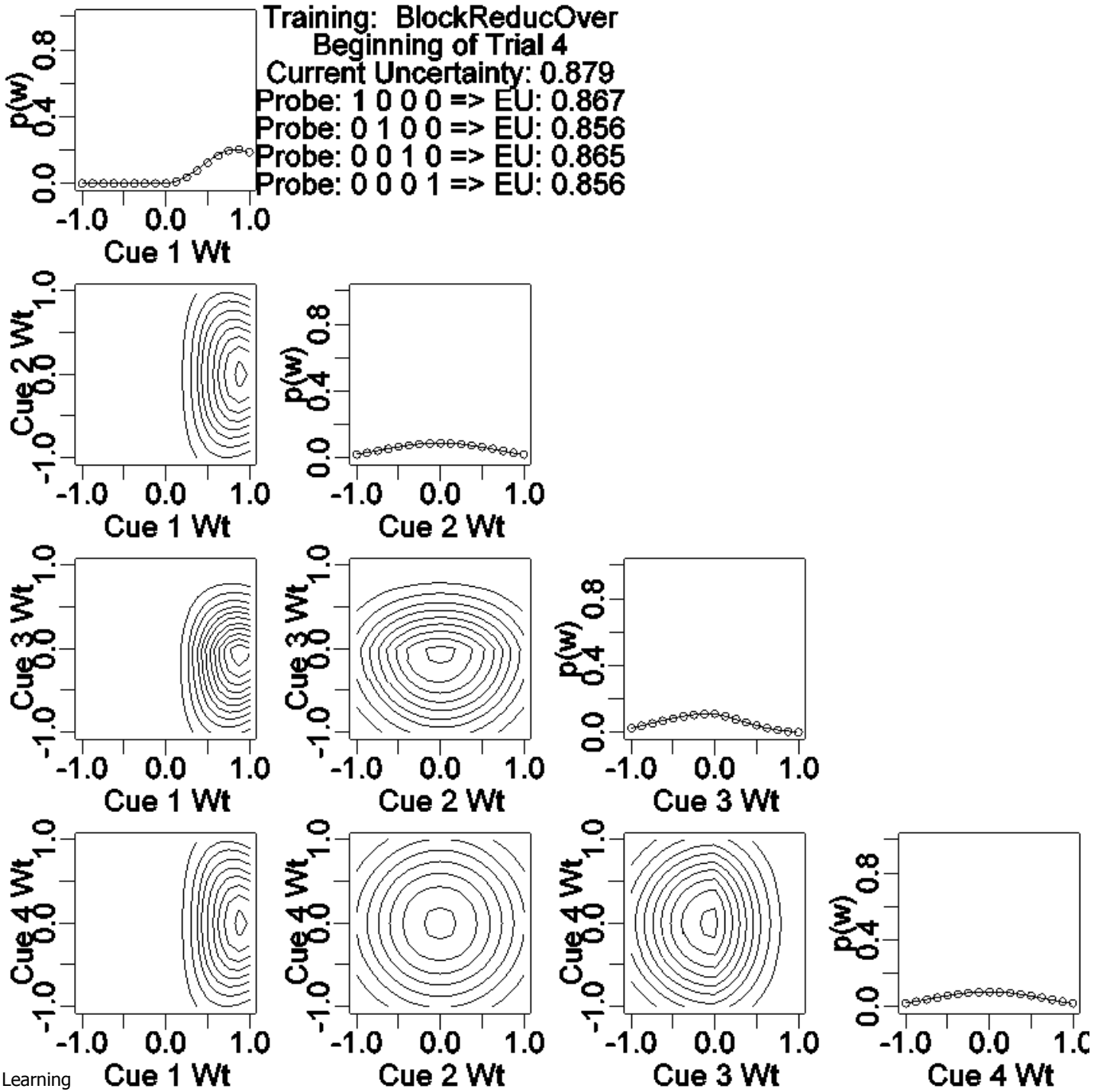
Same for Max P.

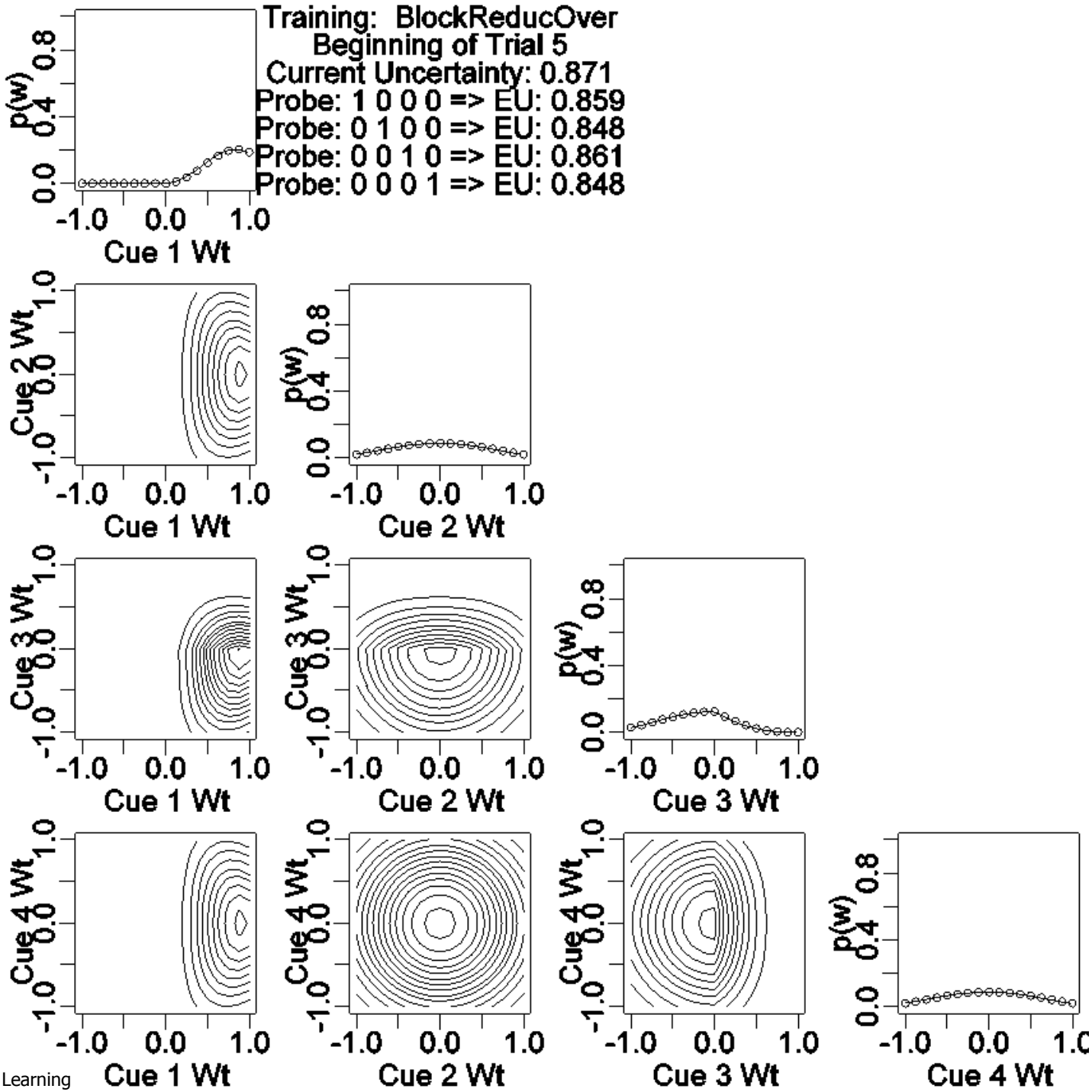


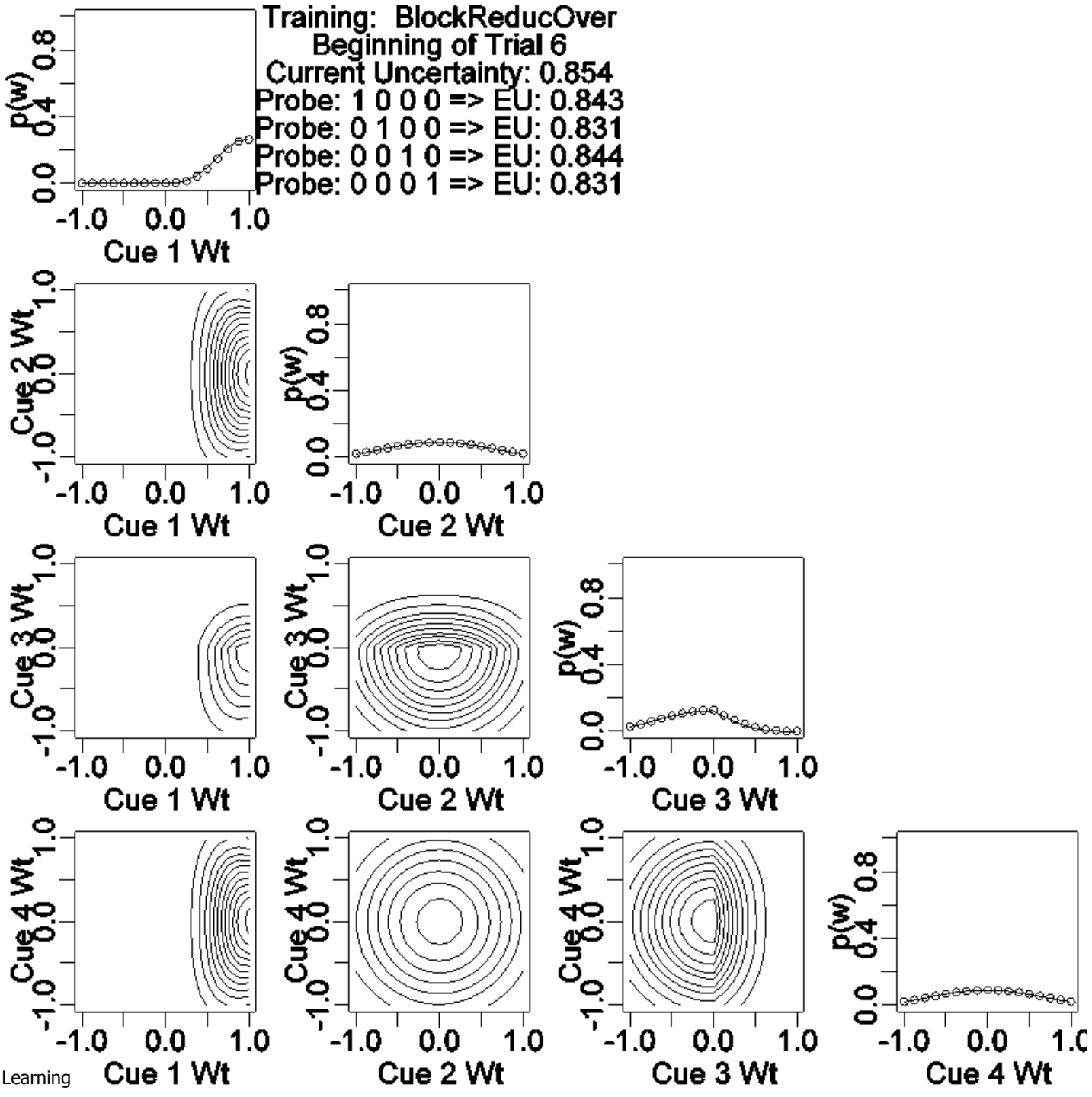
Noisy-logic gate predictions for Expected Uncertainty…

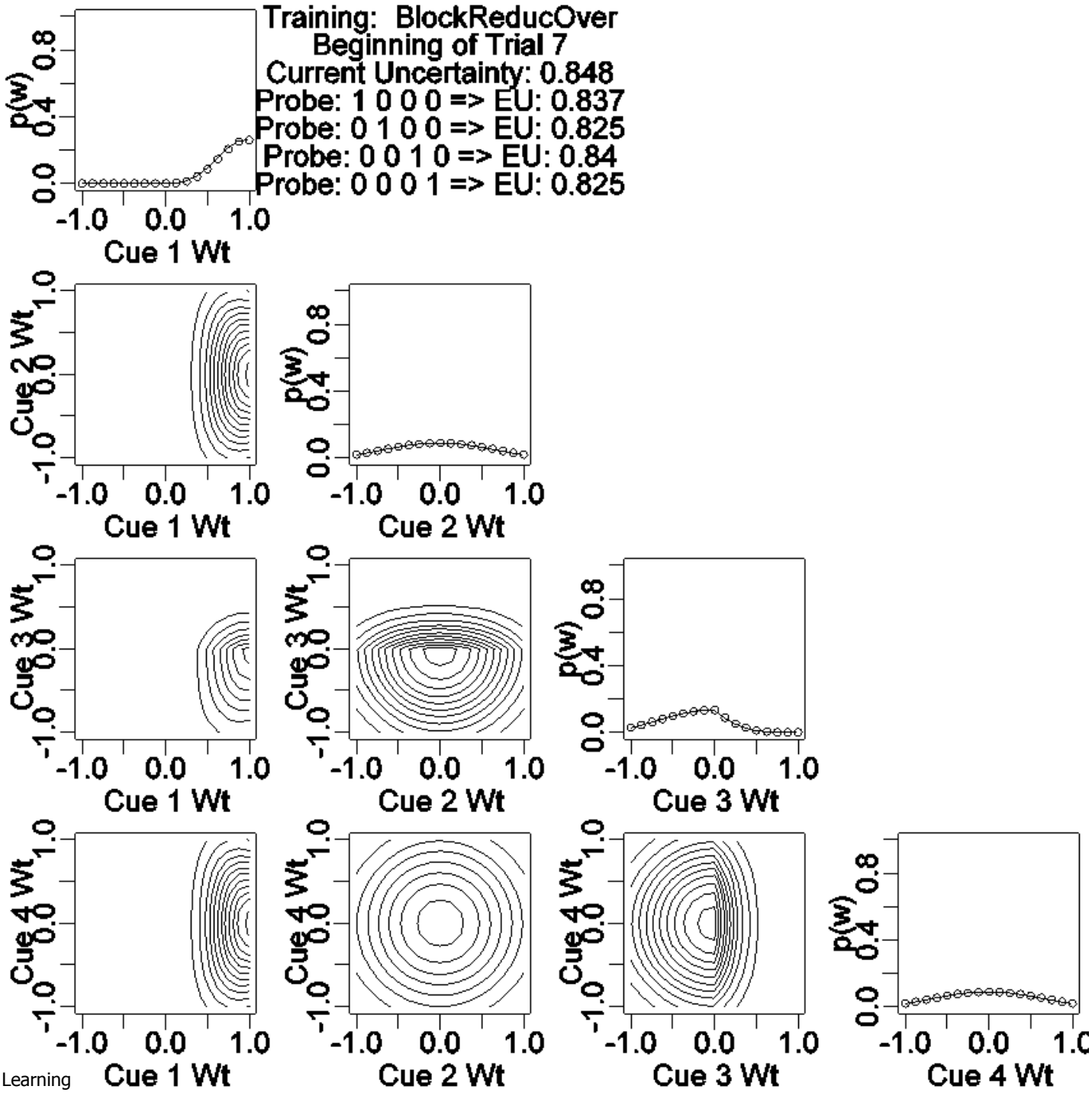

Training: BlockReducOver
Beginning of Trial 2
Current Uncertainty: 0.907
Probe: 1 0 0 0 => EU: 0.893
Probe: 0 1 0 0 => EU: 0.884
Probe: 0 0 1 0 => EU: 0.884
Probe: 0 0 0 1 => EU: 0.884
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8

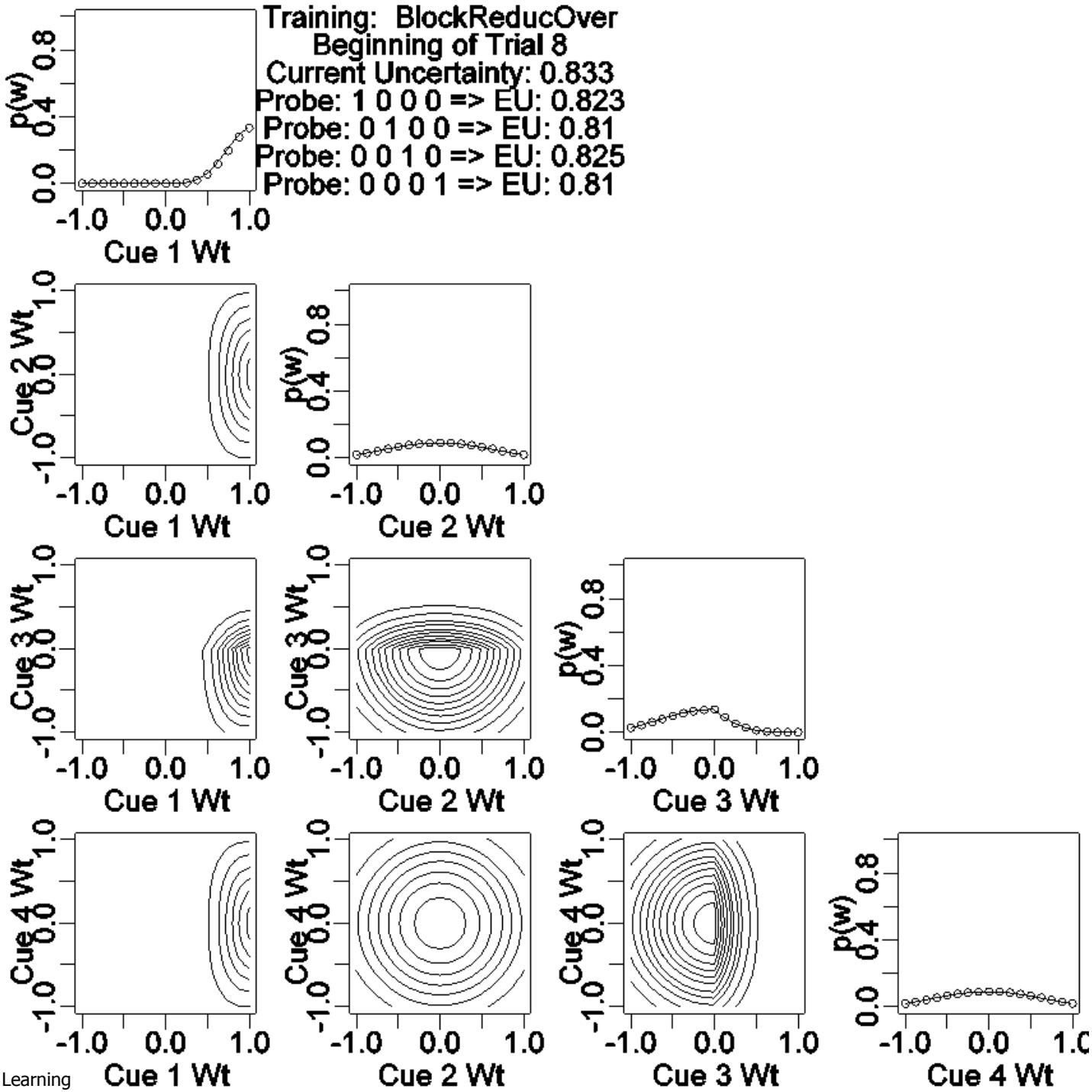

Training: BlockReducOver
Beginning of Trial 3
Current Uncertainty: 0.895
Probe: 1 0 0 0 => EU: 0.881
Probe: 0 1 0 0 => EU: 0.871
Probe: 0 0 1 0 => EU: 0.88
Probe: 0 0 0 1 => EU: 0.871
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8

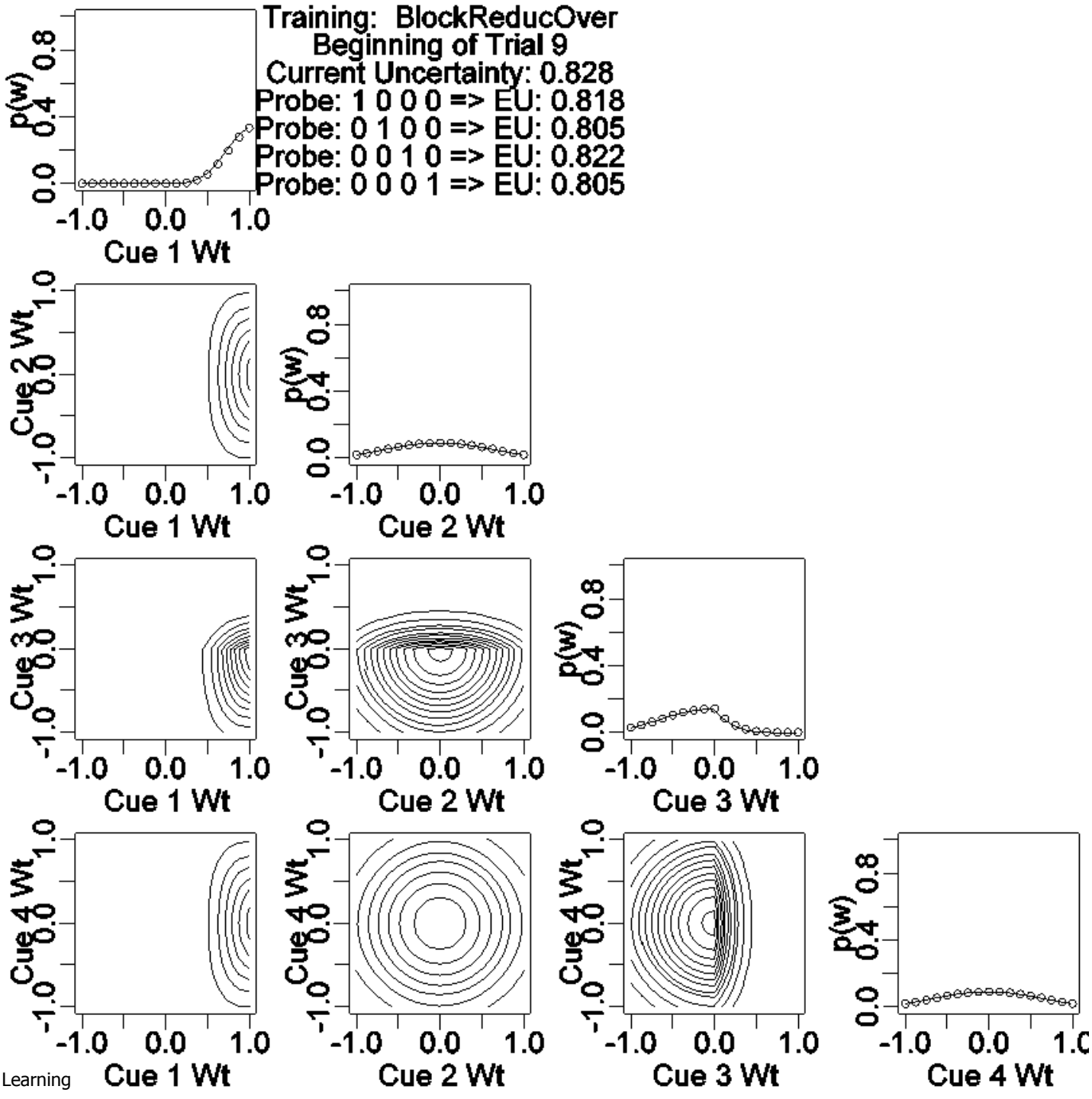

Training: BlockReducOver
Beginning of Trial 4
Current Uncertainty: 0.879
Probe: 1 0 0 0 => EU: 0.867
Probe: 0 1 0 0 => EU: 0.856
Probe: 0 0 1 0 => EU: 0.865
Probe: 0 0 0 1 => EU: 0.856
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8

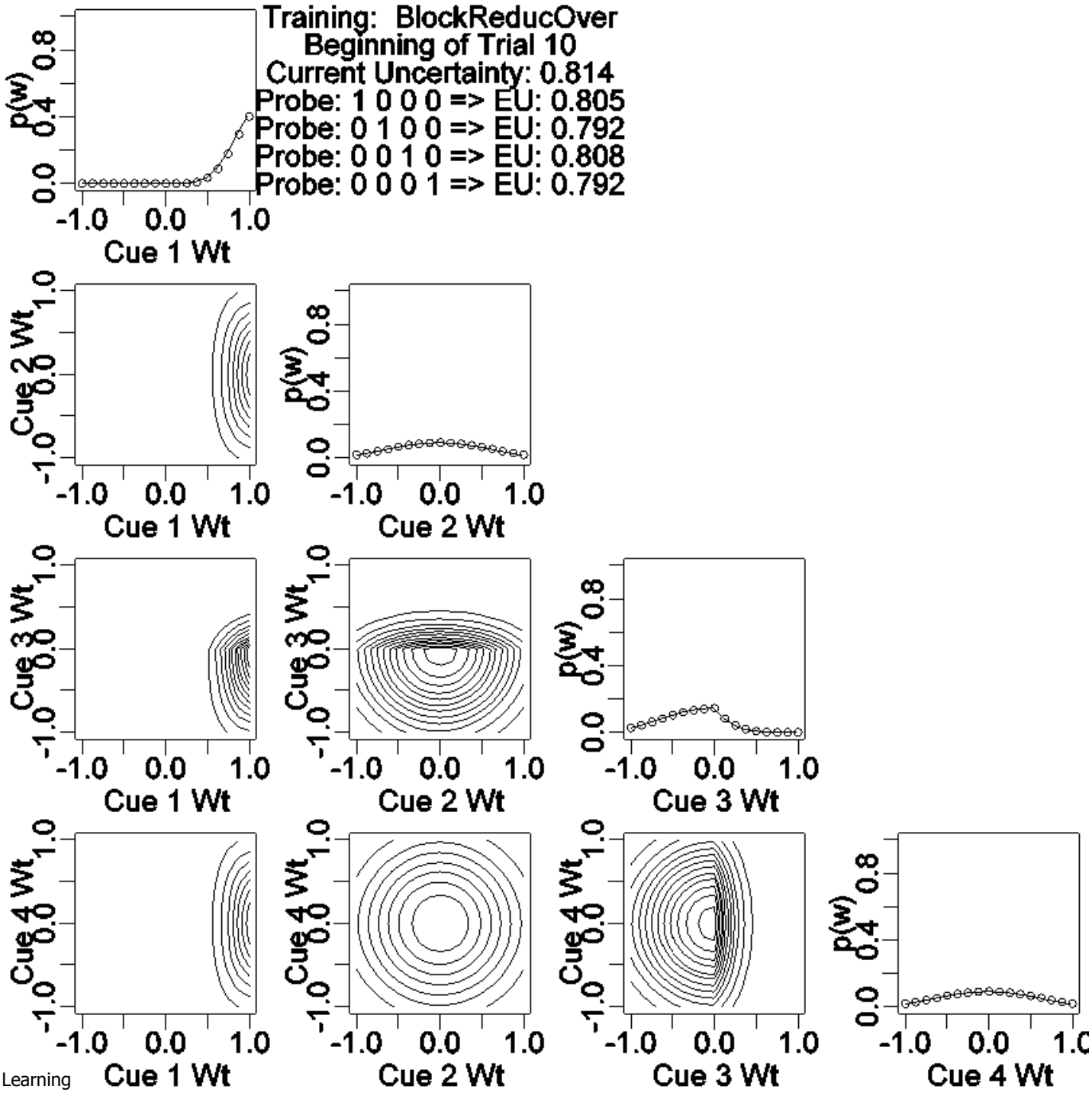

Training: BlockReducOver
Beginning of Trial 5
Current Uncertainty: 0.871
Probe: 1 0 0 0 => EU: 0.859
Probe: 0 1 0 0 => EU: 0.848
Probe: 0 0 1 0 => EU: 0.861
Probe: 0 0 0 1 => EU: 0.848
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8

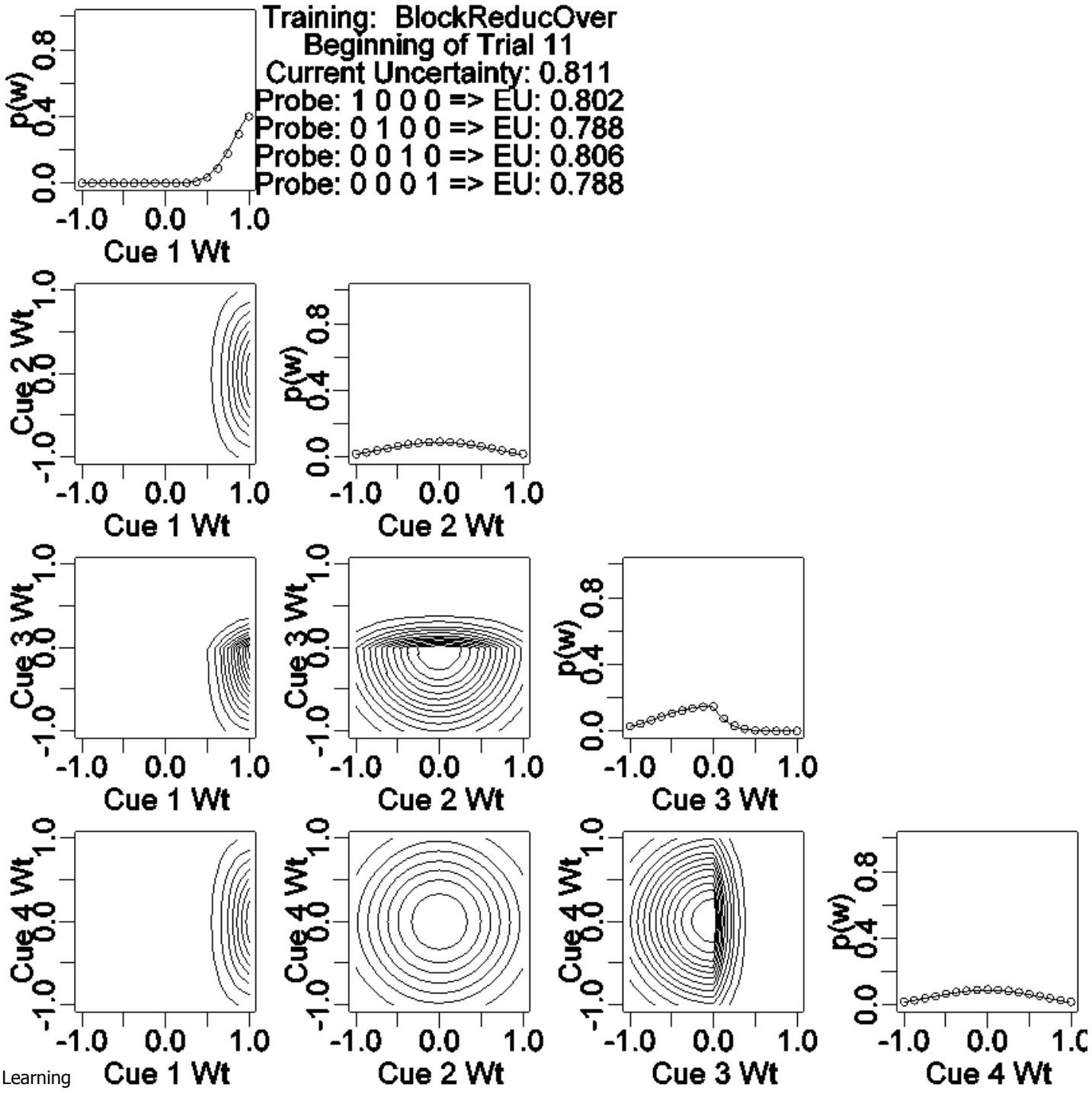

Training: BlockReducOver
Beginning of Trial 6
Current Uncertainty: 0.854
Probe: 1 0 0 0 => EU: 0.843
Probe: 0 1 0 0 => EU: 0.831
Probe: 0 0 1 0 => EU: 0.844
Probe: 0 0 0 1 => EU: 0.831
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8

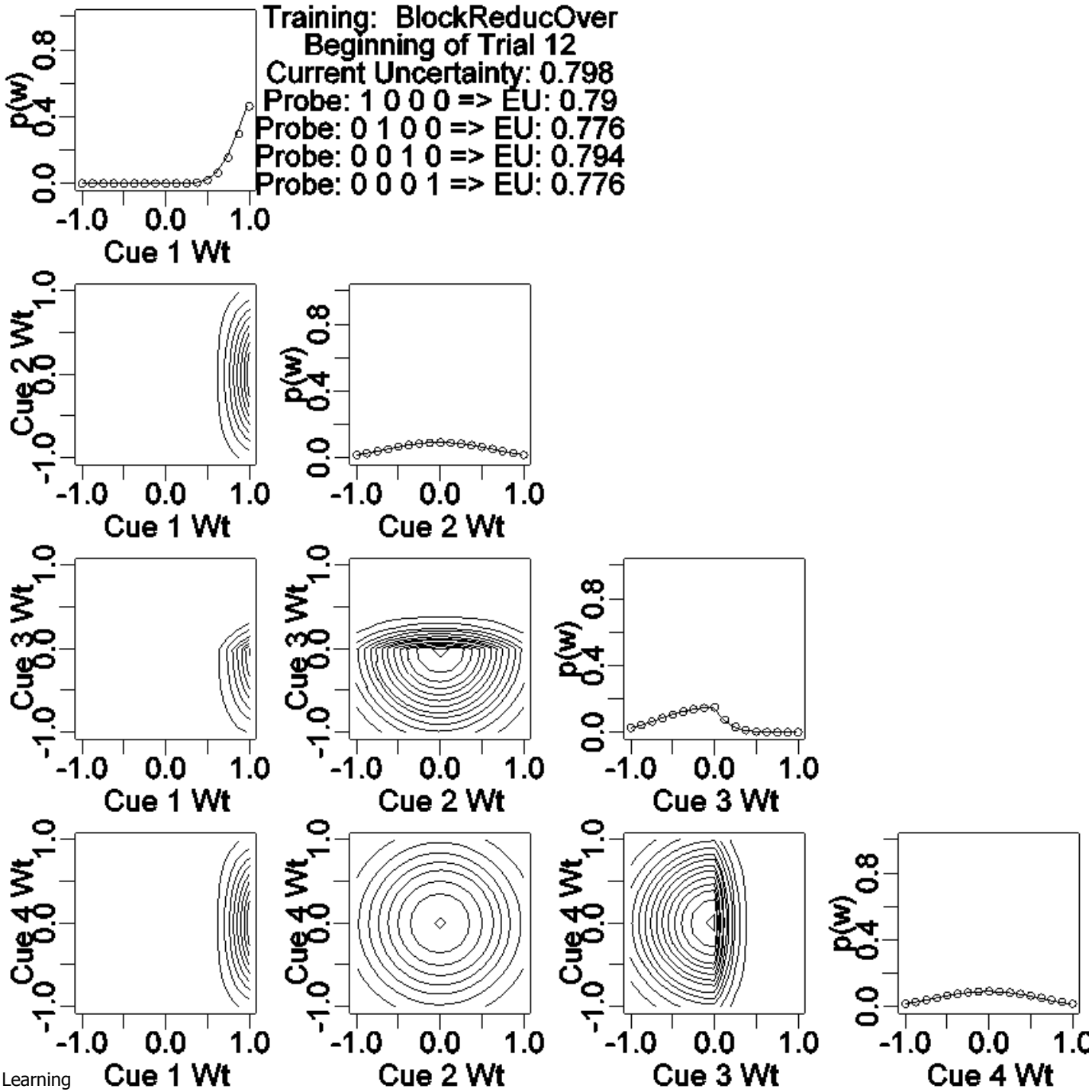

Training: BlockReducOver
Beginning of Trial 7
Current Uncertainty: 0.848
Probe: 1 0 0 0 => EU: 0.837
Probe: 0 1 0 0 => EU: 0.825
Probe: 0 0 1 0 => EU: 0.84
Probe: 0 0 0 1 => EU: 0.825
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8

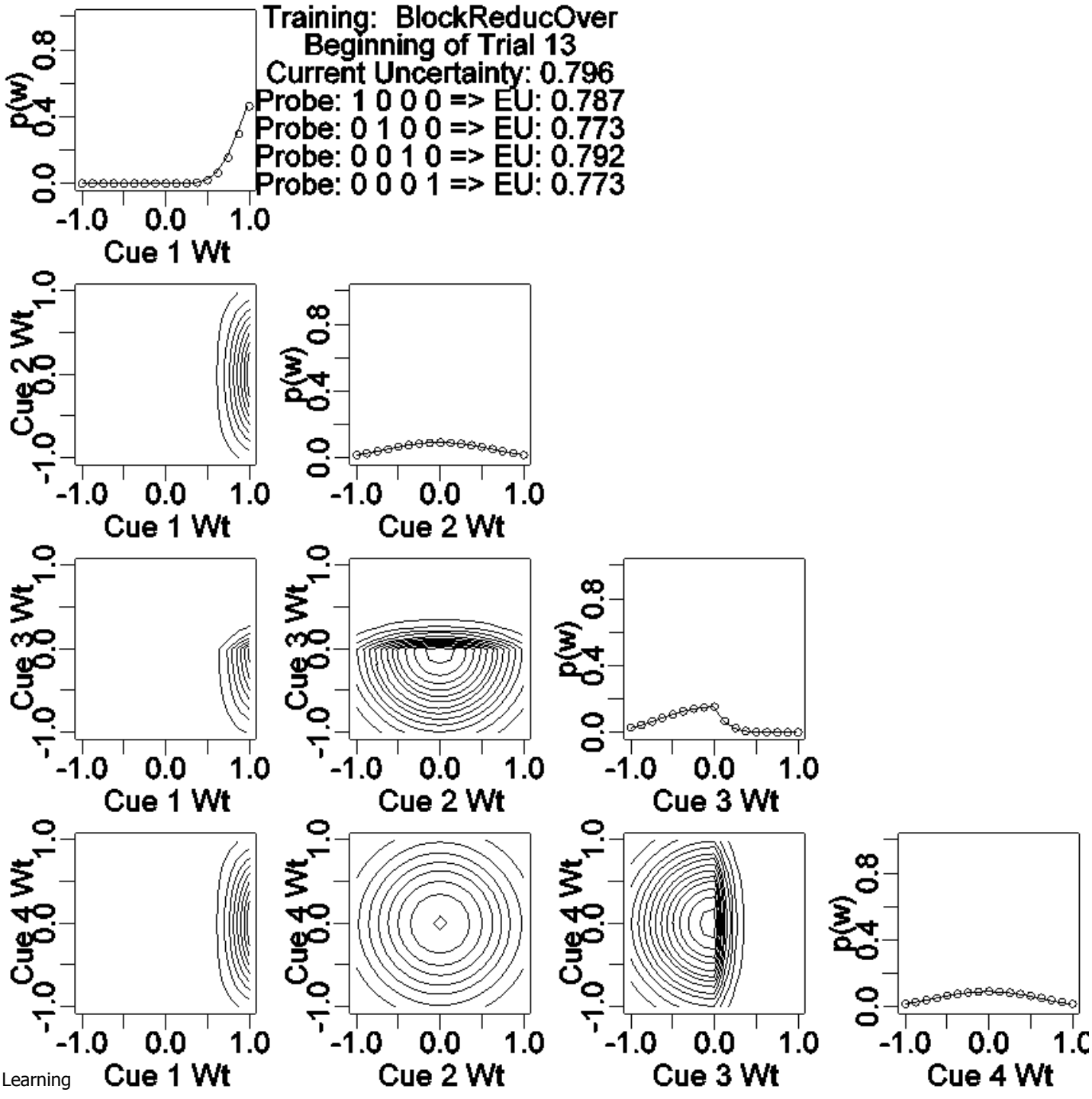

Training: BlockReducOver
Beginning of Trial 8
Current Uncertainty: 0.833
Probe: 1 0 0 0 => EU: 0.823
Probe: 0 1 0 0 => EU: 0.81
Probe: 0 0 1 0 => EU: 0.825
Probe: 0 0 0 1 => EU: 0.81
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8

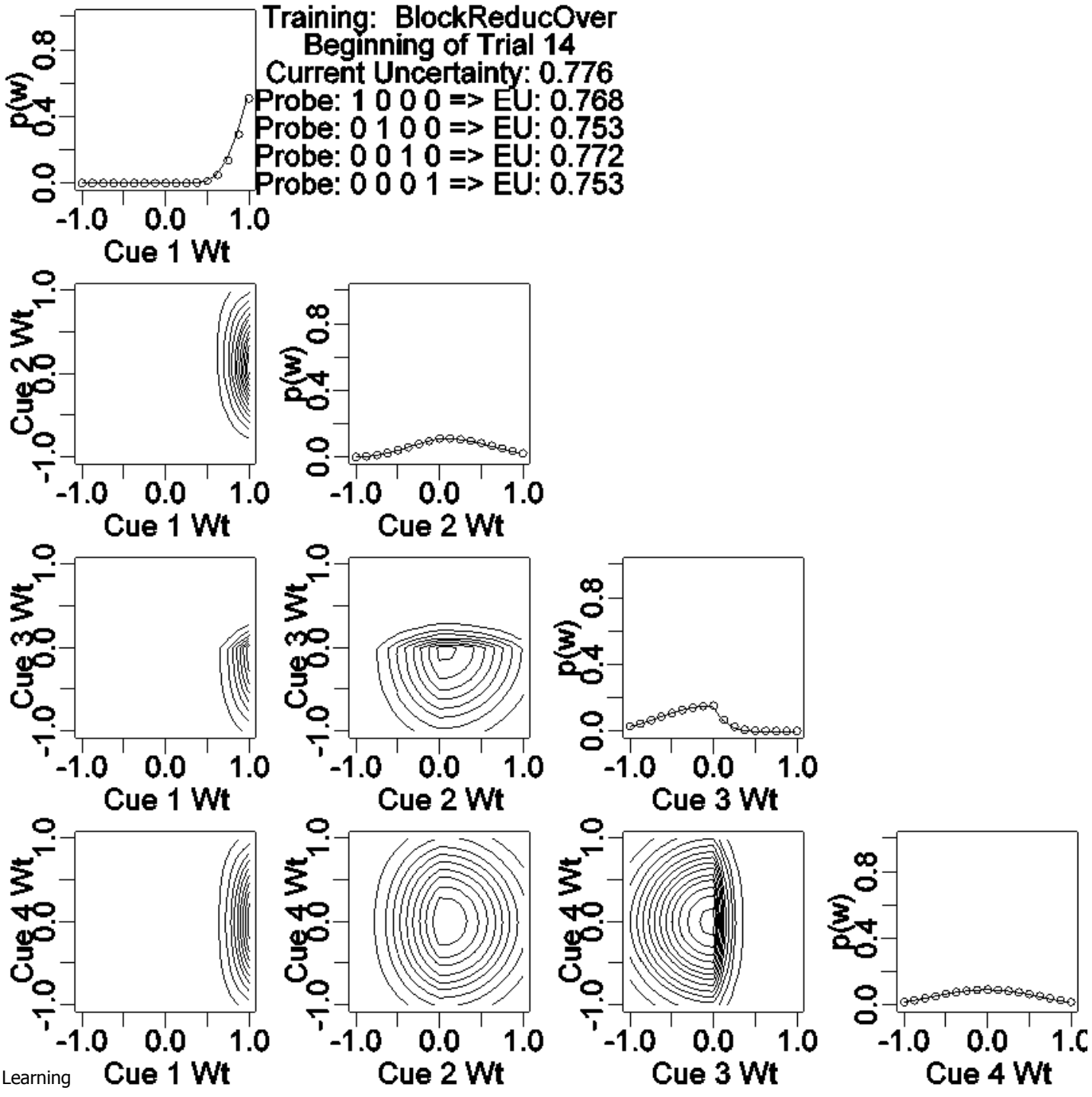

Training: BlockReducOver
Beginning of Trial 9
Current Uncertainty: 0.828
Probe: 1 0 0 0 => EU: 0.818
Probe: 0 1 0 0 => EU: 0.805
Probe: 0 0 1 0 => EU: 0.822
Probe: 0 0 0 1 => EU: 0.805
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8

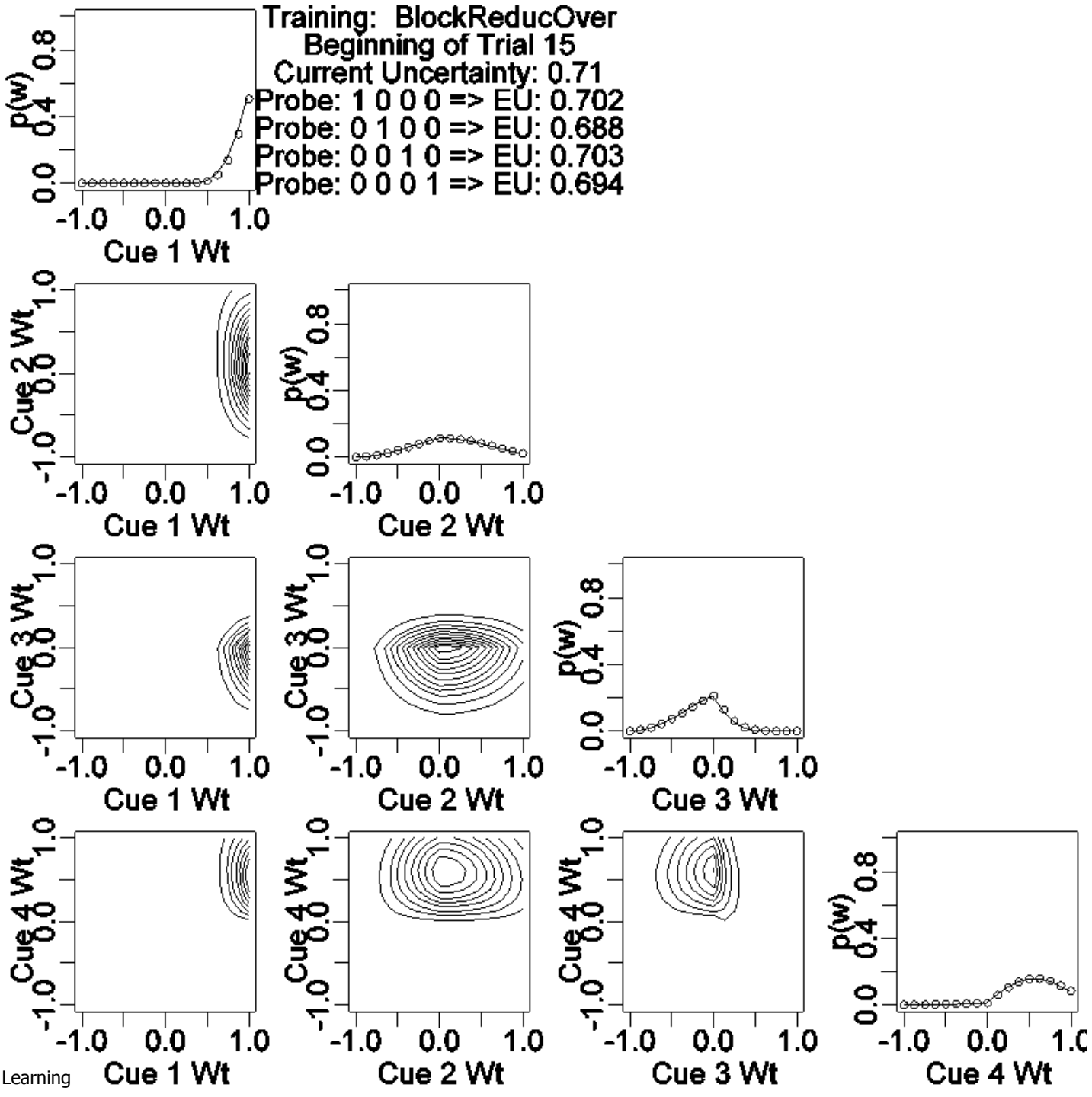

Training: BlockReducOver
Beginning of Trial 10
Current Uncertainty: 0.814
Probe: 1 0 0 0 => EU: 0.805
Probe: 0 1 0 0 => EU: 0.792
Probe: 0 0 1 0 => EU: 0.808
Probe: 0 0 0 1 => EU: 0.792
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8

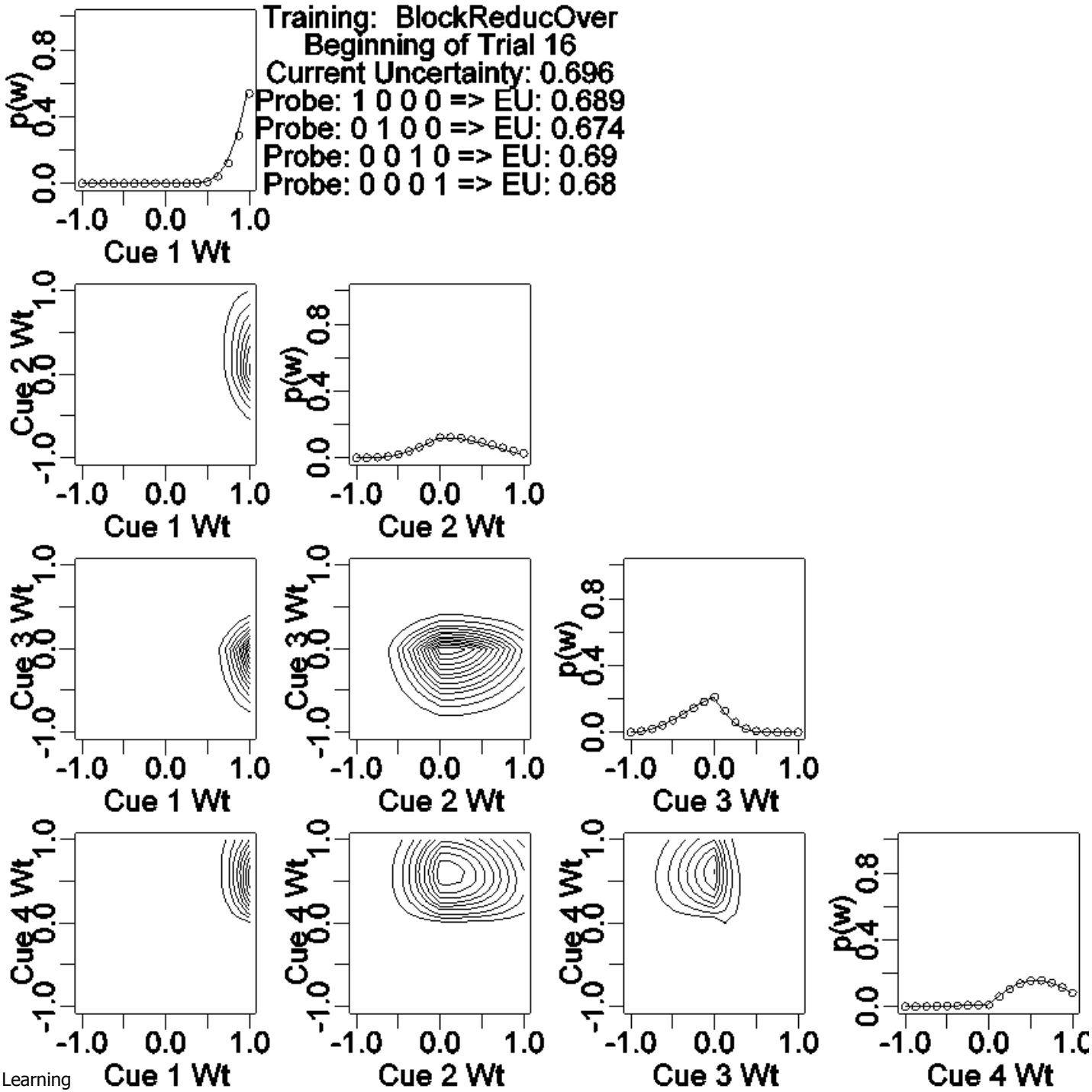

Training: BlockReducOver
Beginning of Trial 11
Current Uncertainty: 0.811
Probe: 1 0 0 0 => EU: 0.802
Probe: 0 1 0 0 => EU: 0.788
Probe: 0 0 1 0 => EU: 0.806
Probe: 0 0 0 1 => EU: 0.788
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8

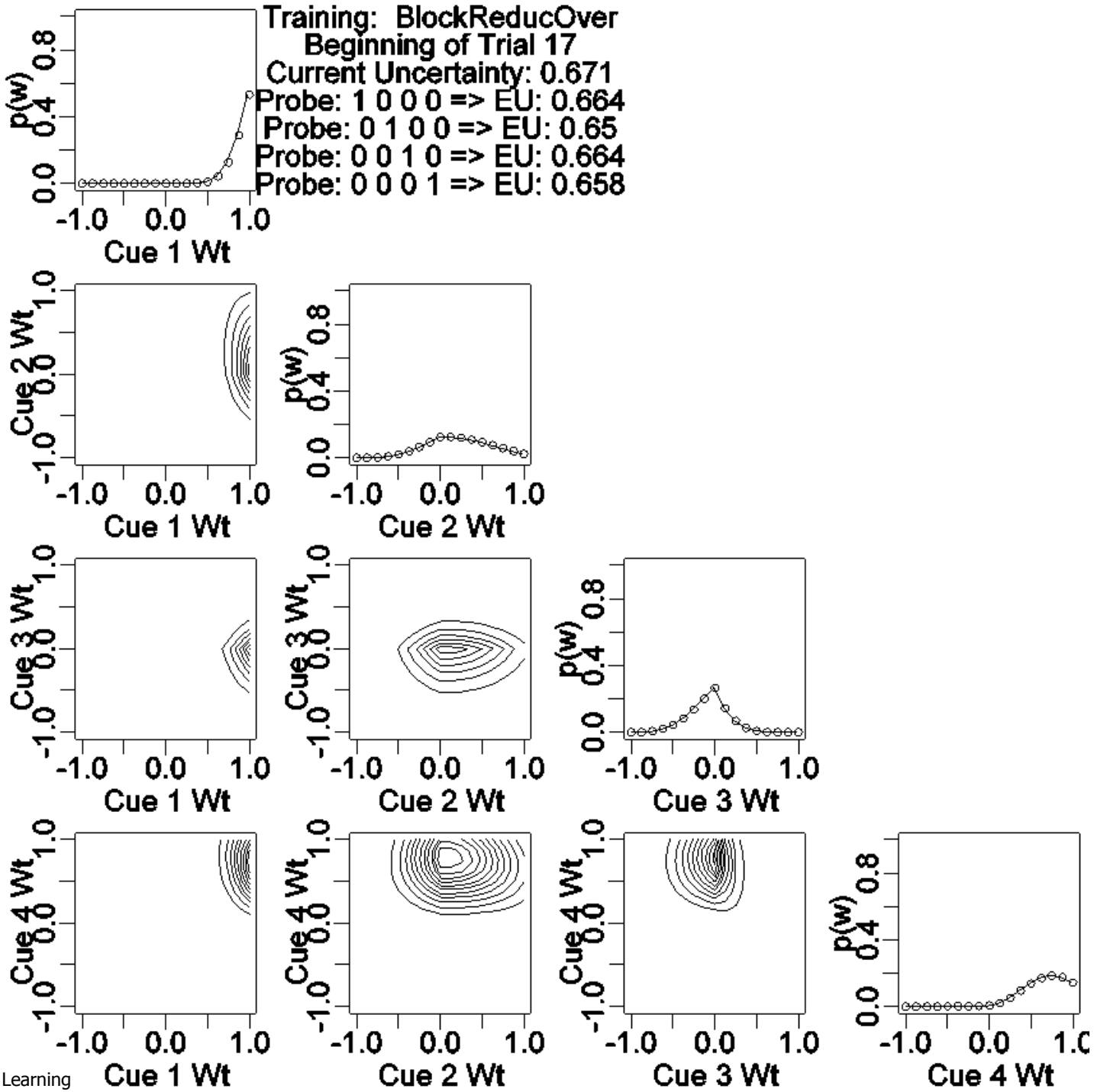

Training: BlockReducOver
Beginning of Trial 12
Current Uncertainty: 0.798
Probe: 1 0 0 0 => EU: 0.79
Probe: 0 1 0 0 => EU: 0.776
Probe: 0 0 1 0 => EU: 0.794
Probe: 0 0 0 1 => EU: 0.776
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8

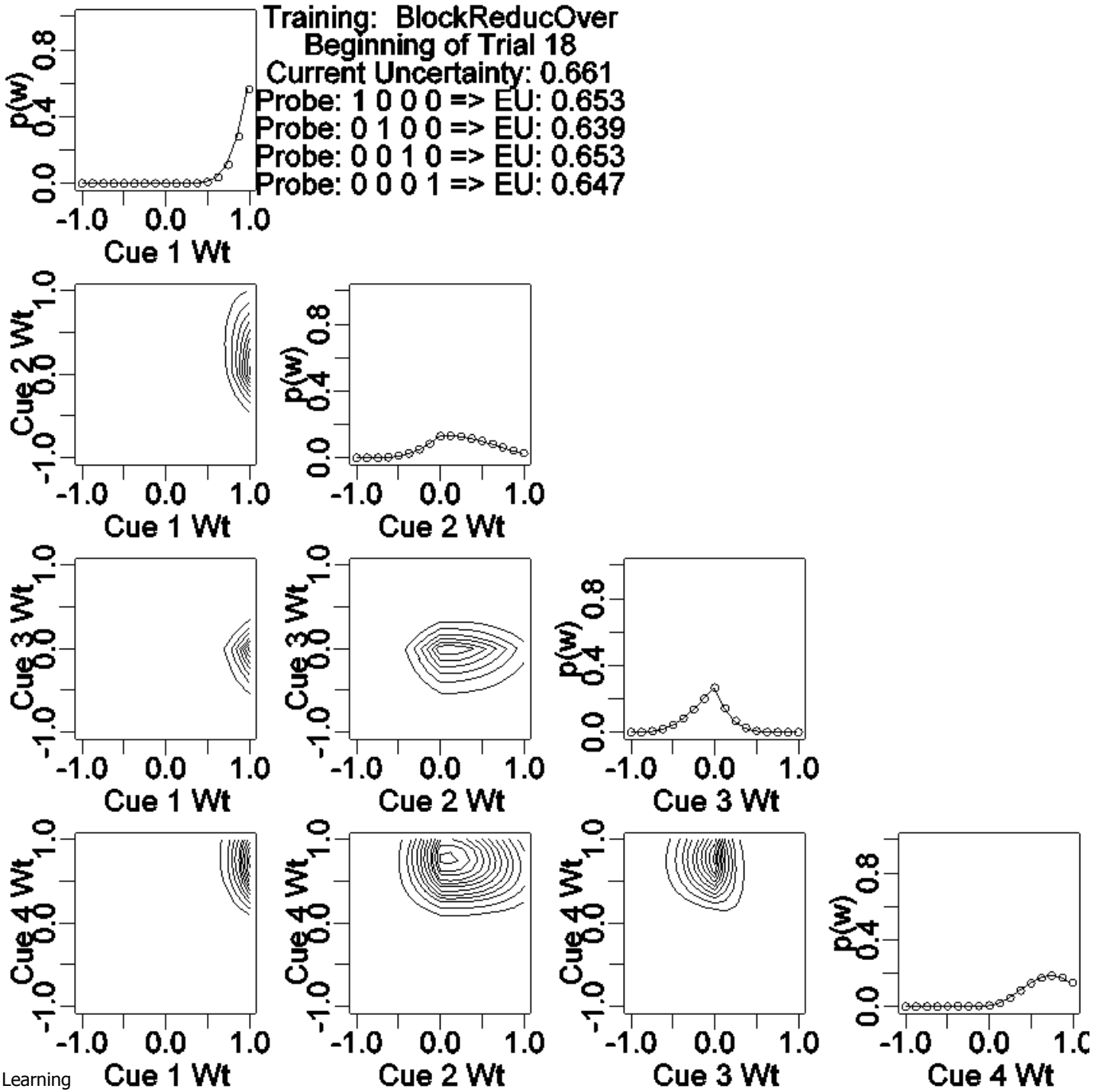

Training: BlockReducOver
Beginning of Trial 13
Current Uncertainty: 0.796
Probe: 1 0 0 0 => EU: 0.787
Probe: 0 1 0 0 => EU: 0.773
Probe: 0 0 1 0 => EU: 0.792
Probe: 0 0 0 1 => EU: 0.773
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8

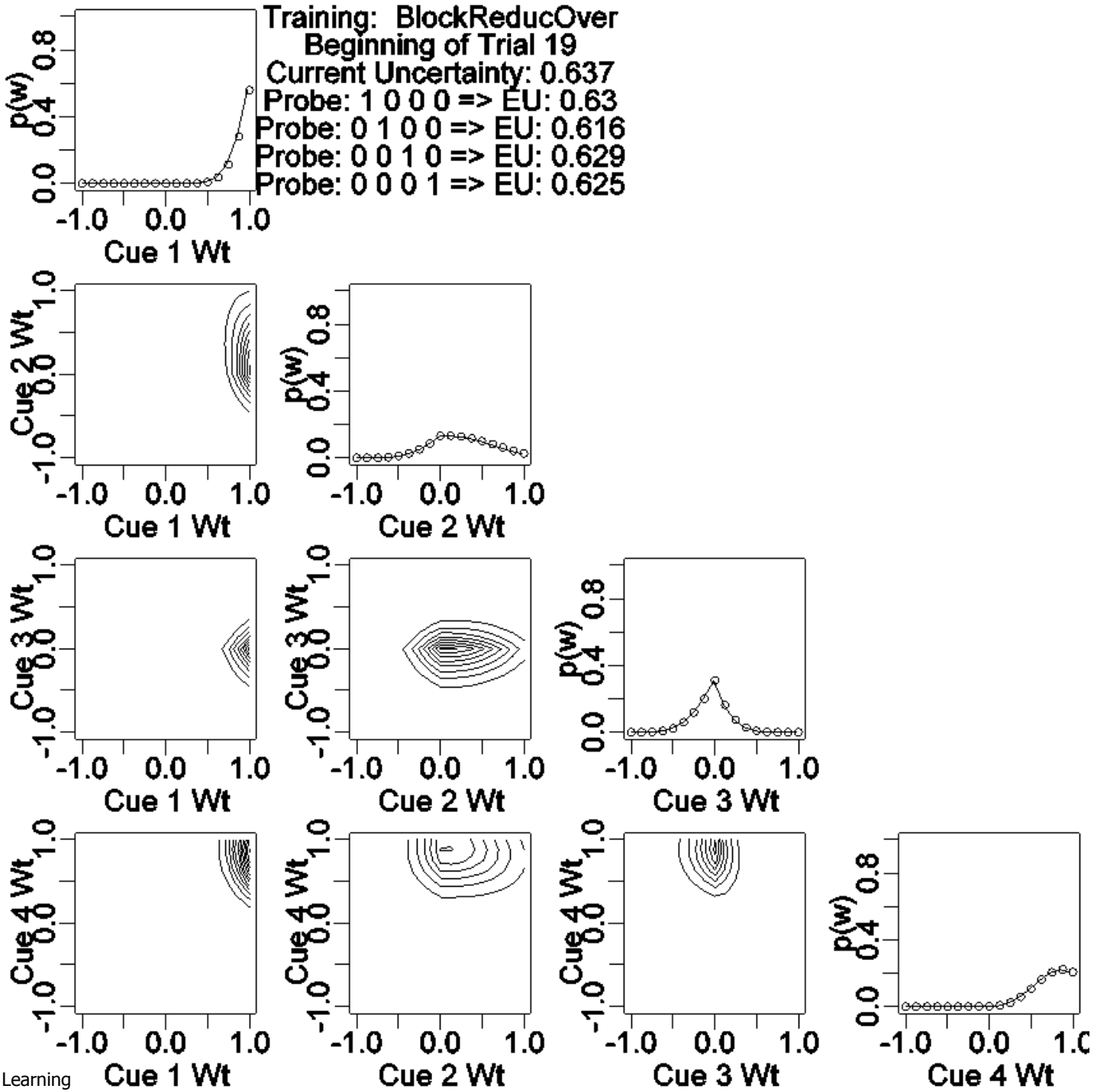

Training: BlockReducOver
Beginning of Trial 14
Current Uncertainty: 0.776
Probe: 1 0 0 0 => EU: 0.768
Probe: 0 1 0 0 => EU: 0.753
Probe: 0 0 1 0 => EU: 0.772
Probe: 0 0 0 1 => EU: 0.753
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8

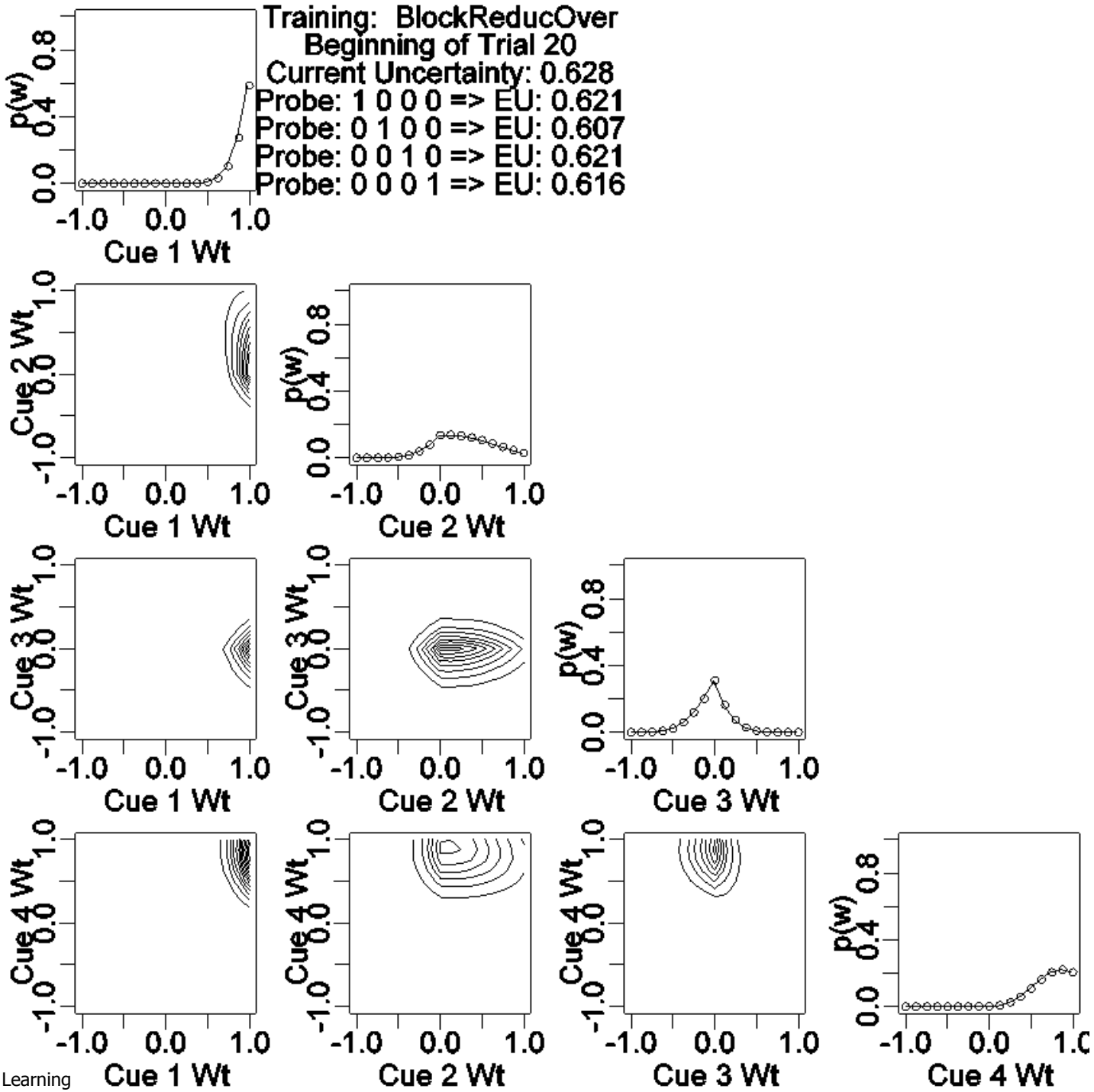

Training: BlockReducOver
Beginning of Trial 15
Current Uncertainty: 0.71
Probe: 1 0 0 0 => EU: 0.702
Probe: 0 1 0 0 => EU: 0.688
Probe: 0 0 1 0 => EU: 0.703
Probe: 0 0 0 1 => EU: 0.694
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8

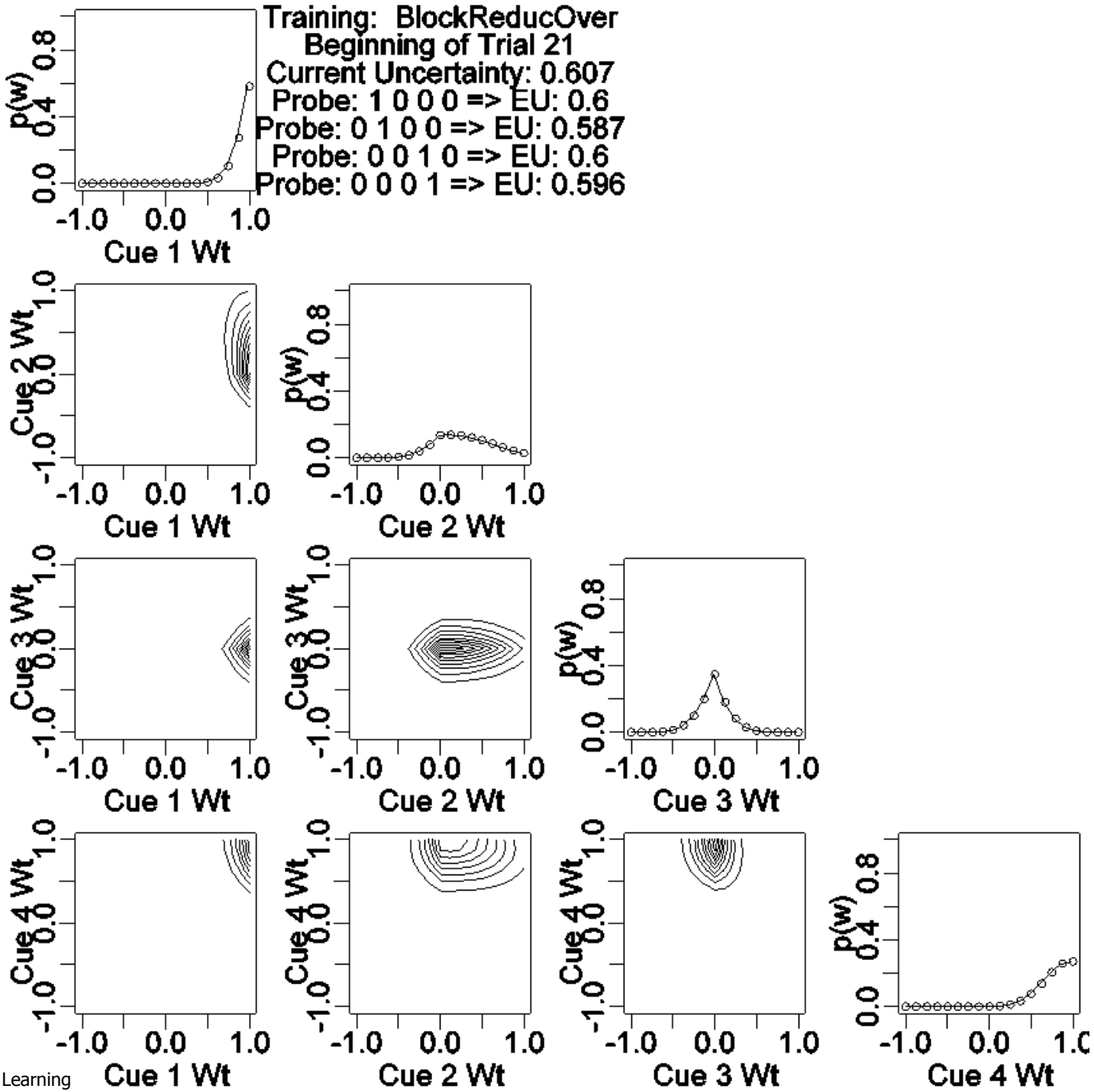

Training: BlockReducOver
Beginning of Trial 16
Current Uncertainty: 0.696
Probe: 1 0 0 0 => EU: 0.689
Probe: 0 1 0 0 => EU: 0.674
Probe: 0 0 1 0 => EU: 0.69
Probe: 0 0 0 1 => EU: 0.68
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8

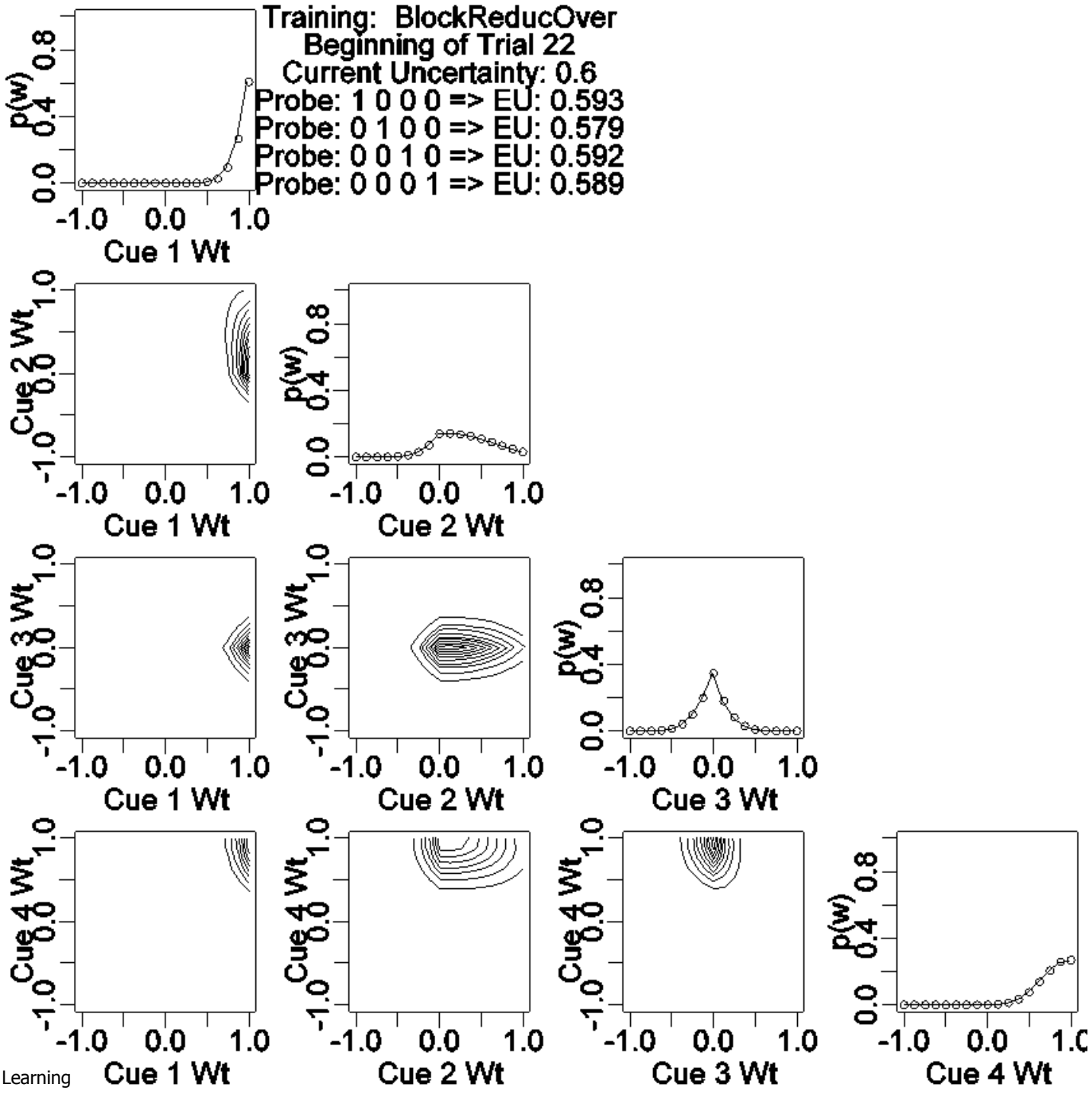

Training: BlockReducOver
Beginning of Trial 17
Current Uncertainty: 0.671
Probe: 1 0 0 0 => EU: 0.664
Probe: 0 1 0 0 => EU: 0.65
Probe: 0 0 1 0 => EU: 0.664
Probe: 0 0 0 1 => EU: 0.658
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8

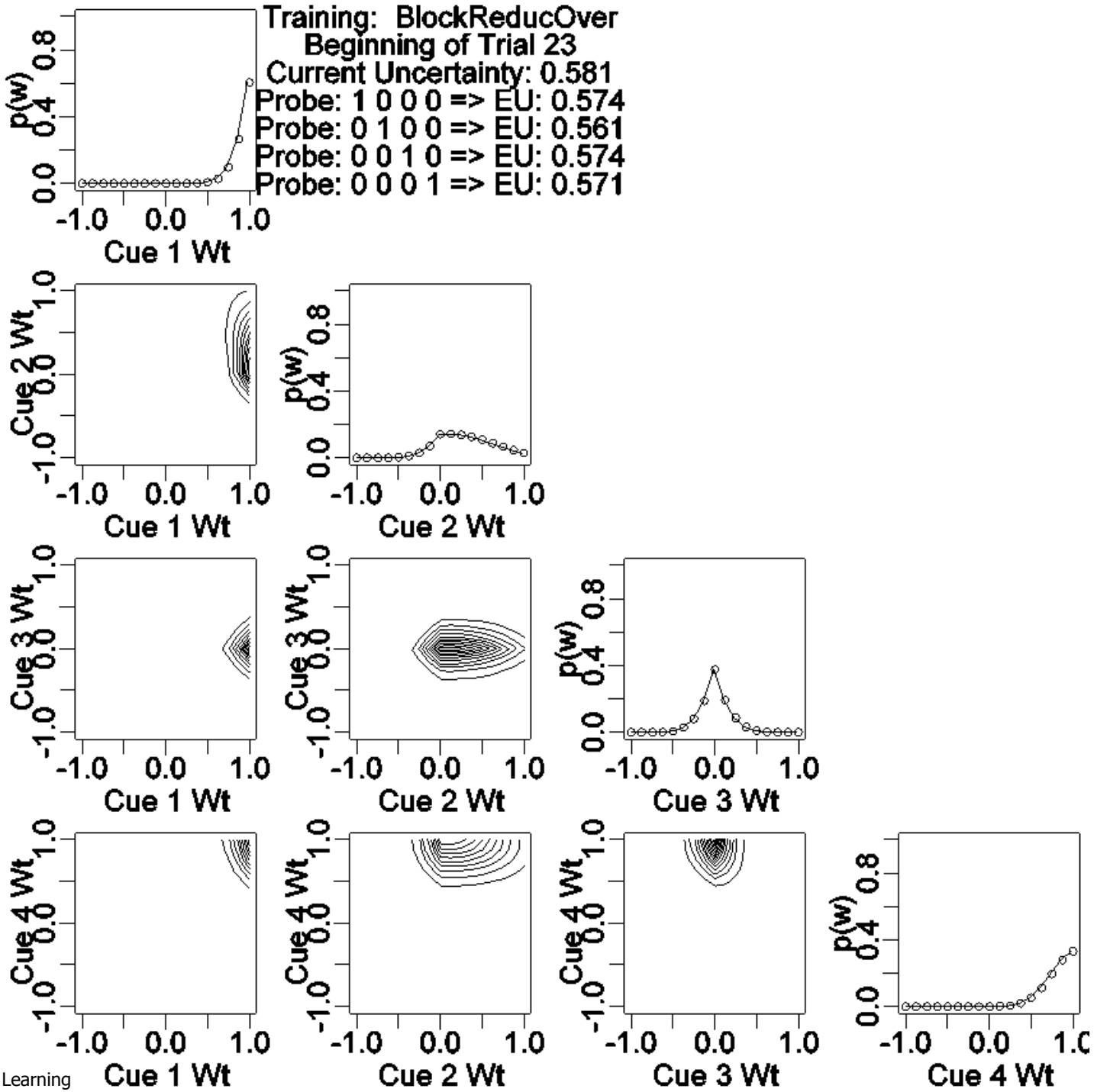

Training: BlockReducOver
Beginning of Trial 18
Current Uncertainty: 0.661
Probe: 1 0 0 0 => EU: 0.653
Probe: 0 1 0 0 => EU: 0.639
Probe: 0 0 1 0 => EU: 0.653
Probe: 0 0 0 1 => EU: 0.647
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8

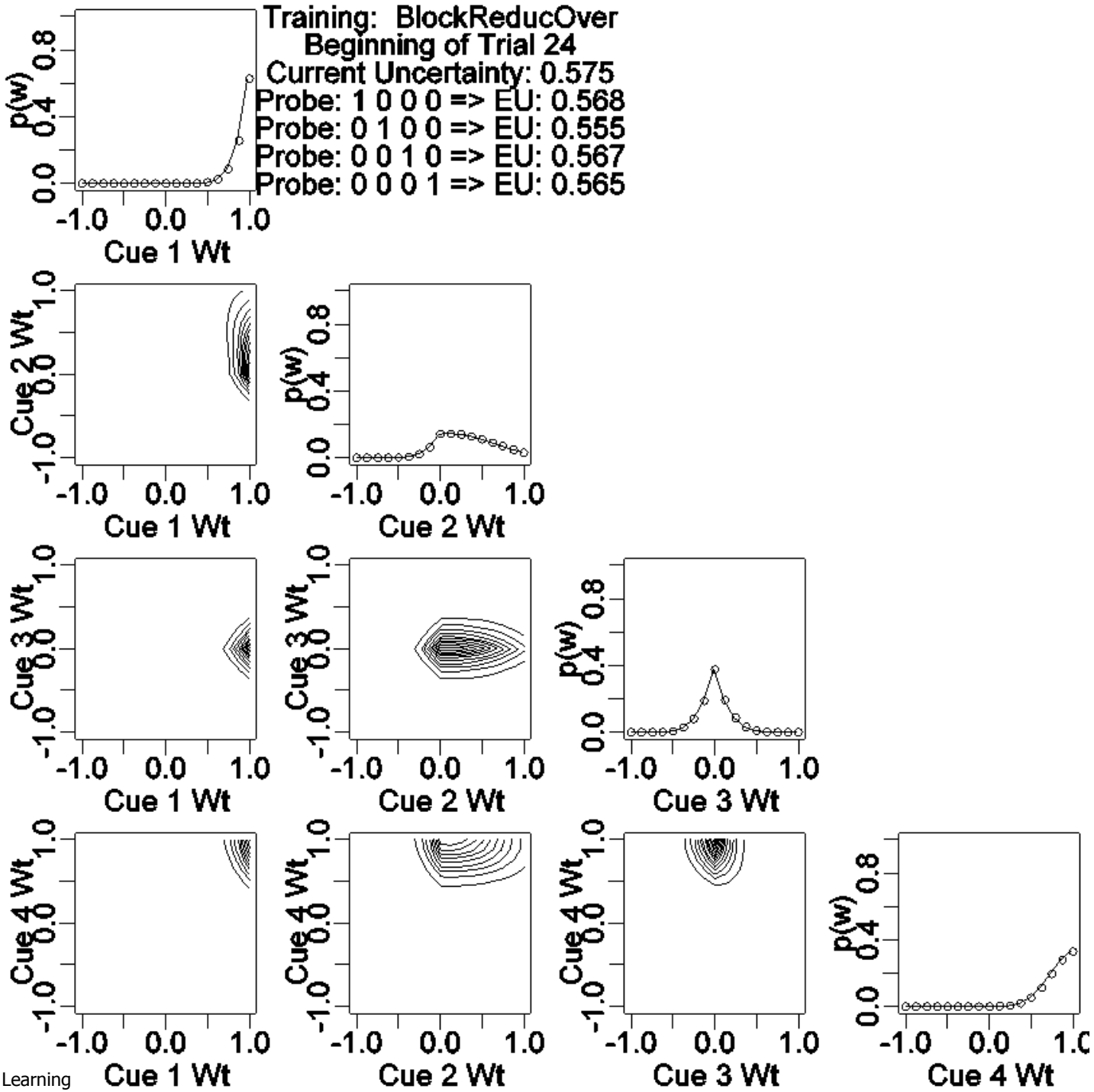

Training: BlockReducOver
Beginning of Trial 19
Current Uncertainty: 0.637
Probe: 1 0 0 0 => EU: 0.63
Probe: 0 1 0 0 => EU: 0.616
Probe: 0 0 1 0 => EU: 0.629
Probe: 0 0 0 1 => EU: 0.625
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.4
0.8

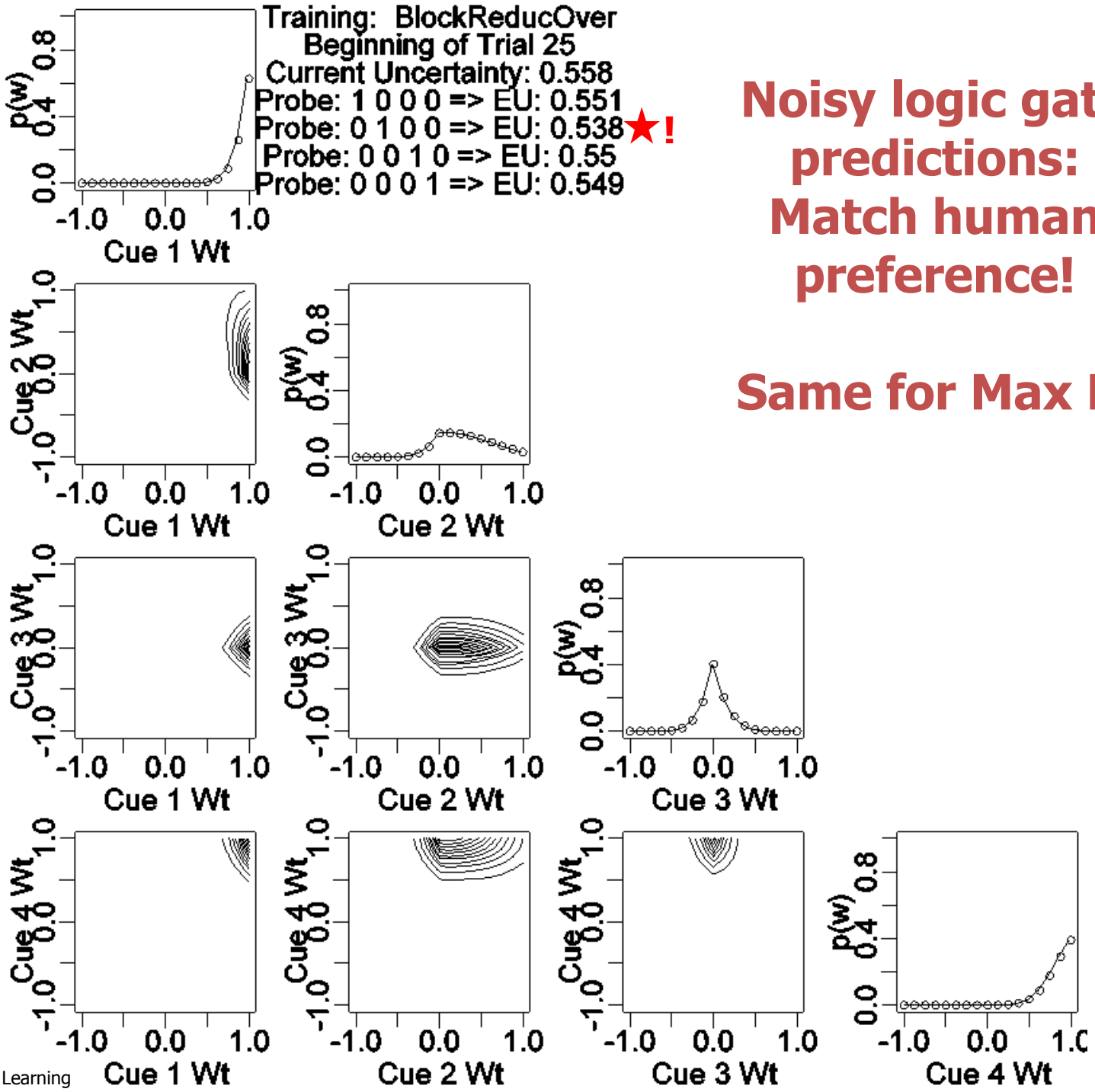

Training: BlockReducOver
Beginning of Trial 20
Current Uncertainty: 0.628
Probe: 1 0 0 0 => EU: 0.621
Probe: 0 1 0 0 => EU: 0.607
Probe: 0 0 1 0 => EU: 0.621
Probe: 0 0 0 1 => EU: 0.616
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8


Training: BlockReducOver
Beginning of Trial 21
Current Uncertainty: 0.607
Probe: 1 0 0 0 => EU: 0.6
Probe: 0 1 0 0 => EU: 0.587
Probe: 0 0 1 0 => EU: 0.6
Probe: 0 0 0 1 => EU: 0.596
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.4
0.8


Training: BlockReducOver
Beginning of Trial 22
Current Uncertainty: 0.6
Probe: 1 0 0 0 => EU: 0.593
Probe: 0 1 0 0 => EU: 0.579
Probe: 0 0 1 0 => EU: 0.592
Probe: 0 0 0 1 => EU: 0.589
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8


Training: BlockReducOver
Beginning of Trial 23
Current Uncertainty: 0.581
Probe: 1 0 0 0 => EU: 0.574
Probe: 0 1 0 0 => EU: 0.561
Probe: 0 0 1 0 => EU: 0.574
Probe: 0 0 0 1 => EU: 0.571
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8


Training: BlockReducOver
Beginning of Trial 24
Current Uncertainty: 0.575
Probe: 1 0 0 0 => EU: 0.568
Probe: 0 1 0 0 => EU: 0.555
Probe: 0 0 1 0 => EU: 0.567
Probe: 0 0 0 1 => EU: 0.565
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
Cue 4 Wt
-1.0
0.0
1.0
0.0
0.4
0.8



Noisy logic gate predictions: Match human preference!

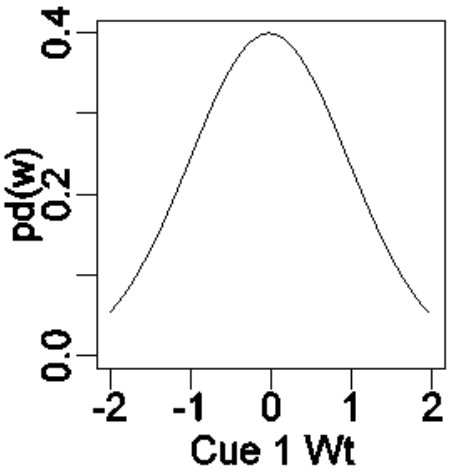
Same for Max P.

Ambiguous Cue

Remember: Foods can be anti-emetic, i.e., prevent nausea.

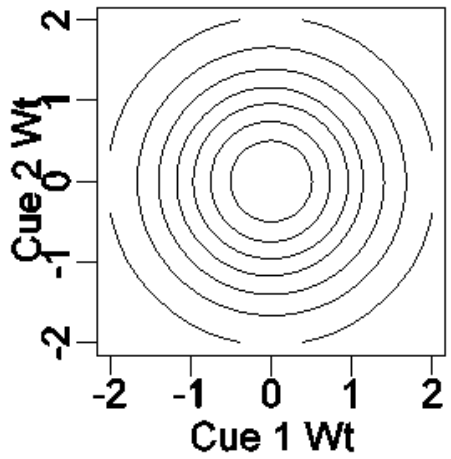
Freq.	Cue 1	Cue 2	Cue 3	Outcome
10	1	1	0	1
10	0	1	1	0

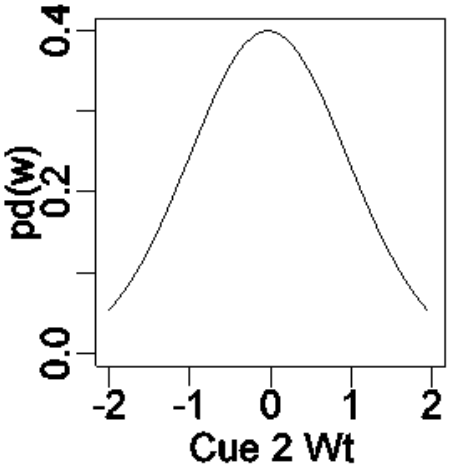
After being trained, your goal is to better determine which cues produce *or prevent* nausea. Which probe below do you prefer to test?

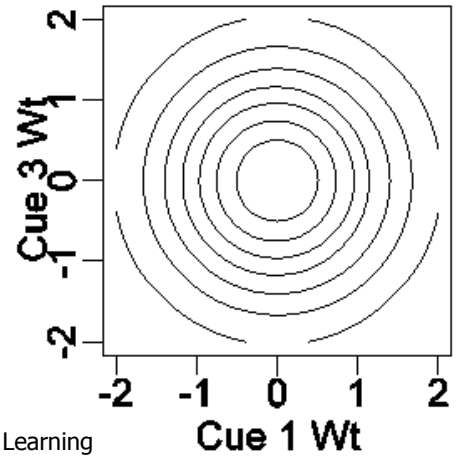


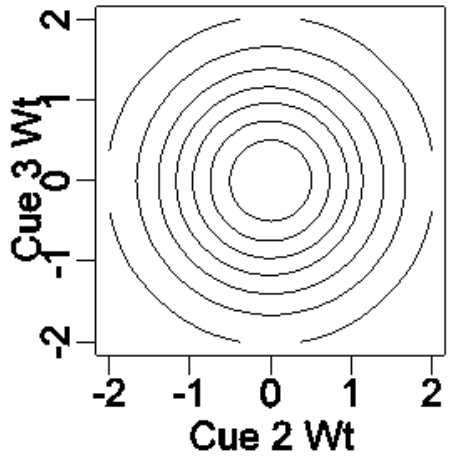
Train: AmbigCue
Beginning of Trial 1
Mean: 0 0 0
Covariance matrix:
1 0 0
0 1 0
0 0 1
Current Uncertainty: 4.257
Probe: 1 0 0 => EU: 4.145
Probe: 0 1 0 => EU: 4.145
Probe: 0 0 1 => EU: 4.145

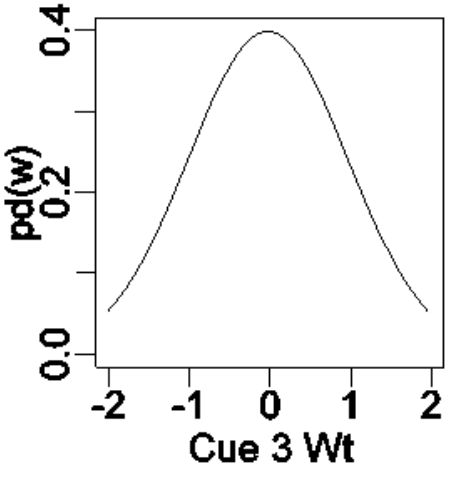
Kalman filter predictions for Expected Uncertainty…

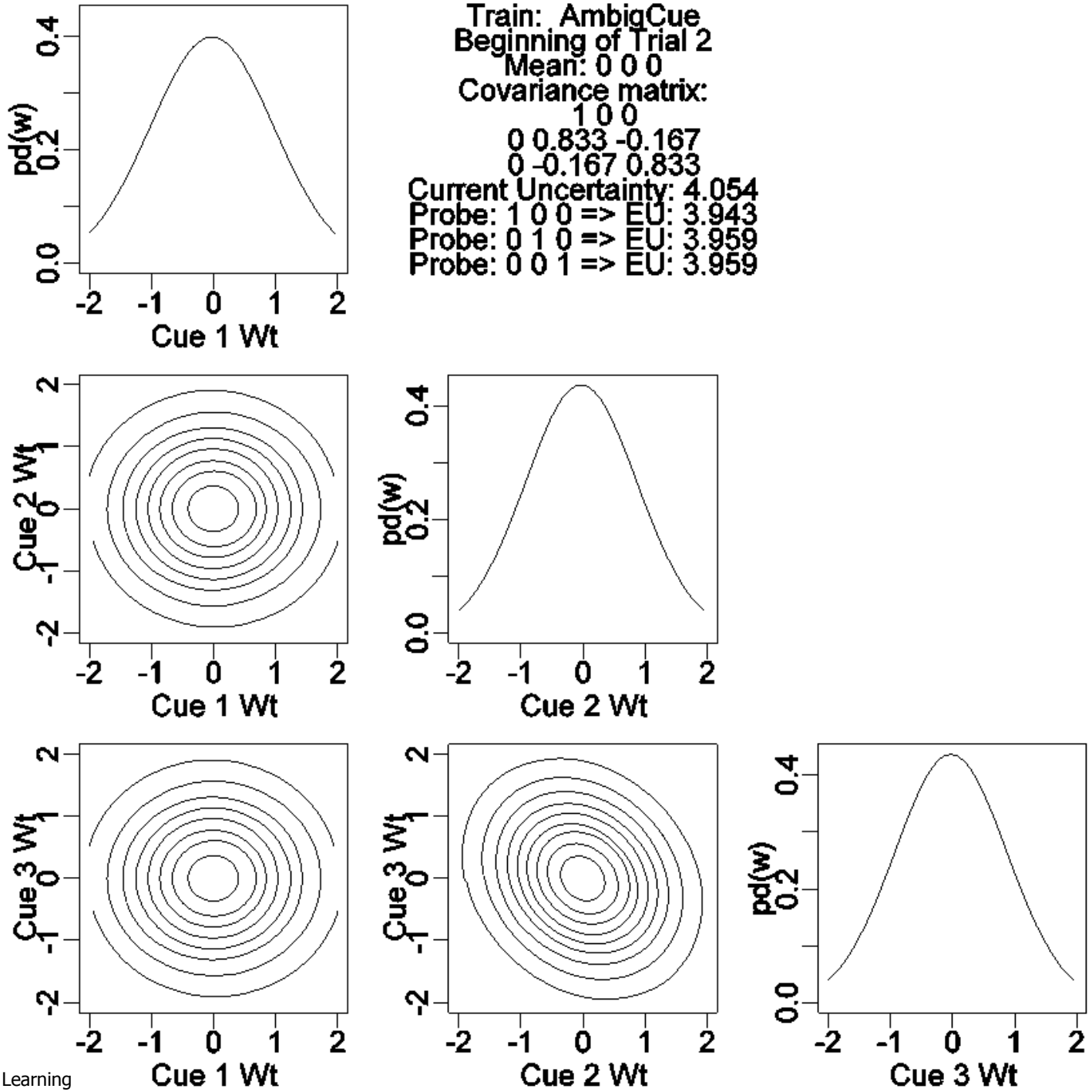


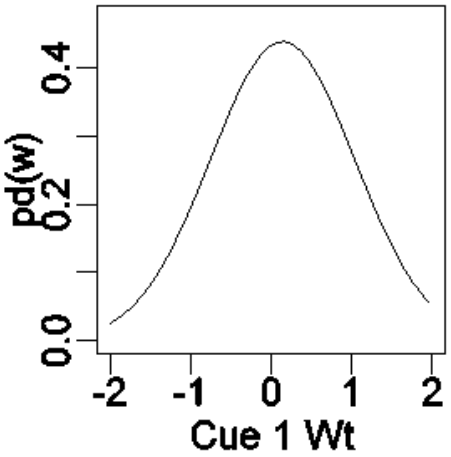




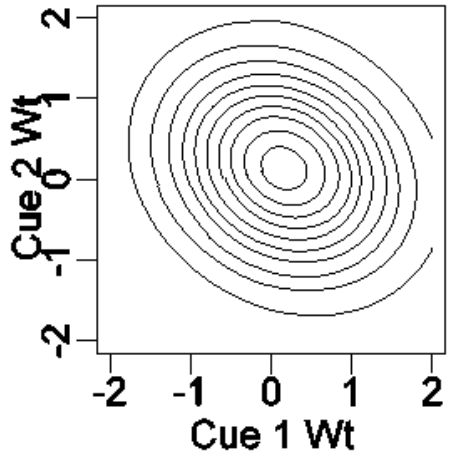


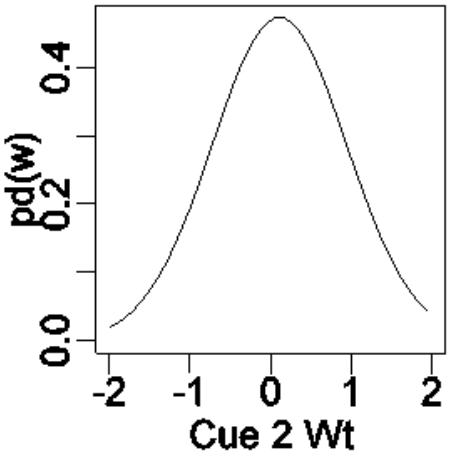


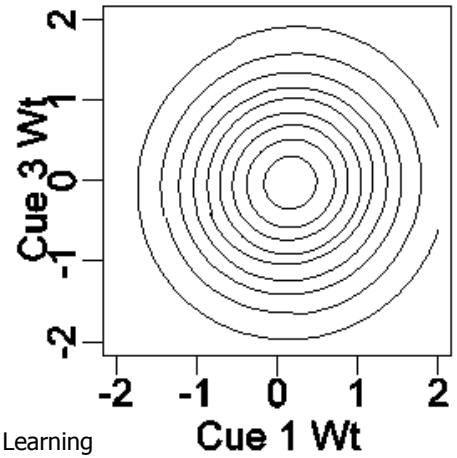

Train: AmbigCue
Beginning of Trial 2
Mean: 0 0 0
Covariance matrix:
1 0 0
0 0.833 -0.167
0 -0.167 0.833
Current Uncertainty: 4.054
Probe: 1 0 0 => EU: 3.943
Probe: 0 1 0 => EU: 3.959
Probe: 0 0 1 => EU: 3.959
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

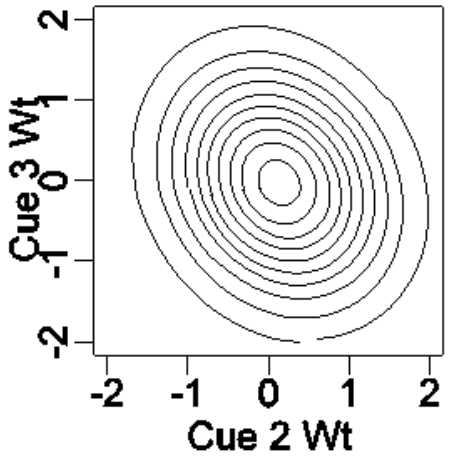


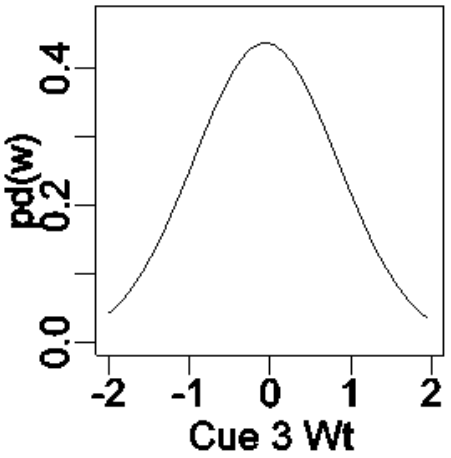
Train: AmbigCue
Beginning of Trial 3
Mean: 0.171 0.143 -0.029
Covariance matrix:
0.829 -0.143 0.0286
-0.143 0.714 -0.143
0.0286 -0.143 0.829
Current Uncertainty: 3.865
Probe: 1 0 0 => EU: 3.771
Probe: 0 1 0 => EU: 3.783
Probe: 0 0 1 => EU: 3.771

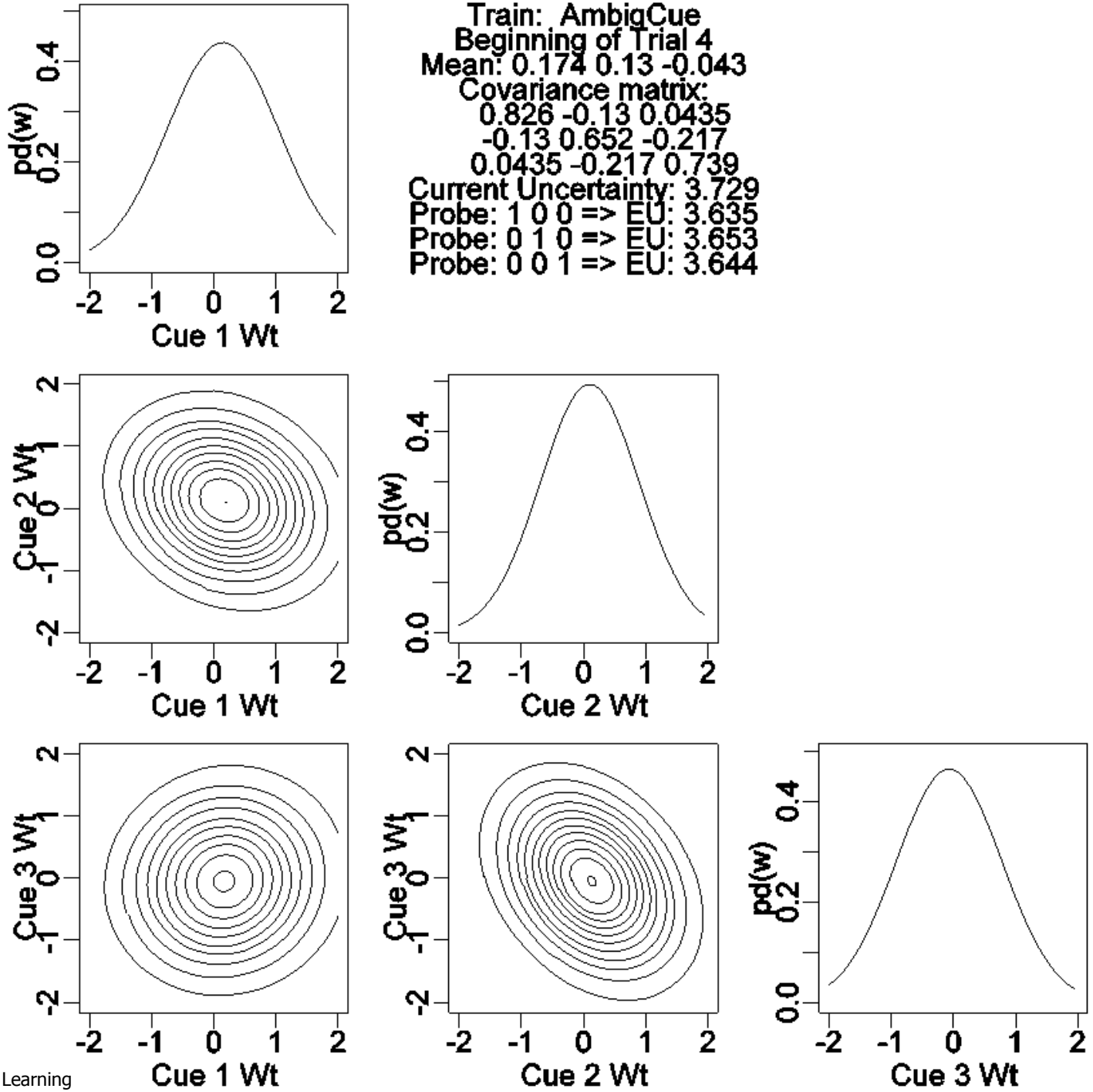


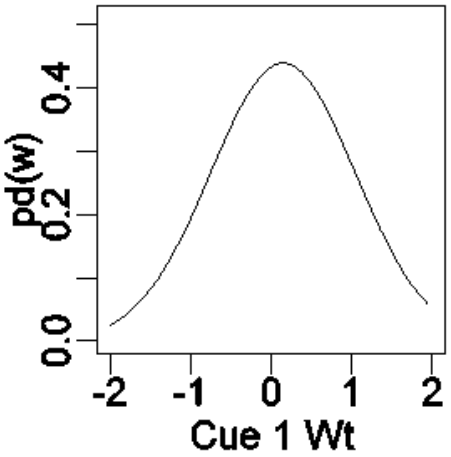




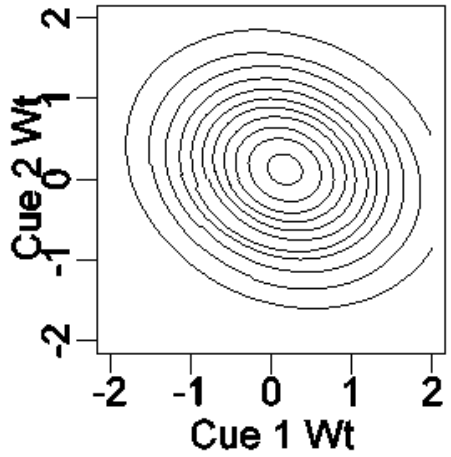


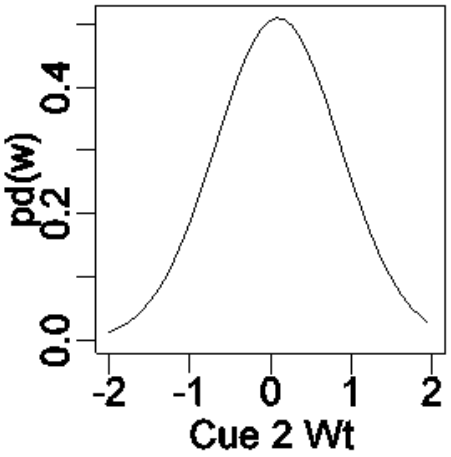


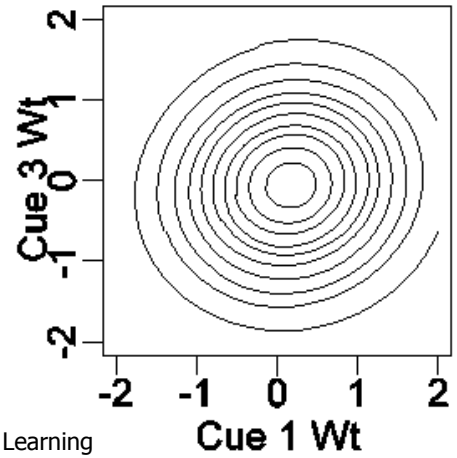

Train: AmbigCue
Beginning of Trial 4
Mean: 0.174 0.13 -0.043
Covariance matrix:
0.826 -0.13 0.0435
-0.13 0.652 -0.217
0.0435 -0.217 0.739
Current Uncertainty: 3.729
Probe: 1 0 0 => EU: 3.635
Probe: 0 1 0 => EU: 3.653
Probe: 0 0 1 => EU: 3.644
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

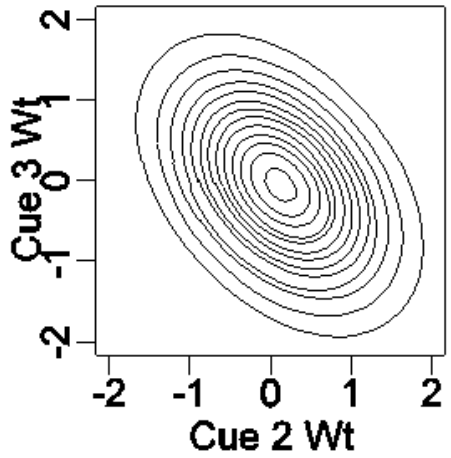


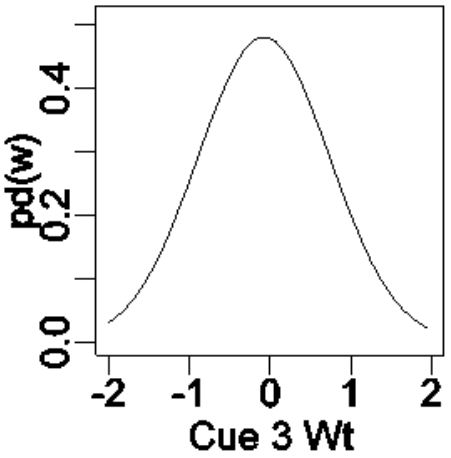
Train: AmbigCue
Beginning of Trial 5
Mean: 0.175 0.123 -0.053
Covariance matrix:
0.825 -0.123 0.0526
-0.123 0.614 -0.263
0.0526 -0.263 0.684
Current Uncertainty: 3.622
Probe: 1 0 0 => EU: 3.528
Probe: 0 1 0 => EU: 3.55
Probe: 0 0 1 => EU: 3.543

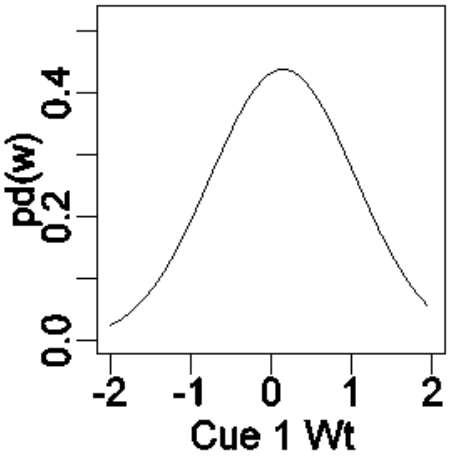




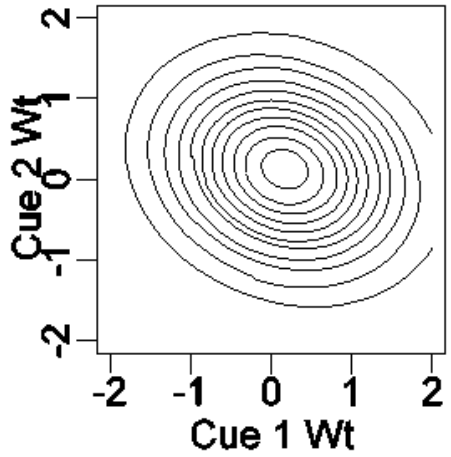


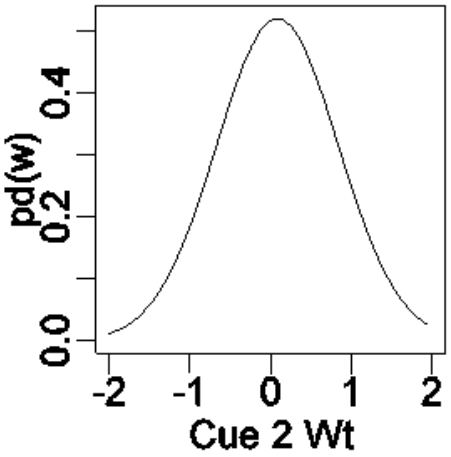


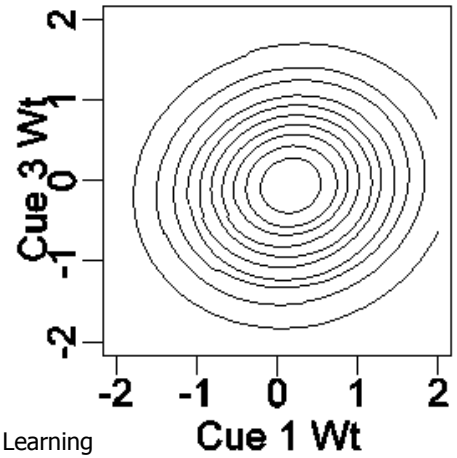


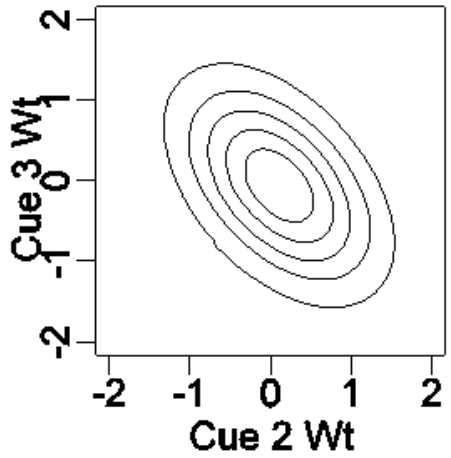


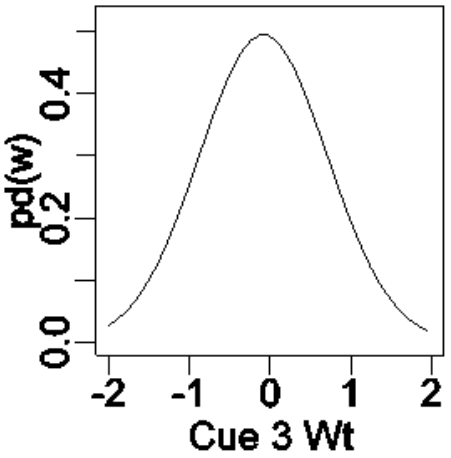
Train: AmbigCue
Beginning of Trial 6
Mean: 0.176 0.118 -0.059
Covariance matrix:
0.824 -0.118 0.0588
-0.118 0.588 -0.294
0.0588 -0.294 0.647
Current Uncertainty: 3.533
Probe: 1 0 0 => EU: 3.44
Probe: 0 1 0 => EU: 3.465
Probe: 0 0 1 => EU: 3.458

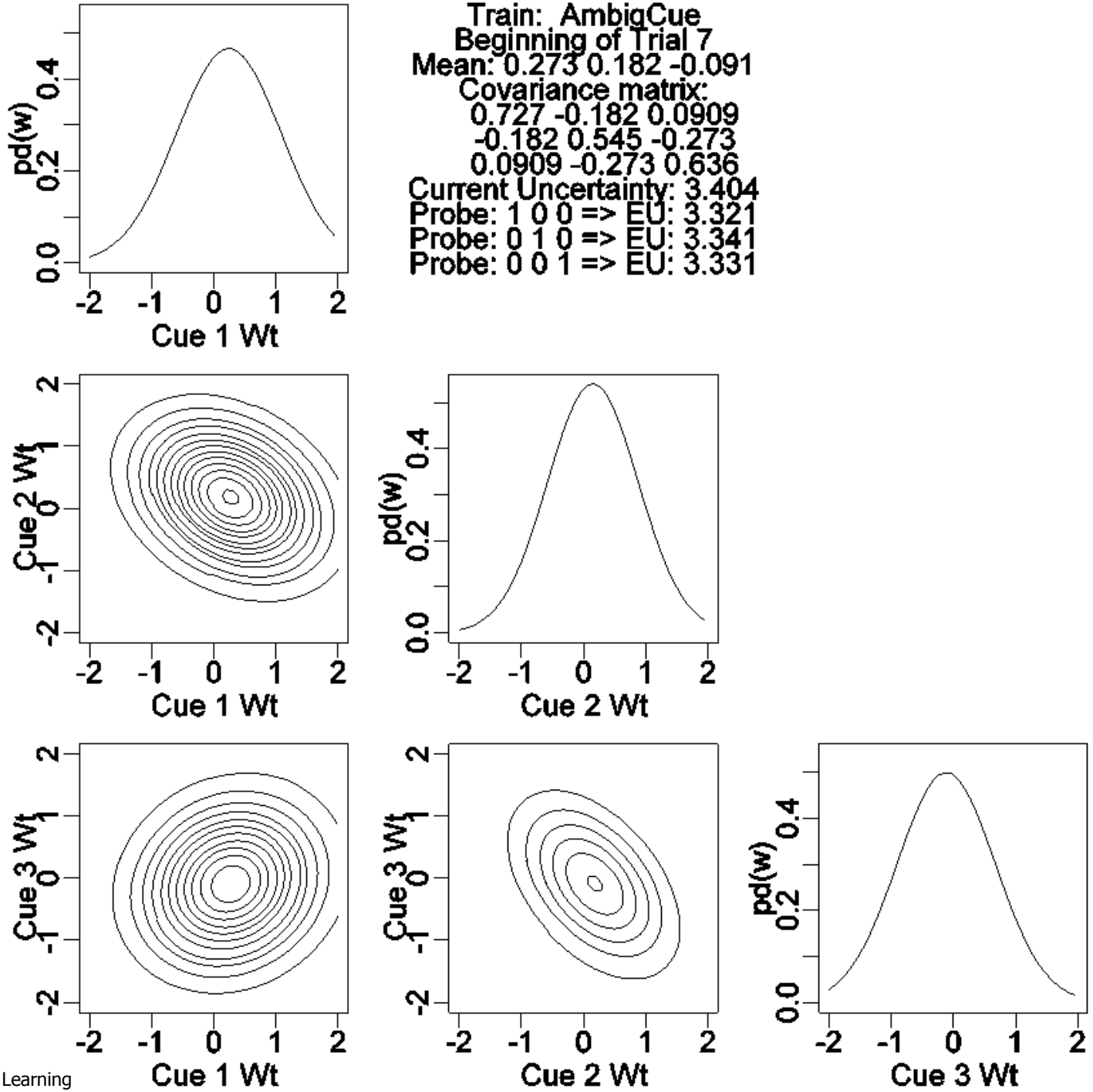


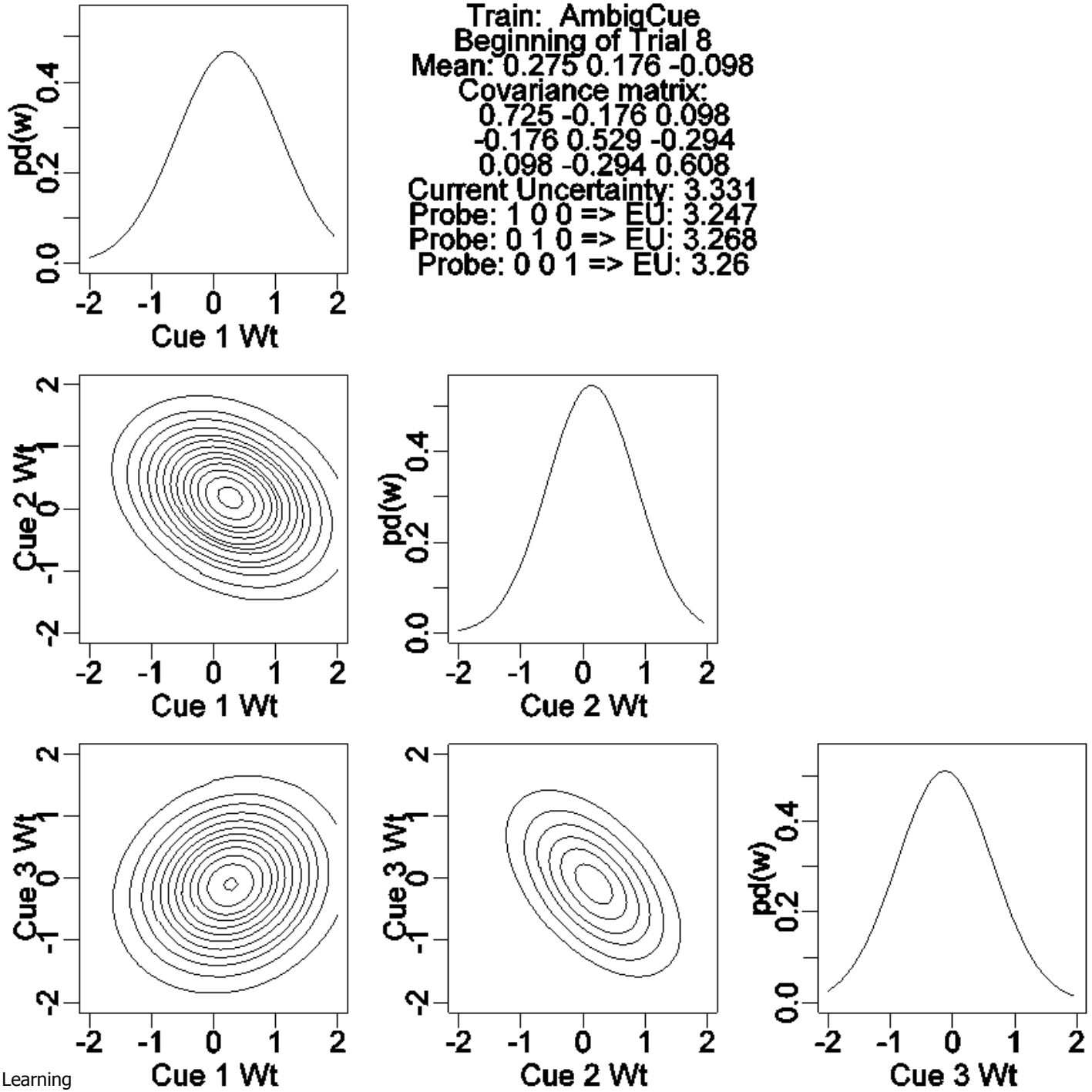


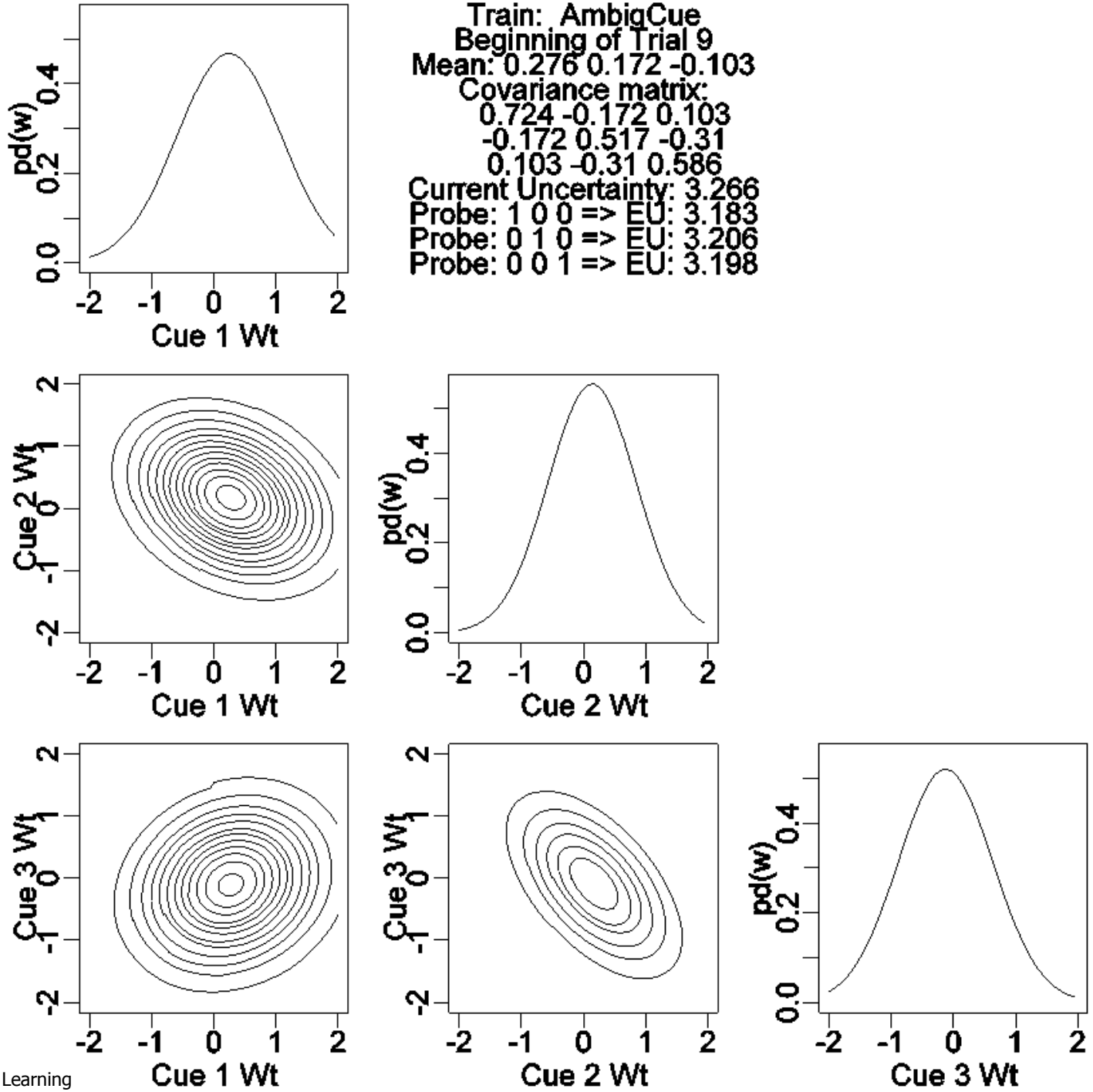


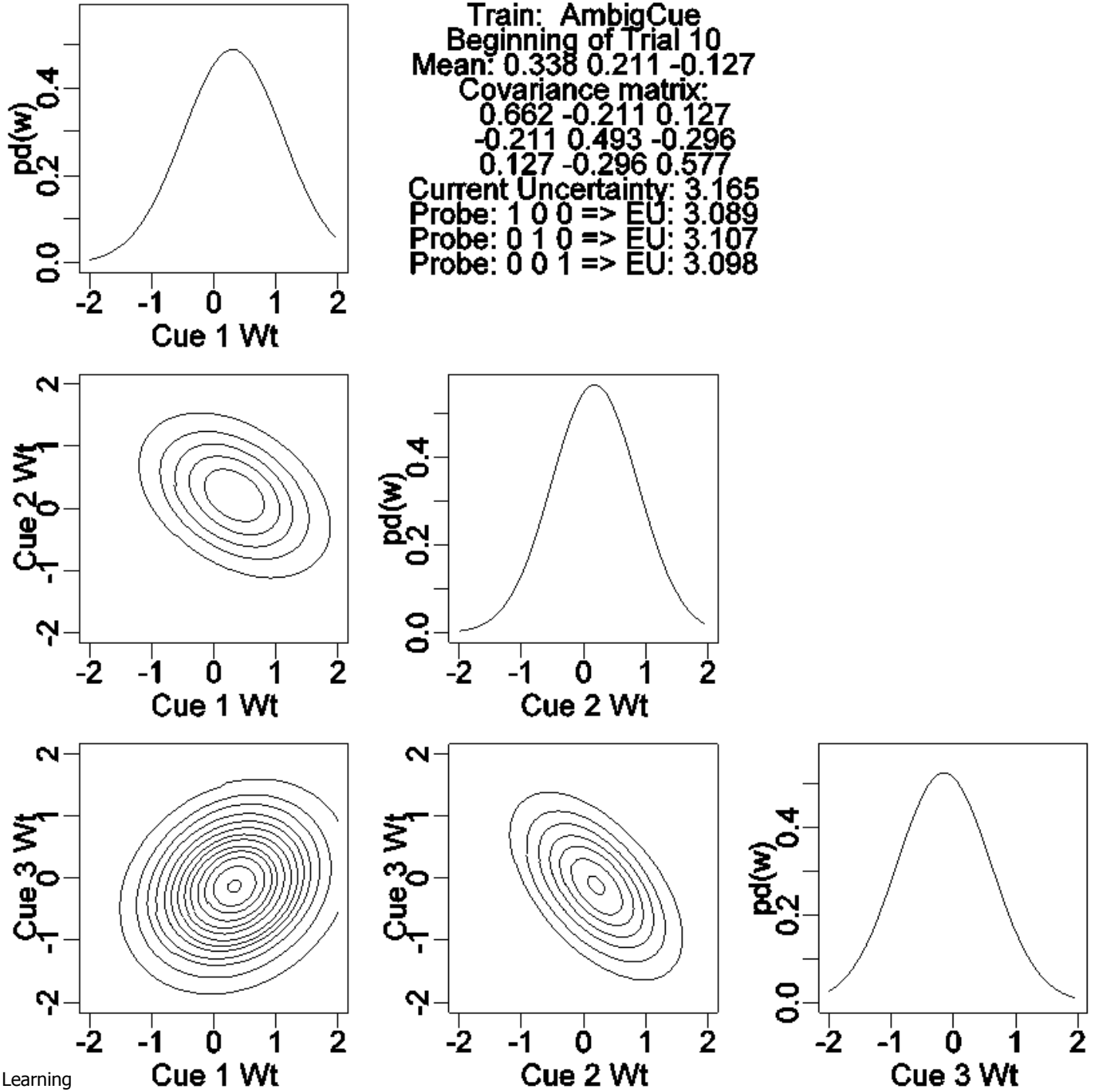


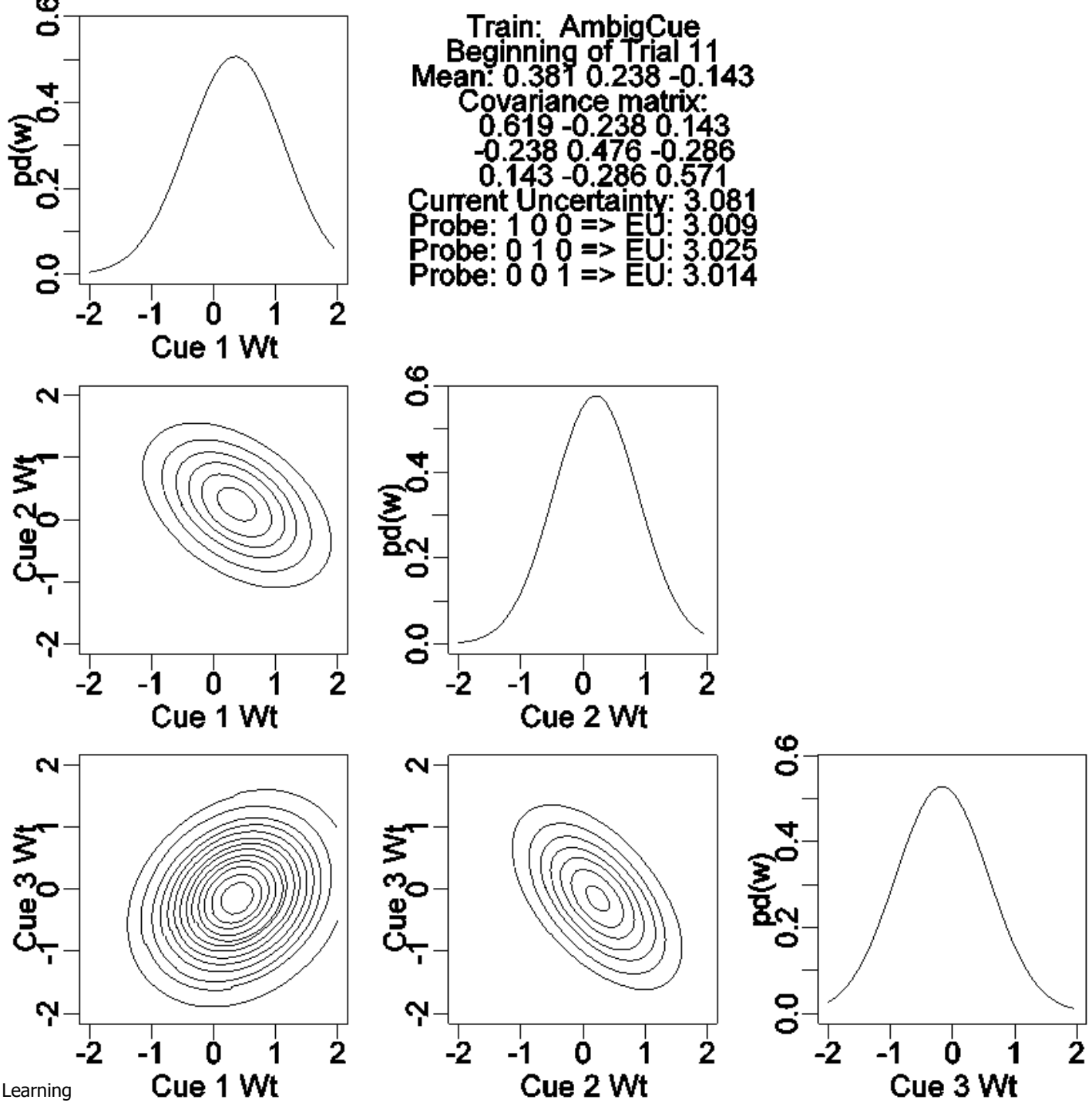


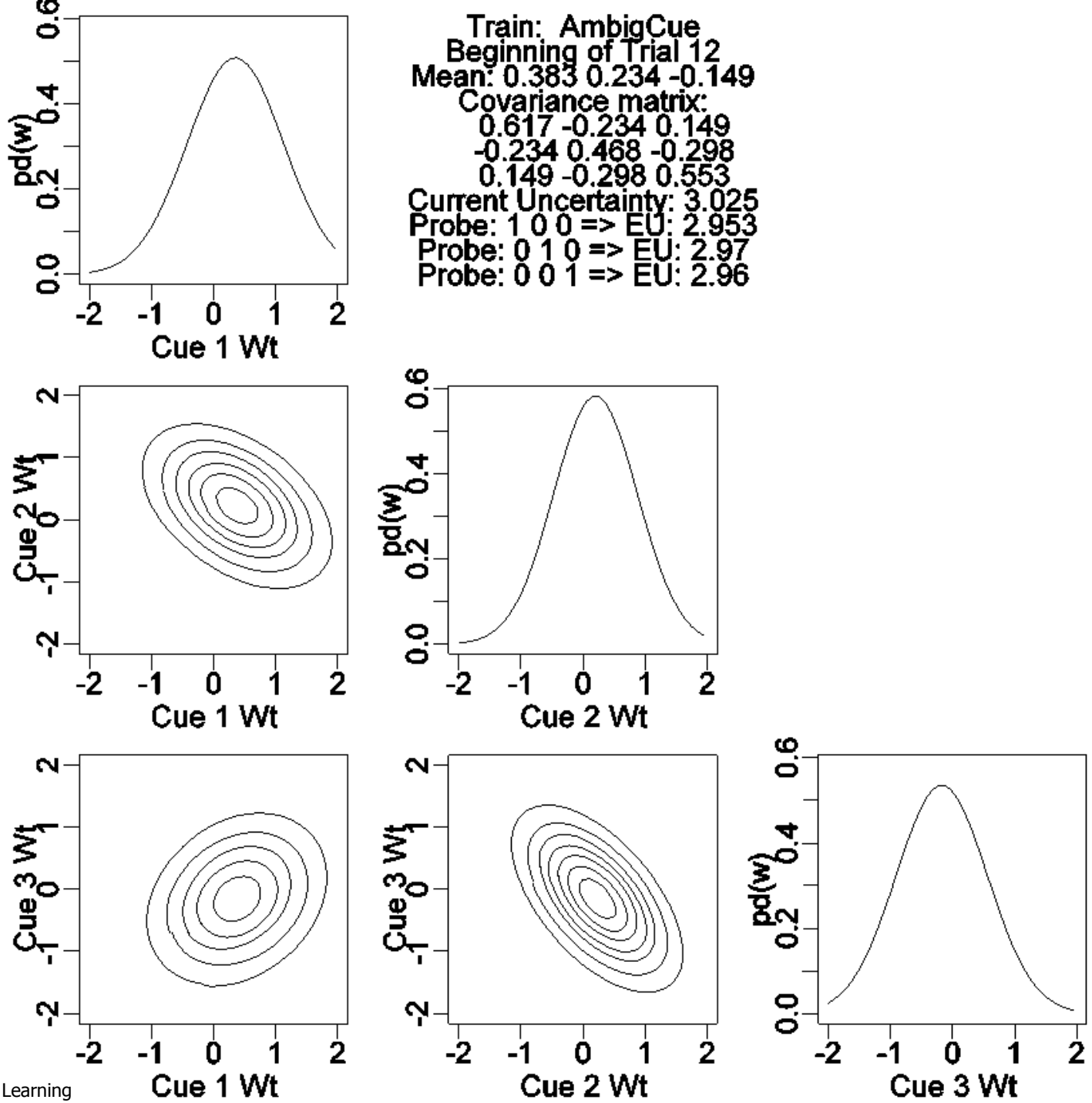

Train: AmbigCue
Beginning of Trial 7
Mean: 0.273 0.182 -0.091
Covariance matrix:
0.727 -0.182 0.0909
-0.182 0.545 -0.273
0.0909 -0.273 0.636
Current Uncertainty: 3.404
Probe: 1 0 0 => EU: 3.321
Probe: 0 1 0 => EU: 3.341
Probe: 0 0 1 => EU: 3.331
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

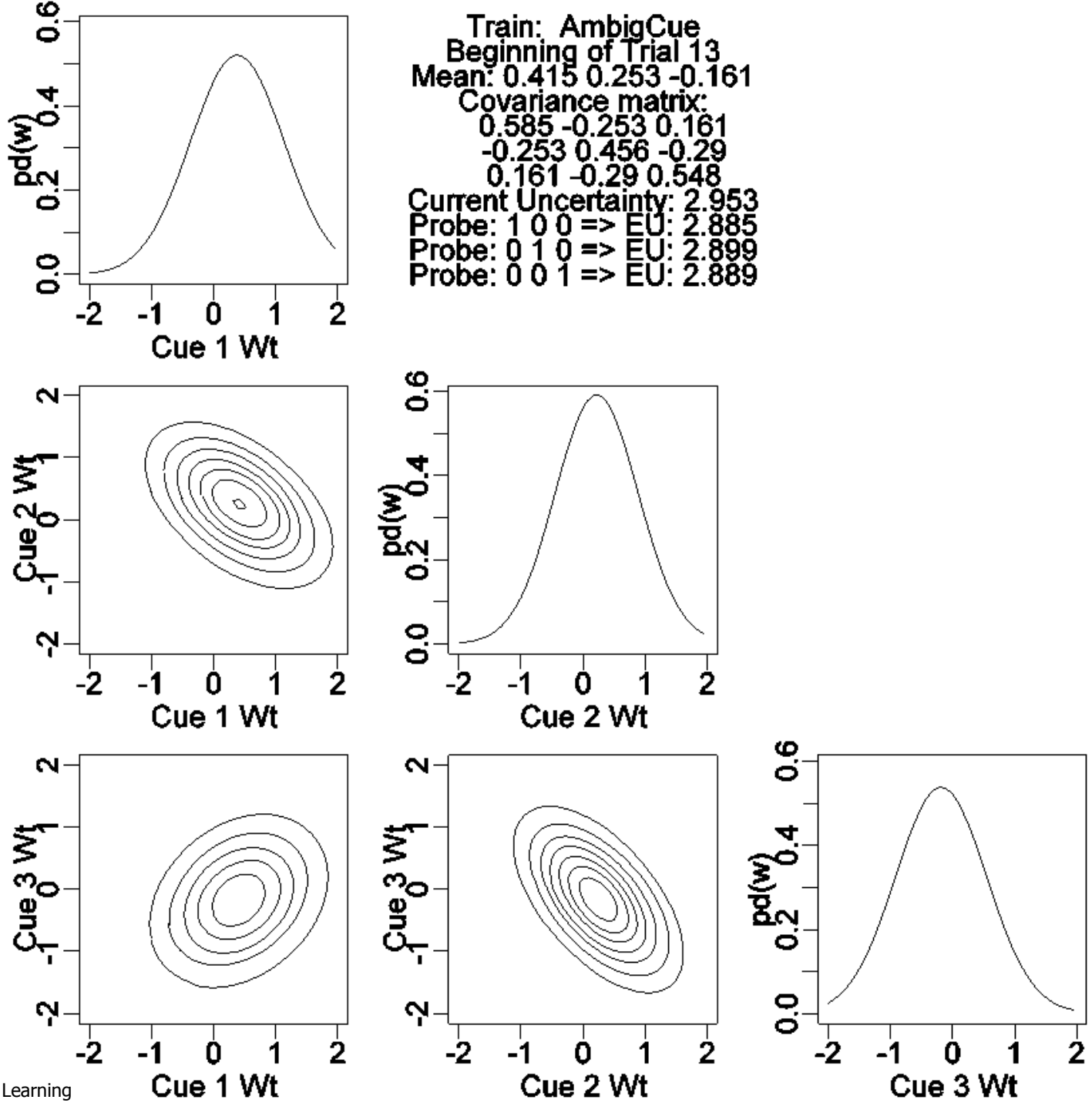

Train: AmbigCue
Beginning of Trial 8
Mean: 0.275 0.176 -0.098
Covariance matrix:
0.725 -0.176 0.098
-0.176 0.529 -0.294
0.098 -0.294 0.608
Current Uncertainty: 3.331
Probe: 1 0 0 => EU: 3.247
Probe: 0 1 0 => EU: 3.268
Probe: 0 0 1 => EU: 3.26
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

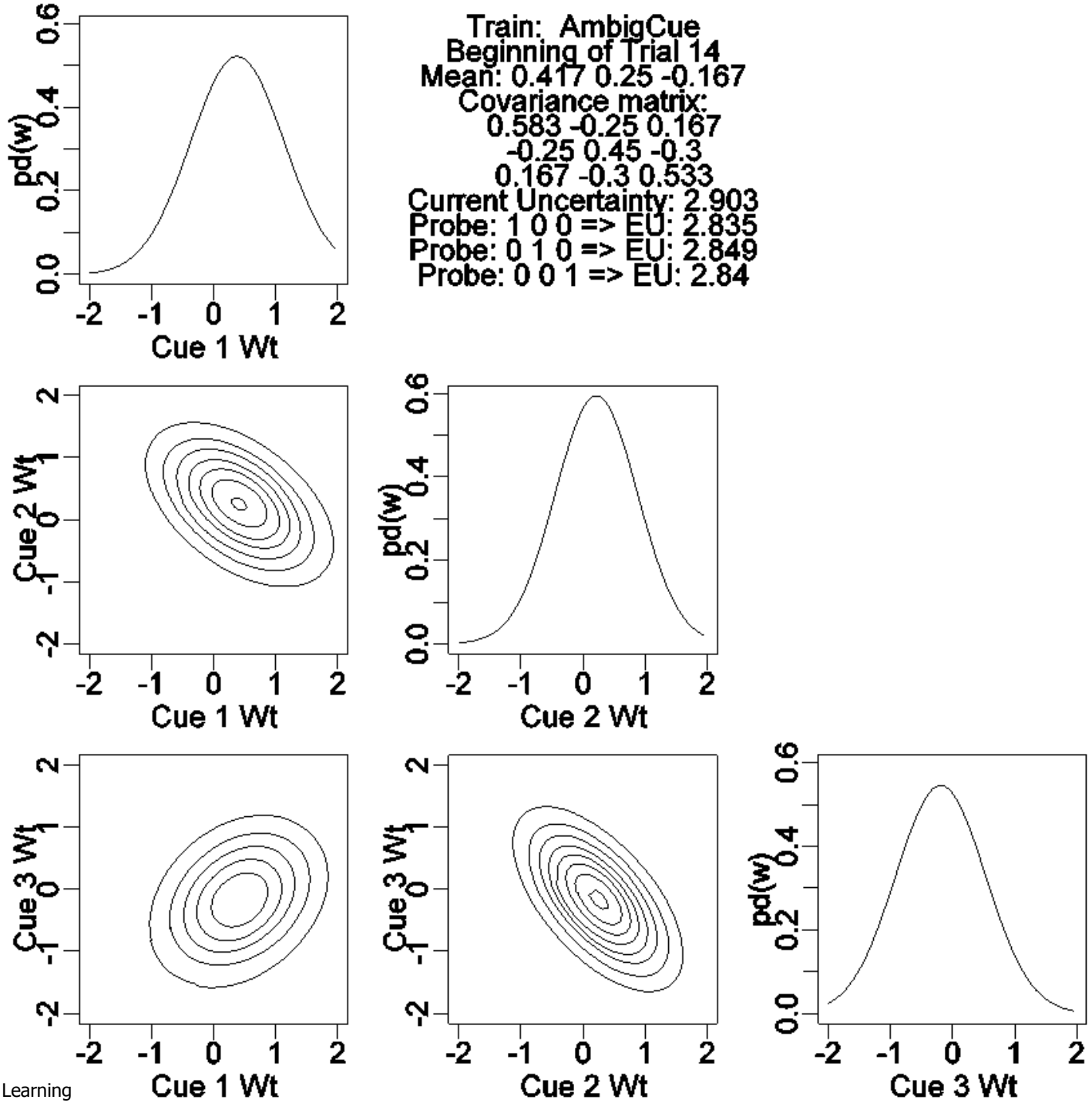

Train: AmbigCue
Beginning of Trial 9
Mean: 0.276 0.172 -0.103
Covariance matrix:
0.724 -0.172 0.103
-0.172 0.517 -0.31
0.103 -0.31 0.586
Current Uncertainty: 3.266
Probe: 1 0 0 => EU: 3.183
Probe: 0 1 0 => EU: 3.206
Probe: 0 0 1 => EU: 3.198
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

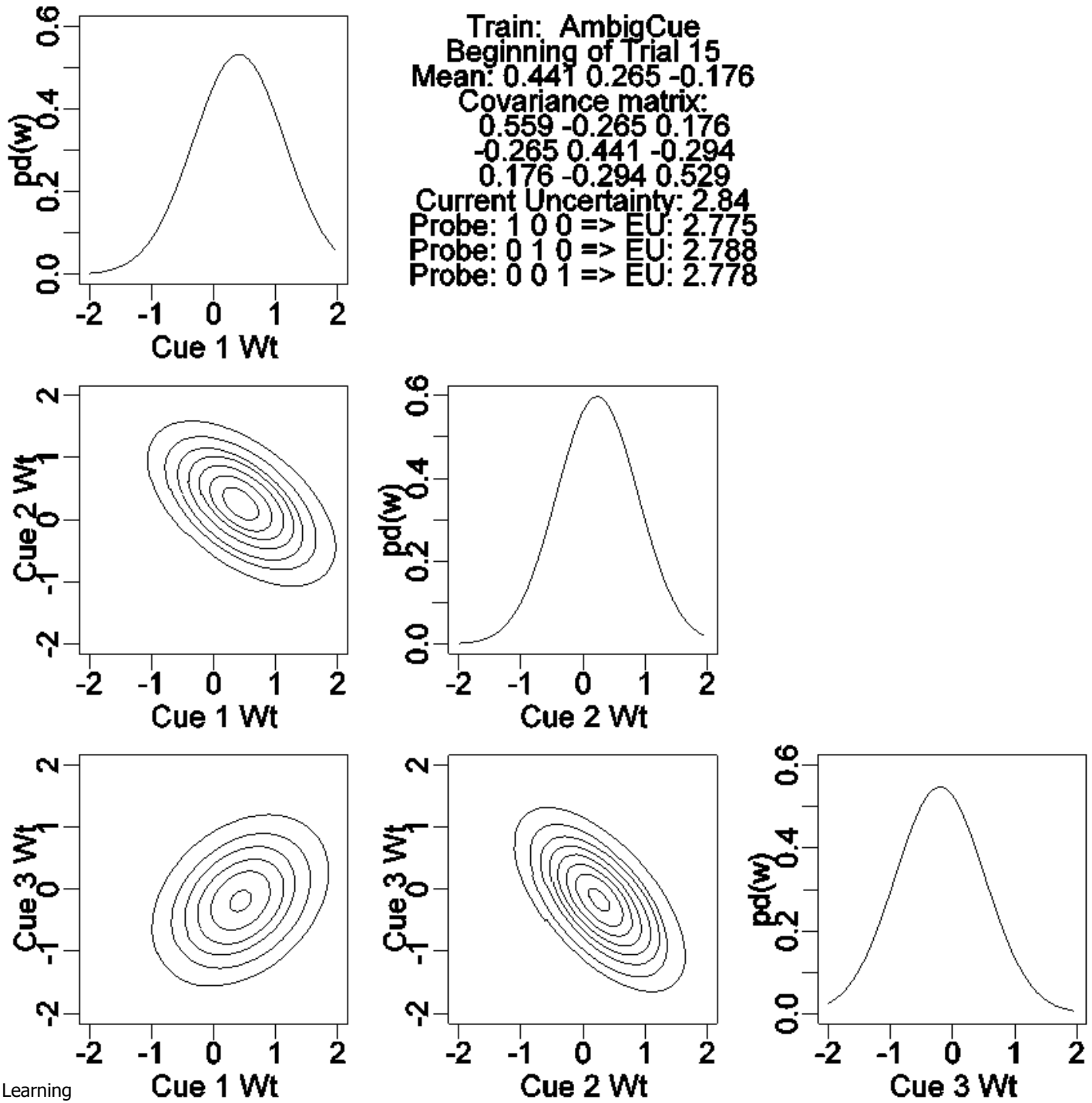

Train: AmbigCue
Beginning of Trial 10
Mean: 0.338 0.211 -0.127
Covariance matrix:
0.662 -0.211 0.127
-0.211 0.493 -0.296
0.127 -0.296 0.577
Current Uncertainty: 3.165
Probe: 1 0 0 => EU: 3.089
Probe: 0 1 0 => EU: 3.107
Probe: 0 0 1 => EU: 3.098
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

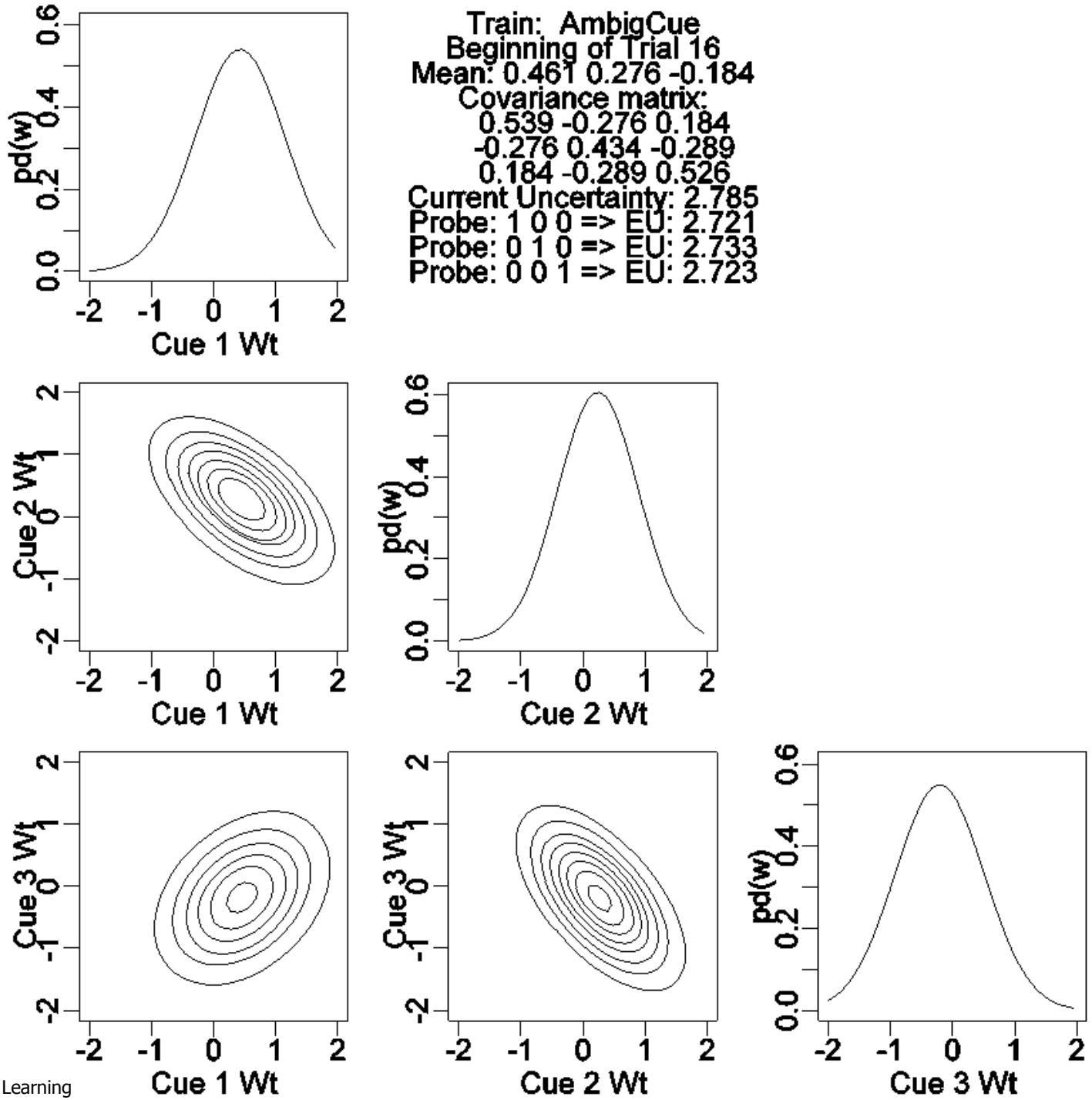

Train: AmbigCue
Beginning of Trial 11
Mean: 0.381 0.238 -0.143
Covariance matrix:
0.619 -0.238 0.143
-0.238 0.476 -0.286
0.143 -0.286 0.571
Current Uncertainty: 3.081
Probe: 1 0 0 => EU: 3.009
Probe: 0 1 0 => EU: 3.025
Probe: 0 0 1 => EU: 3.014
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

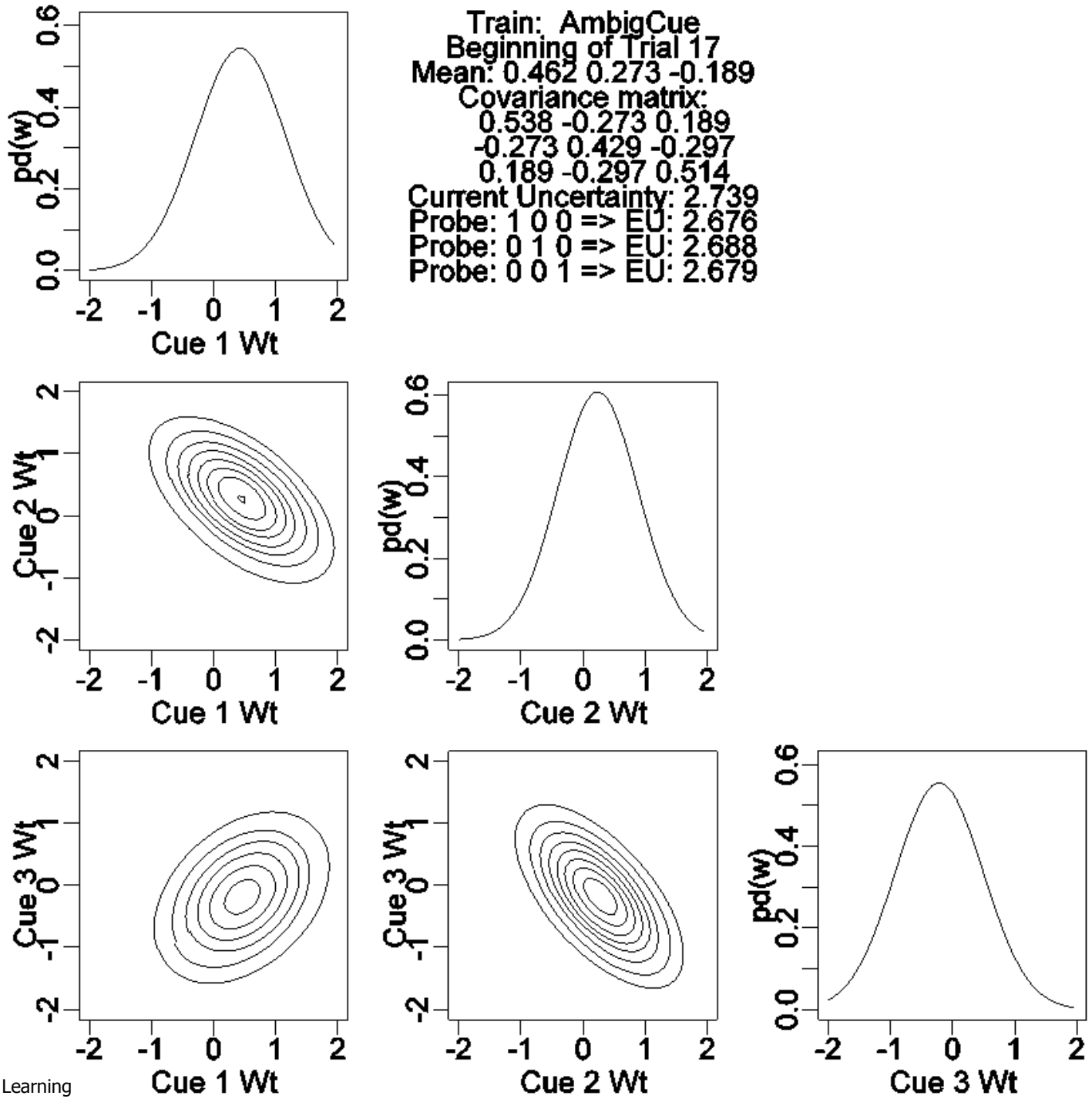

Train: AmbigCue
Beginning of Trial 12
Mean: 0.383 0.234 -0.149
Covariance matrix:
0.617 -0.234 0.149
-0.234 0.468 -0.298
0.149 -0.298 0.553
Current Uncertainty: 3.025
Probe: 1 0 0 => EU: 2.953
Probe: 0 1 0 => EU: 2.97
Probe: 0 0 1 => EU: 2.96
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

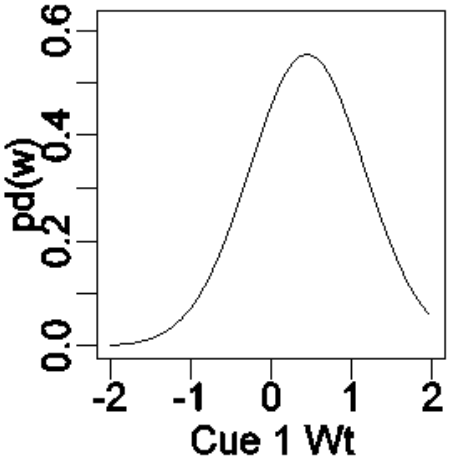

Train: AmbigCue
Beginning of Trial 13
Mean: 0.415 0.253 -0.161
Covariance matrix:
0.585 -0.253 0.161
-0.253 0.456 -0.29
0.161 -0.29 0.548
Current Uncertainty: 2.953
Probe: 1 0 0 => EU: 2.885
Probe: 0 1 0 => EU: 2.899
Probe: 0 0 1 => EU: 2.889
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt


Train: AmbigCue
Beginning of Trial 14
Mean: 0.417 0.25 -0.167
Covariance matrix:
0.583 -0.25 0.167
-0.25 0.45 -0.3
0.167 -0.3 0.533
Current Uncertainty: 2.903
Probe: 1 0 0 => EU: 2.835
Probe: 0 1 0 => EU: 2.849
Probe: 0 0 1 => EU: 2.84
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

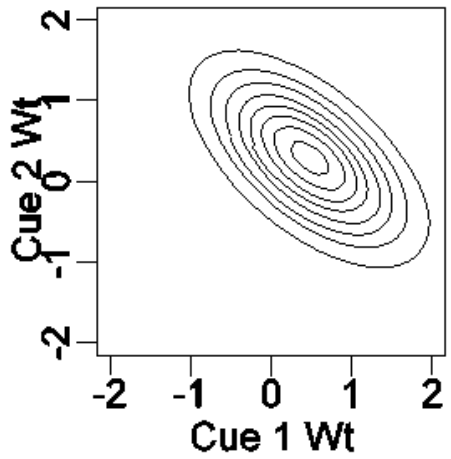

Train: AmbigCue
Beginning of Trial 15
Mean: 0.441 0.265 -0.176
Covariance matrix:
0.559 -0.265 0.176
-0.265 0.441 -0.294
0.176 -0.294 0.529
Current Uncertainty: 2.84
Probe: 1 0 0 => EU: 2.775
Probe: 0 1 0 => EU: 2.788
Probe: 0 0 1 => EU: 2.778
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

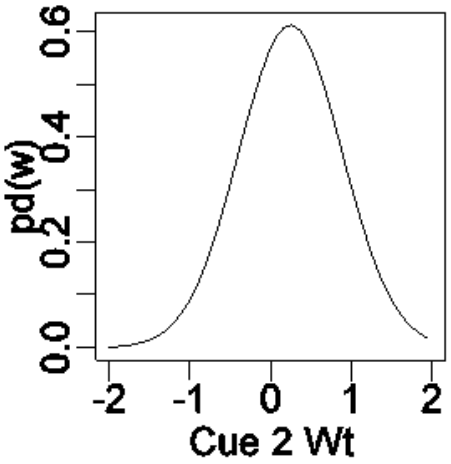

Train: AmbigCue
Beginning of Trial 16
Mean: 0.461 0.276 -0.184
Covariance matrix:
0.539 -0.276 0.184
-0.276 0.434 -0.289
0.184 -0.289 0.526
Current Uncertainty: 2.785
Probe: 1 0 0 => EU: 2.721
Probe: 0 1 0 => EU: 2.733
Probe: 0 0 1 => EU: 2.723
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

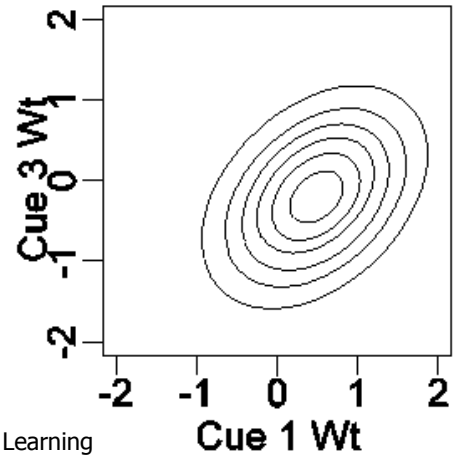

Train: AmbigCue
Beginning of Trial 17
Mean: 0.462 0.273 -0.189
Covariance matrix:
0.538 -0.273 0.189
-0.273 0.429 -0.297
0.189 -0.297 0.514
Current Uncertainty: 2.739
Probe: 1 0 0 => EU: 2.676
Probe: 0 1 0 => EU: 2.688
Probe: 0 0 1 => EU: 2.679
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

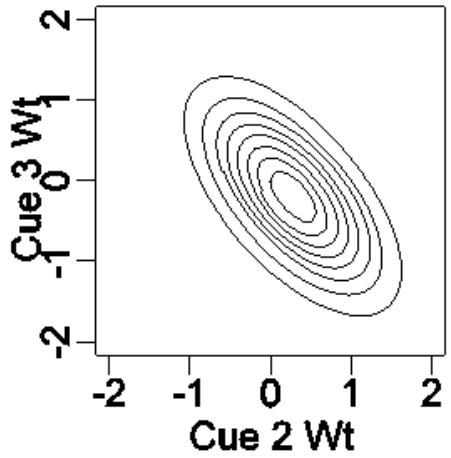


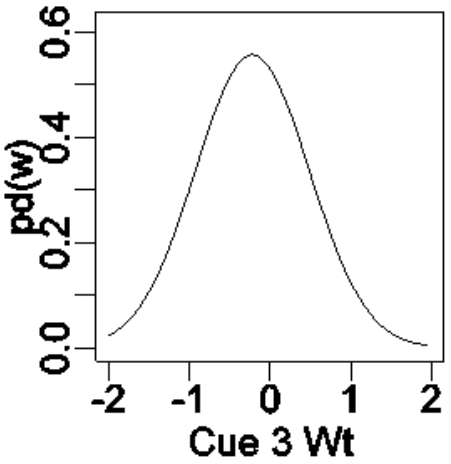
Train: AmbigCue
Beginning of Trial 18
Mean: 0.478 0.283 -0.196
Covariance matrix:
0.522 -0.283 0.196
-0.283 0.424 -0.293
0.196 -0.293 0.511
Current Uncertainty: 2.689
Probe: 1 0 0 => EU: 2.628
Probe: 0 1 0 => EU: 2.639
Probe: 0 0 1 => EU: 2.629

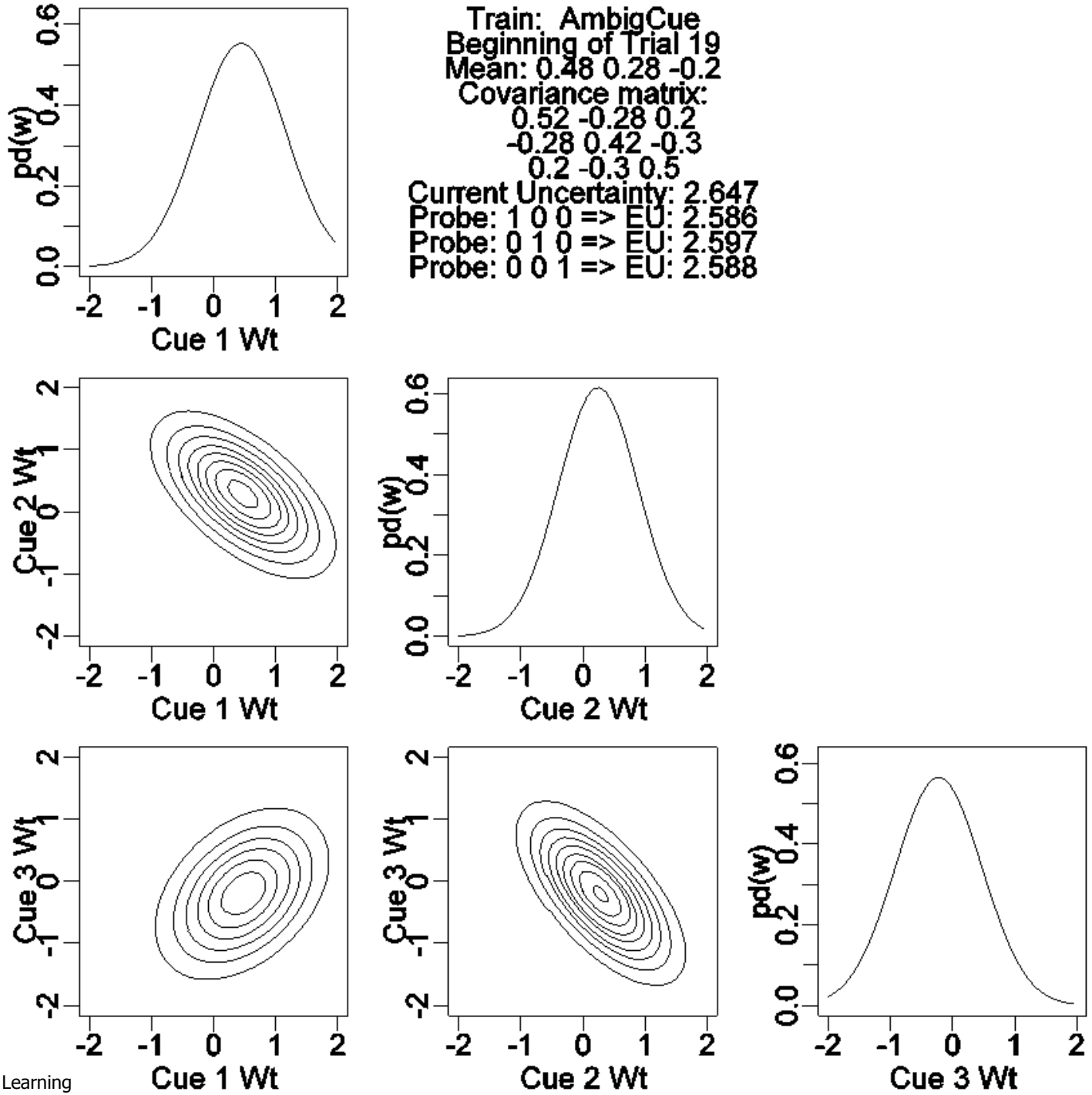


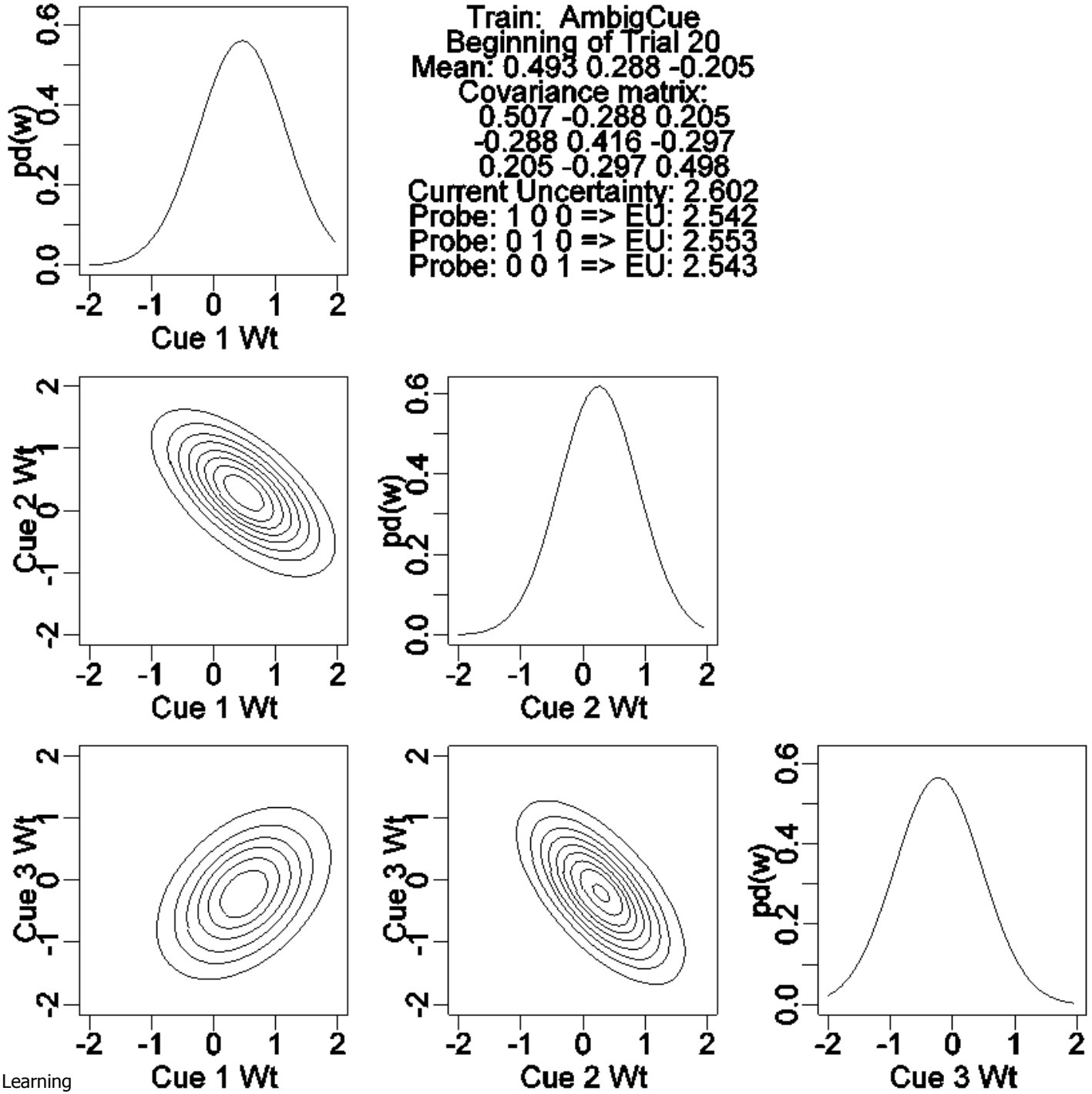


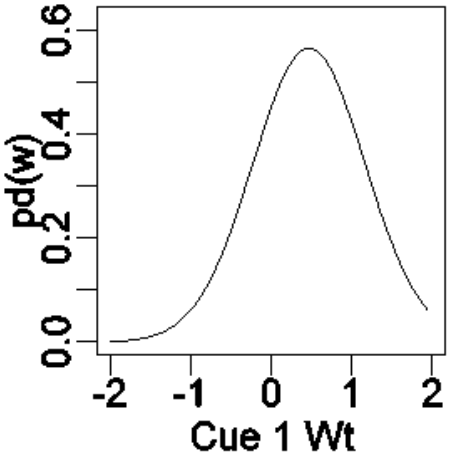






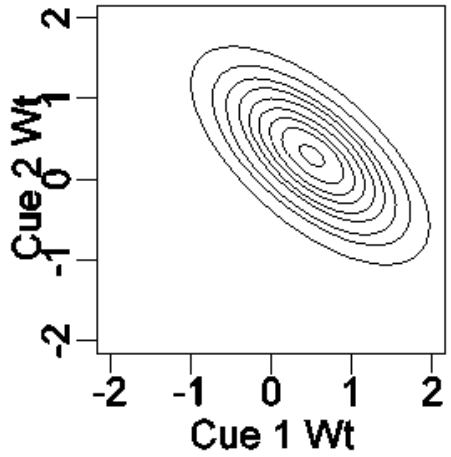

Train: AmbigCue
Beginning of Trial 19
Mean: 0.48 0.28 -0.2
Covariance matrix:
0.52 -0.28 0.2
-0.28 0.42 -0.3
0.2 -0.3 0.5
Current Uncertainty: 2.647
Probe: 1 0 0 => EU: 2.586
Probe: 0 1 0 => EU: 2.597
Probe: 0 0 1 => EU: 2.588
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

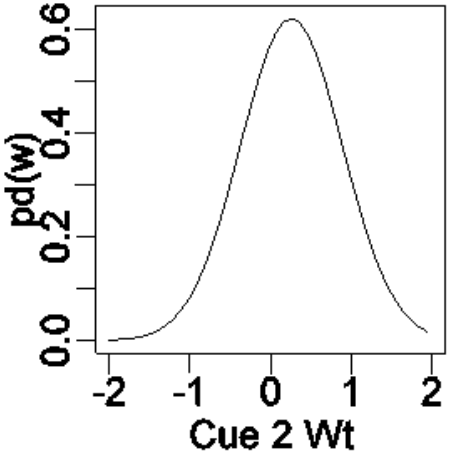

Train: AmbigCue
Beginning of Trial 20
Mean: 0.493 0.288 -0.205
Covariance matrix:
0.507 -0.288 0.205
-0.288 0.416 -0.297
0.205 -0.297 0.498
Current Uncertainty: 2.602
Probe: 1 0 0 => EU: 2.542
Probe: 0 1 0 => EU: 2.553
Probe: 0 0 1 => EU: 2.543
pd(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

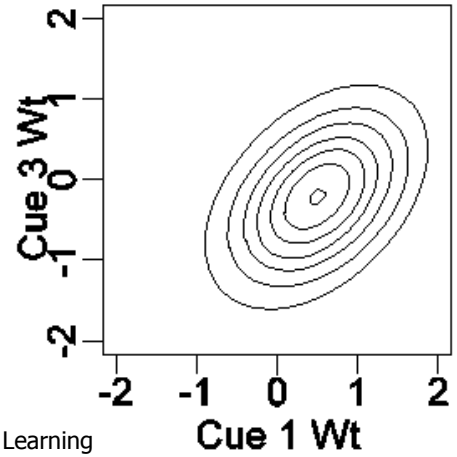


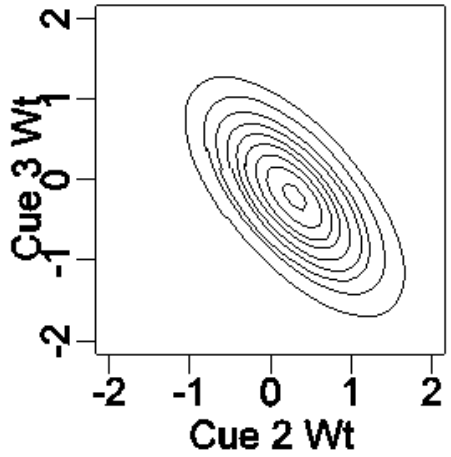
Train: AmbigCue
Beginning of Trial 21
Mean: 0.504 0.294 -0.21
Covariance matrix:
0.496 -0.294 0.21
-0.294 0.412 -0.294
0.21 -0.294 0.496
Current Uncertainty: 2.56
Probe: 1 0 0 => EU: 2.502
Probe: 0 1 0 => EU: 2.511
Probe: 0 0 1 => EU: 2.502

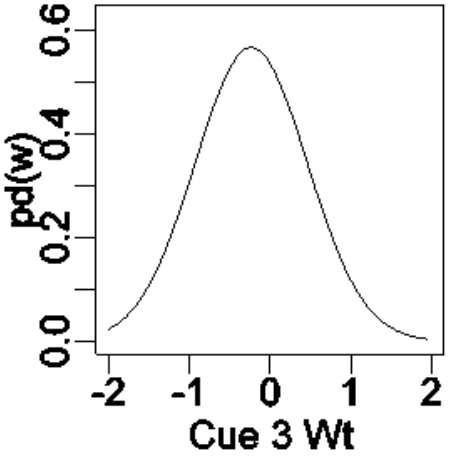
Cue 2 is *least* preferred!

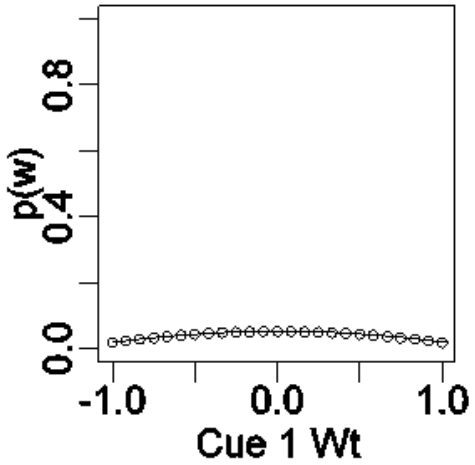












Training: AmbigCue

Beginning of Trial 1

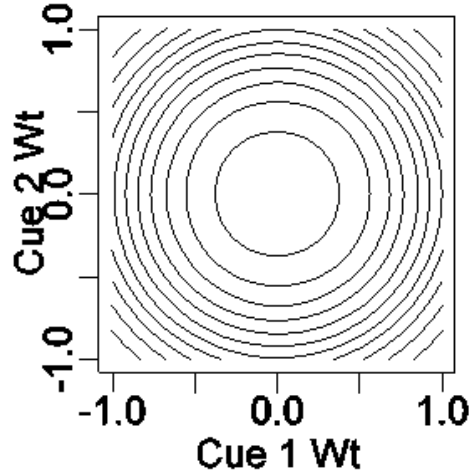
Current Uncertainty: 0.987

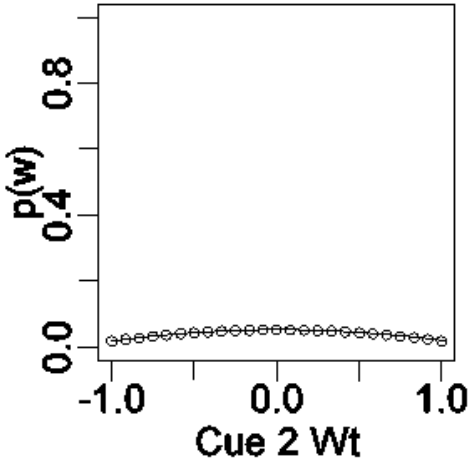
Probe: 1 0 0 => EU: 0.958

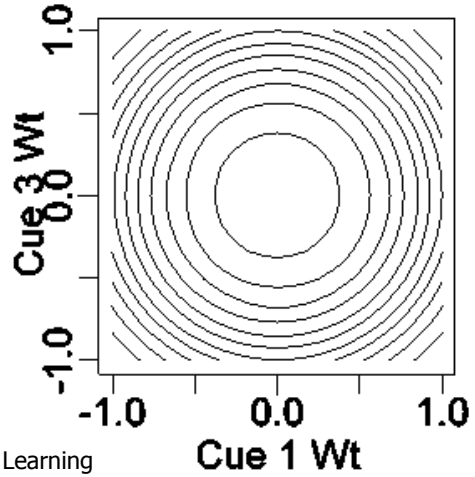
Probe: 0 1 0 => EU: 0.958

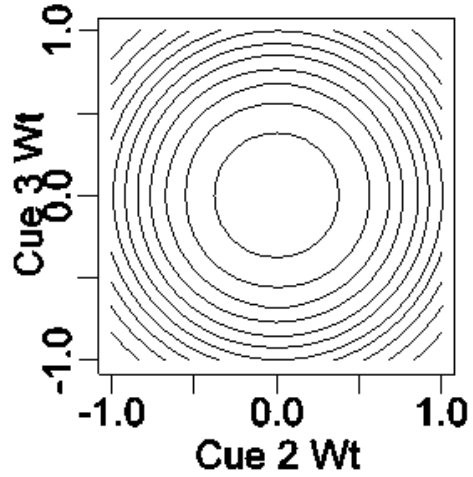
Probe: 0 0 1 => EU: 0.958

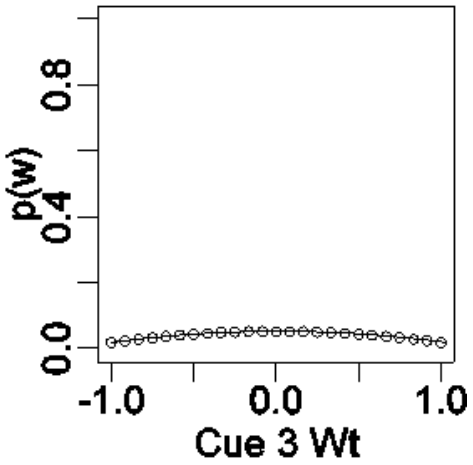
Noisy logic gate predictions for Expected Uncertainty…

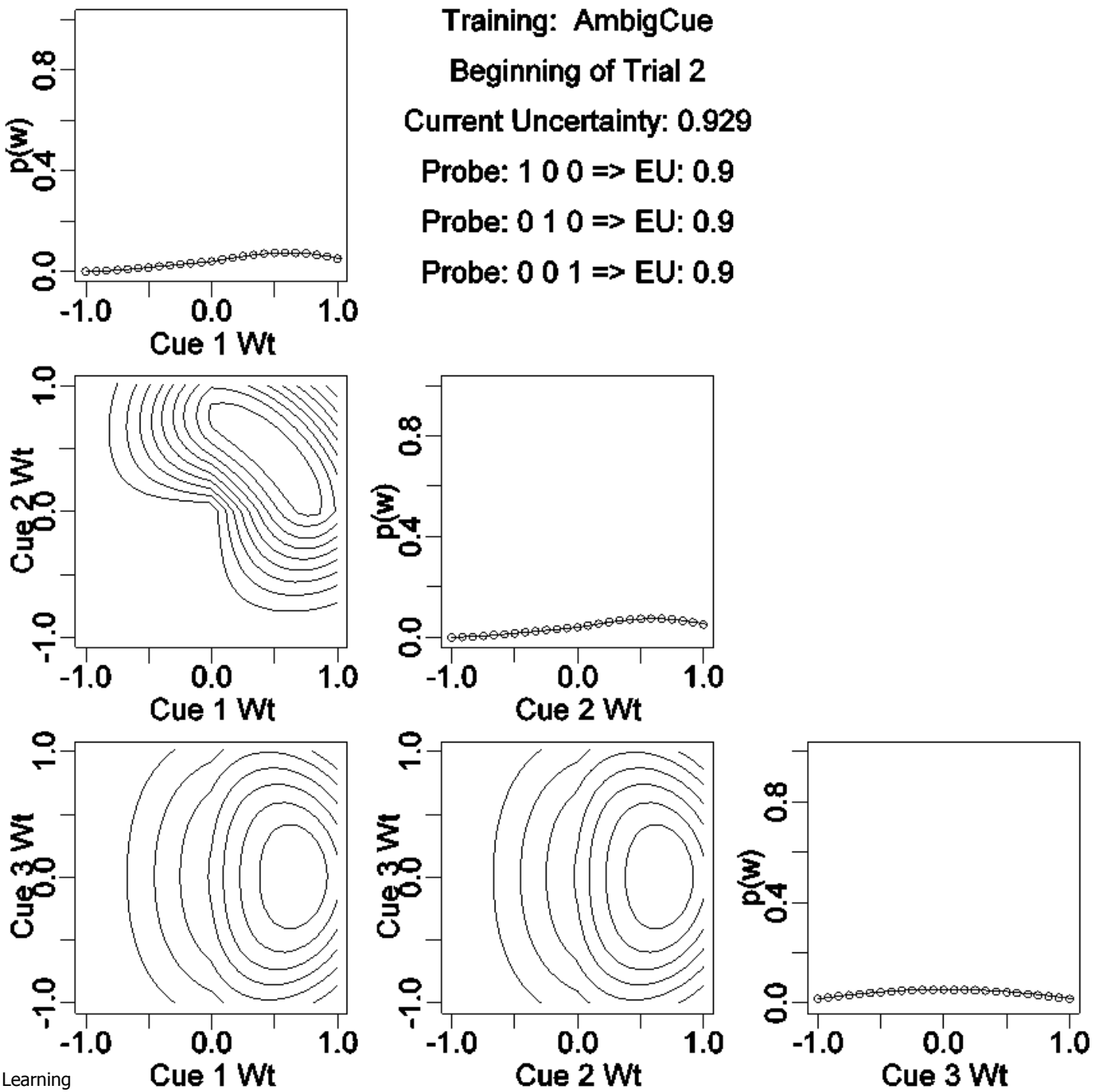


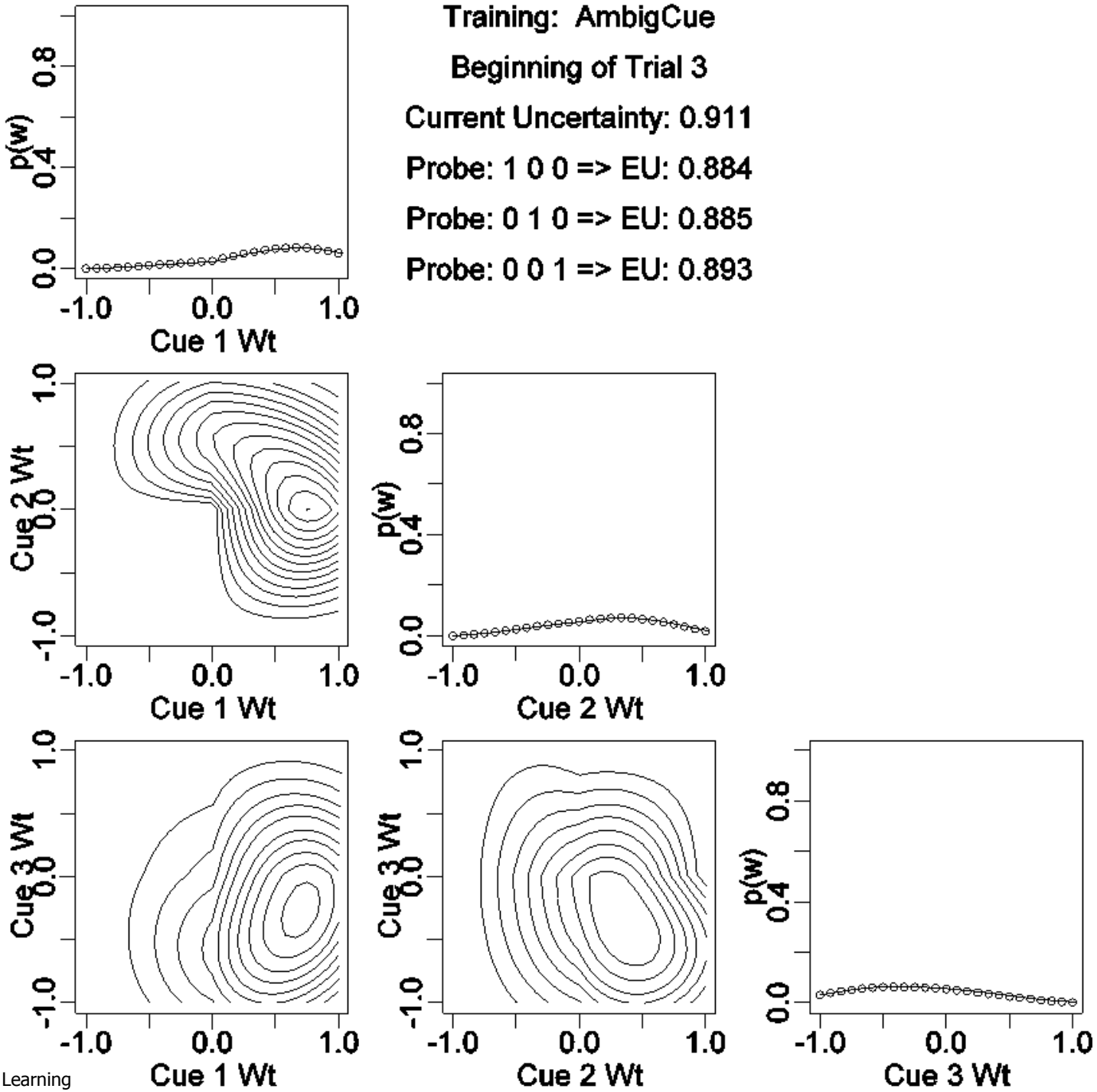


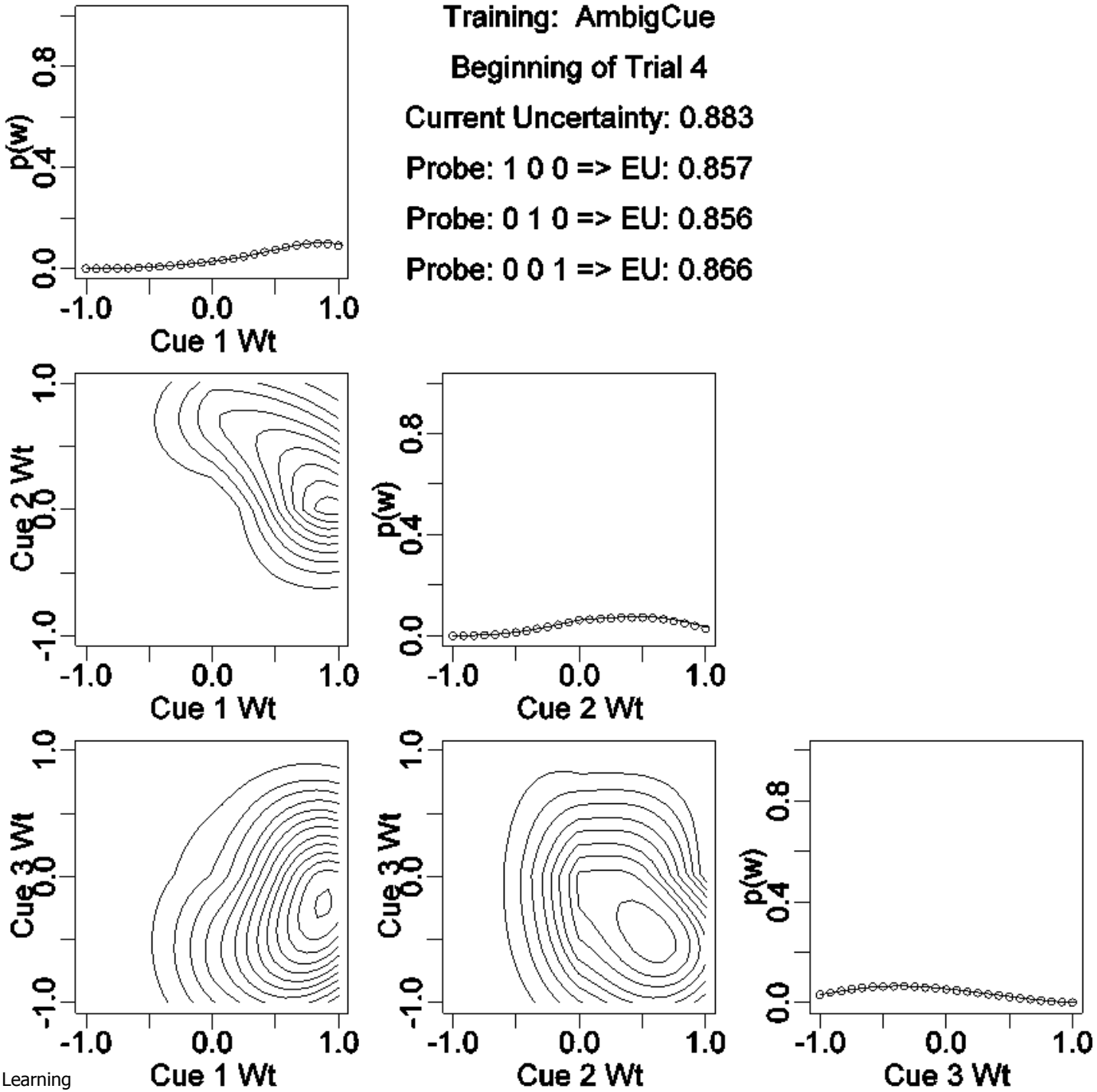


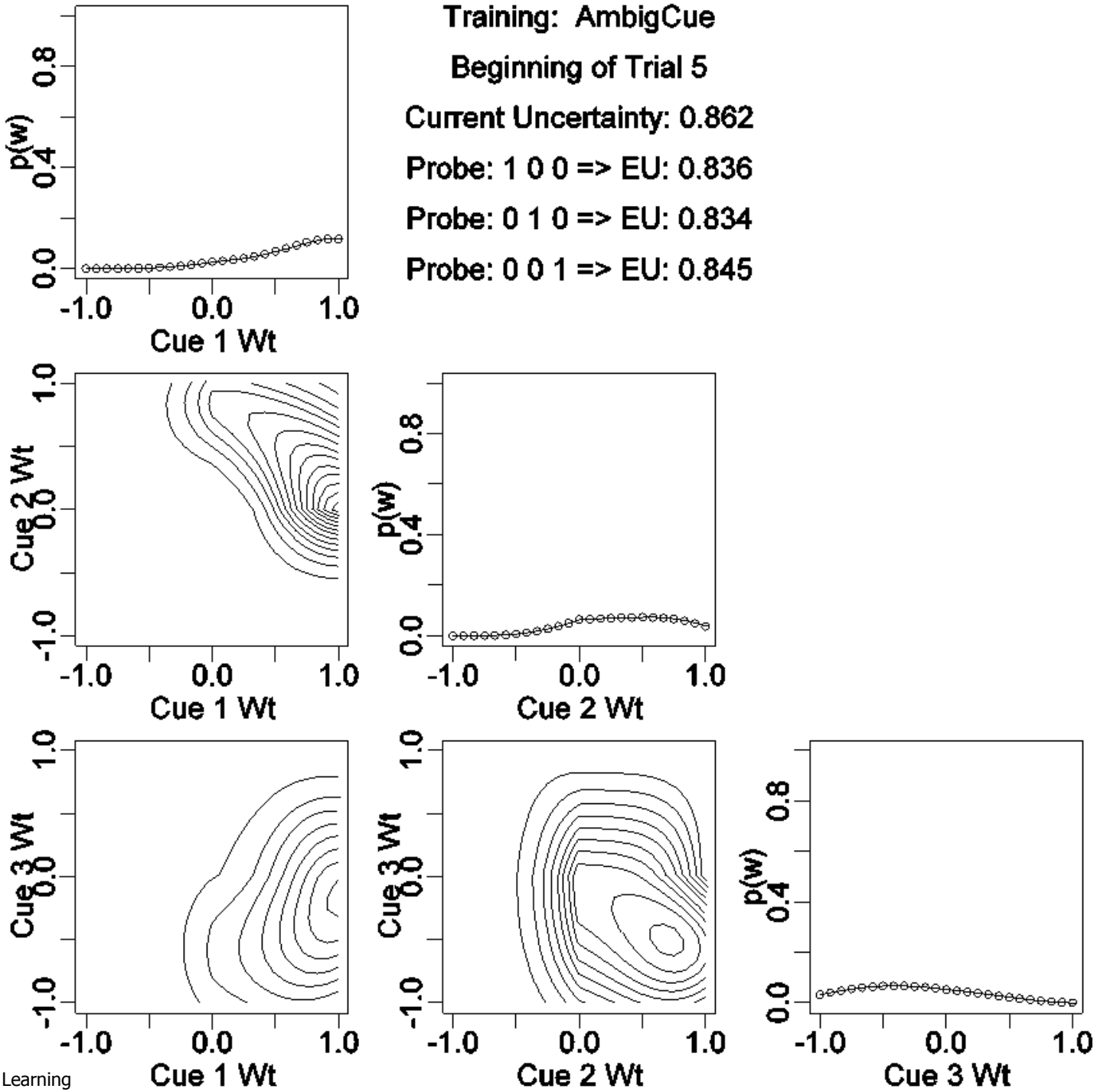


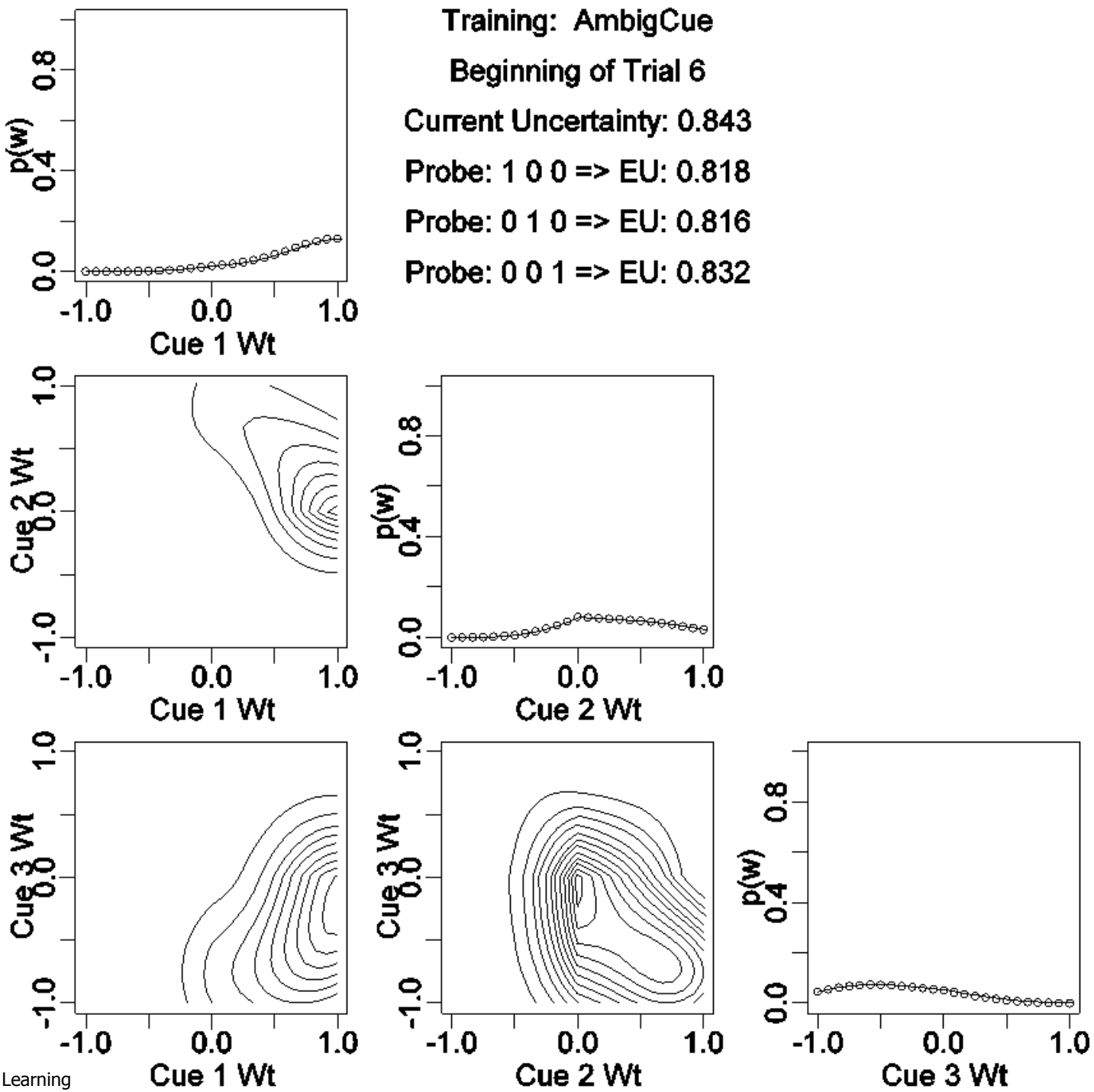


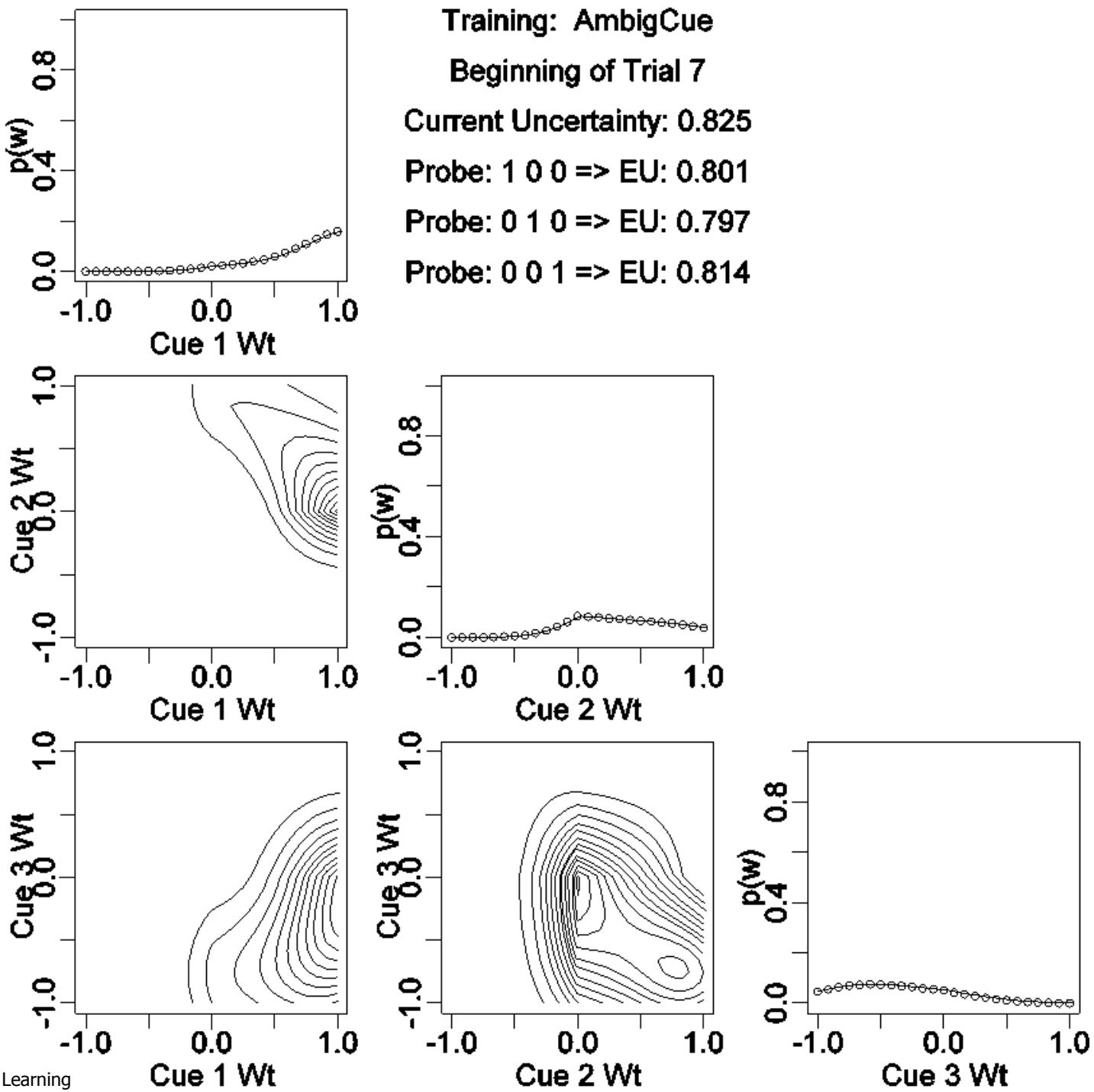

Training: AmbigCue
Beginning of Trial 2
Current Uncertainty: 0.929
Probe: 1 0 0 => EU: 0.9
Probe: 0 1 0 => EU: 0.9
Probe: 0 0 1 => EU: 0.9
p(w)
0.0
0.4
0.8
-1.0
0.0
1.0
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

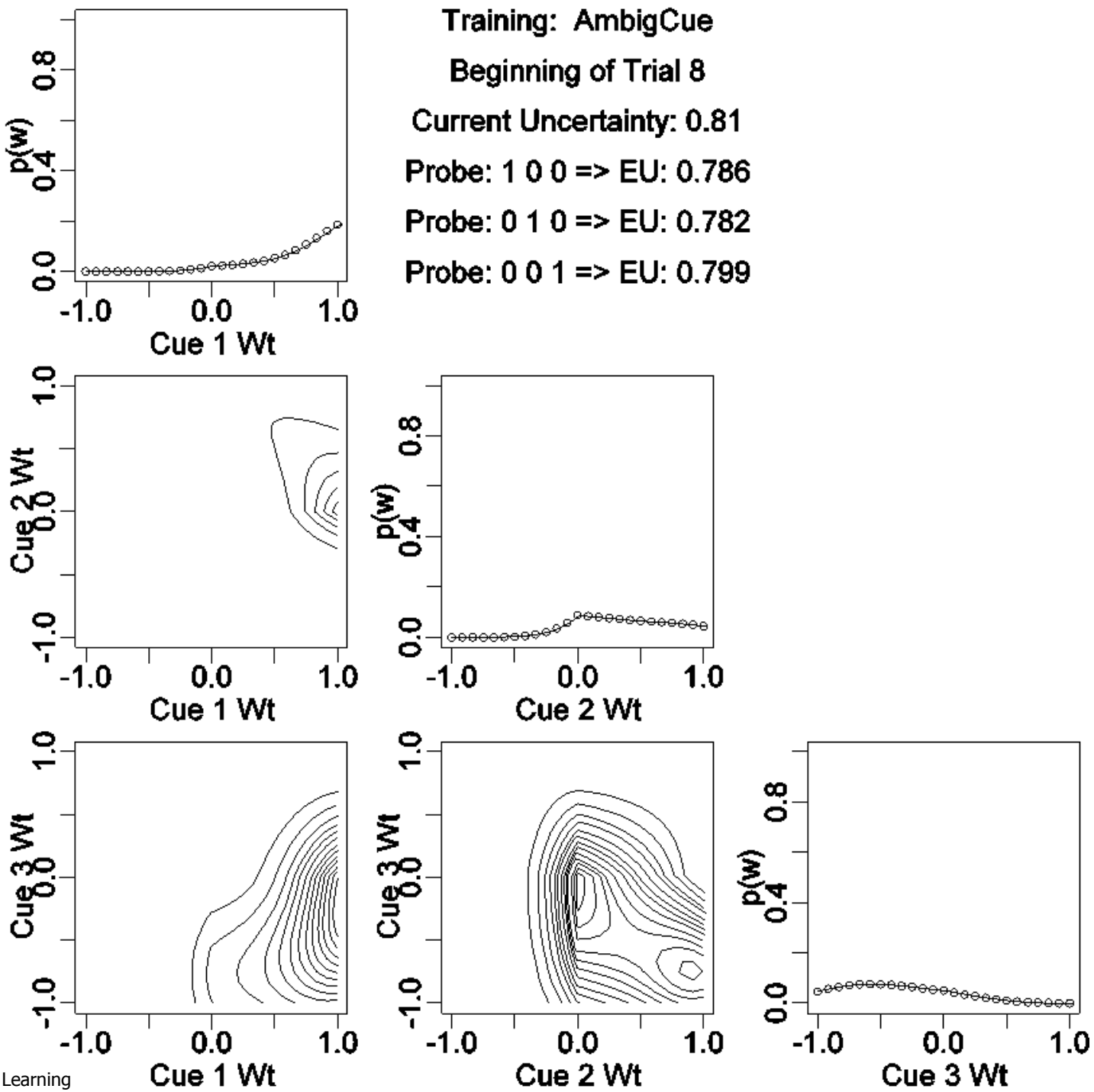

Training: AmbigCue
Beginning of Trial 3
Current Uncertainty: 0.911
Probe: 1 0 0 => EU: 0.884
Probe: 0 1 0 => EU: 0.885
Probe: 0 0 1 => EU: 0.893
p(w)
0.0
0.4
0.8
-1.0
0.0
1.0
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

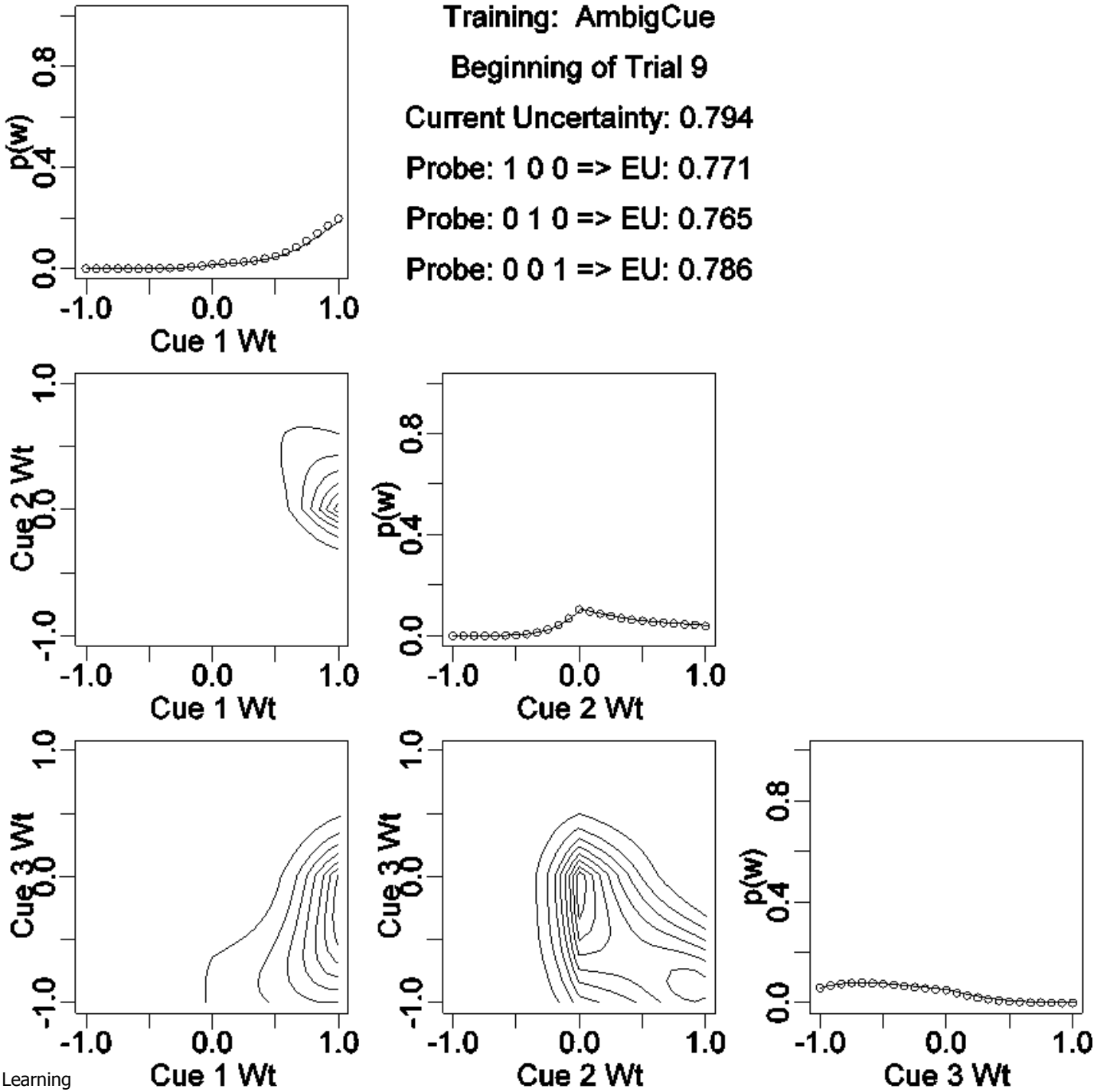

Training: AmbigCue
Beginning of Trial 4
Current Uncertainty: 0.883
Probe: 1 0 0 => EU: 0.857
Probe: 0 1 0 => EU: 0.856
Probe: 0 0 1 => EU: 0.866
p(w)
0.0
0.4
0.8
-1.0
0.0
1.0
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

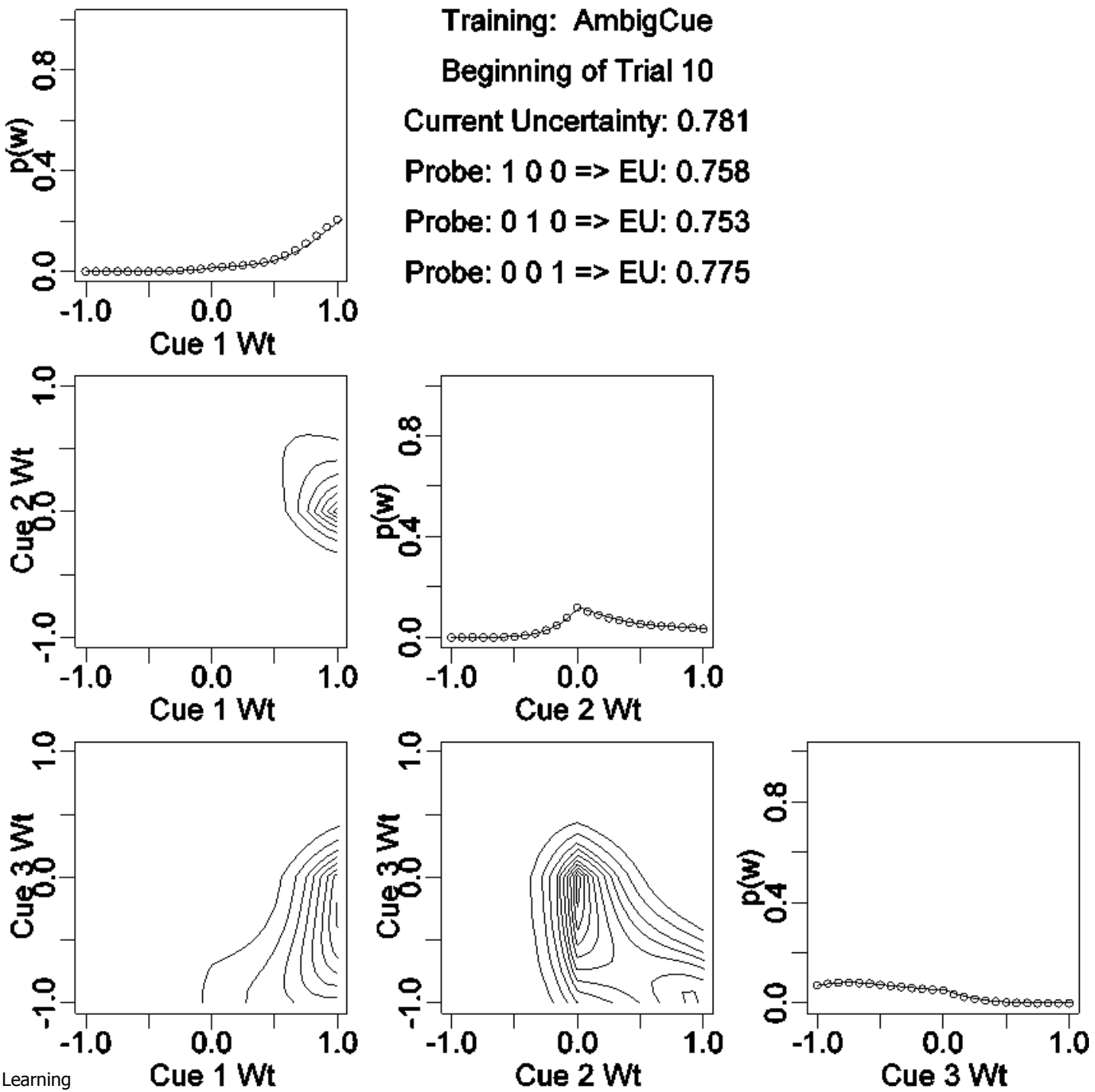

Training: AmbigCue
Beginning of Trial 5
Current Uncertainty: 0.862
Probe: 1 0 0 => EU: 0.836
Probe: 0 1 0 => EU: 0.834
Probe: 0 0 1 => EU: 0.845
p(w)
0.0
0.4
0.8
-1.0
0.0
1.0
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

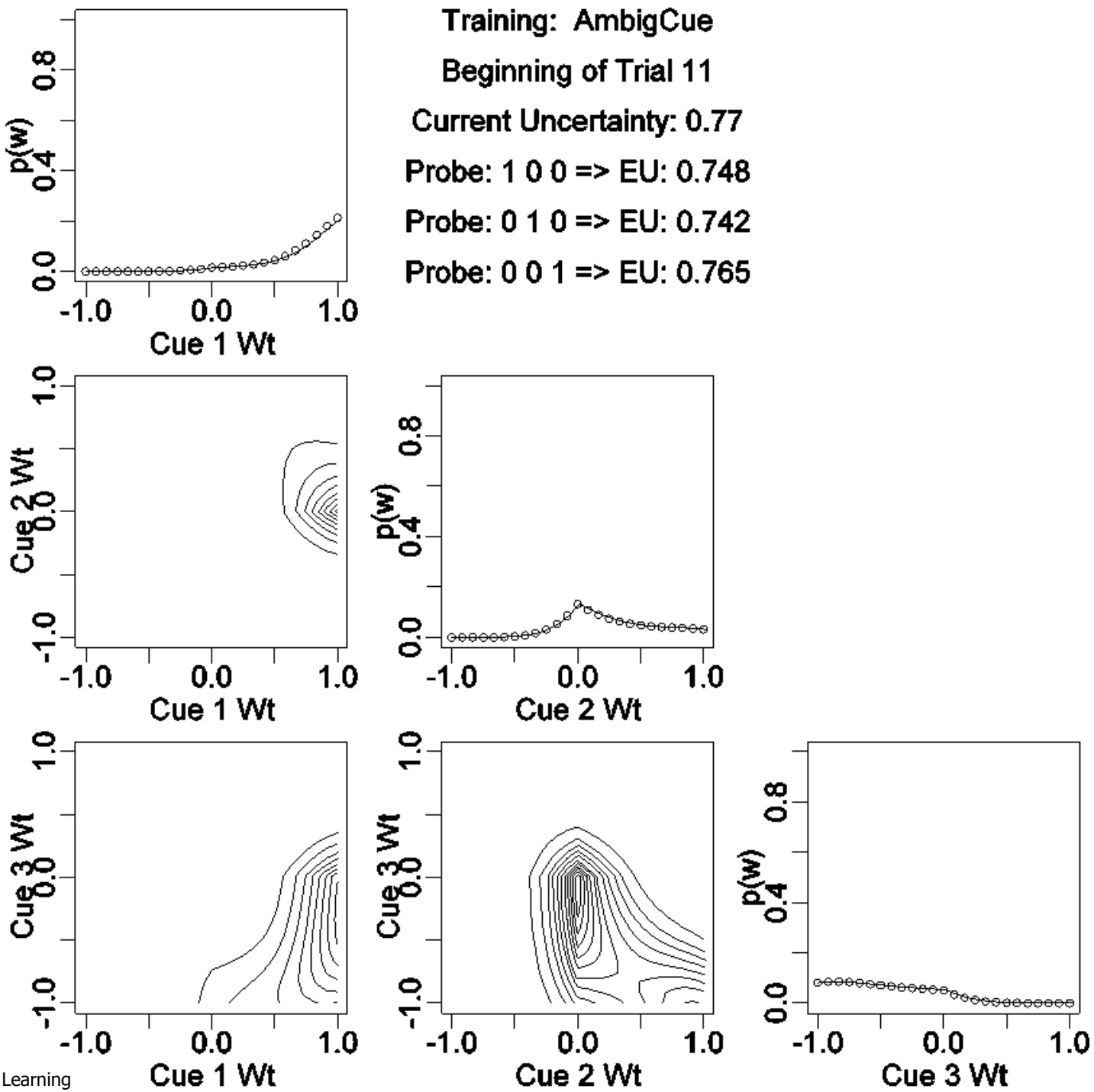

Training: AmbigCue
Beginning of Trial 6
Current Uncertainty: 0.843
Probe: 1 0 0 => EU: 0.818
Probe: 0 1 0 => EU: 0.816
Probe: 0 0 1 => EU: 0.832
p(w)
0.0
0.4
0.8
-1.0
0.0
1.0
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

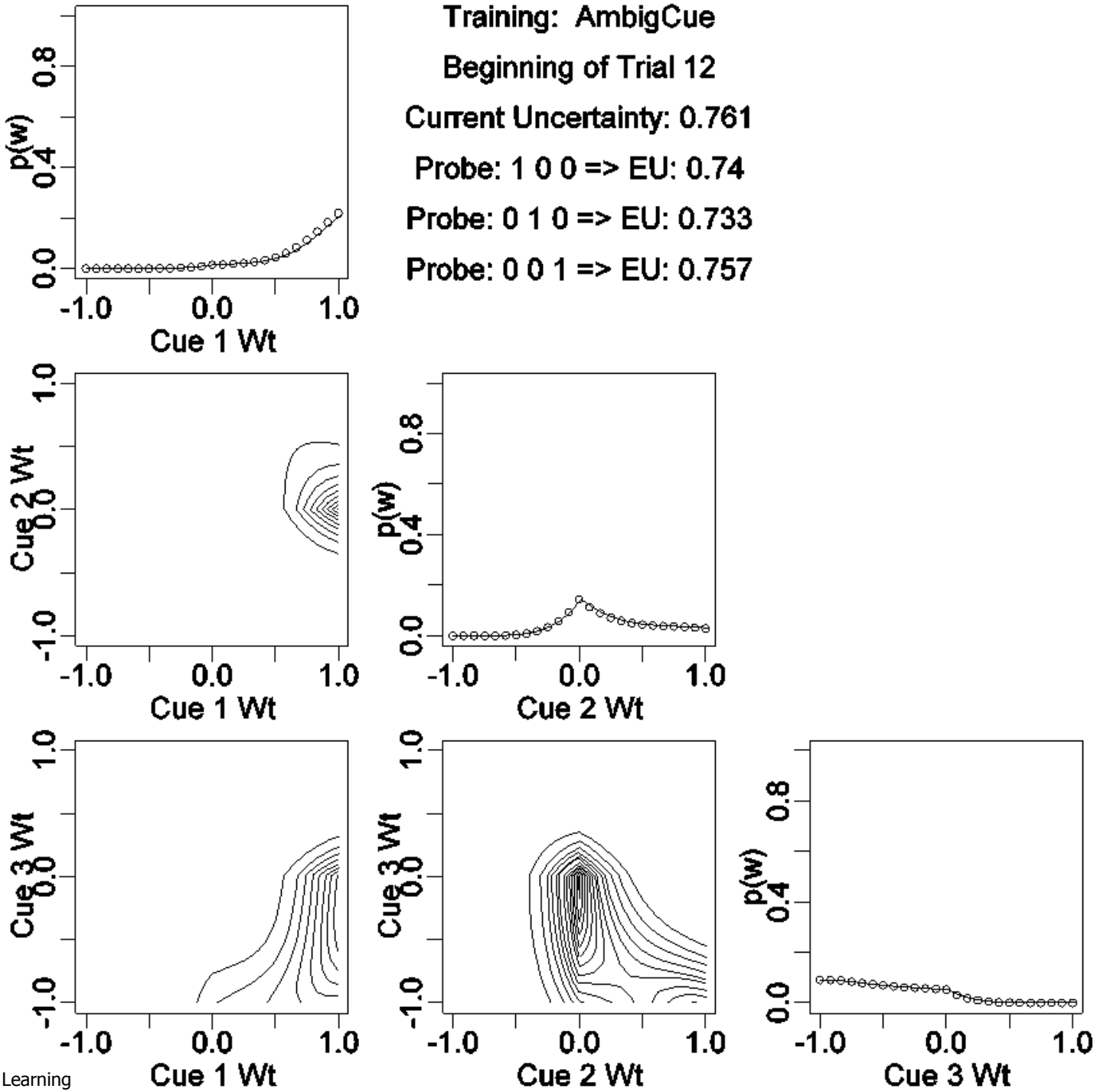

Training: AmbigCue
Beginning of Trial 7
Current Uncertainty: 0.825
Probe: 1 0 0 => EU: 0.801
Probe: 0 1 0 => EU: 0.797
Probe: 0 0 1 => EU: 0.814
p(w)
0.0
0.4
0.8
-1.0
0.0
1.0
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

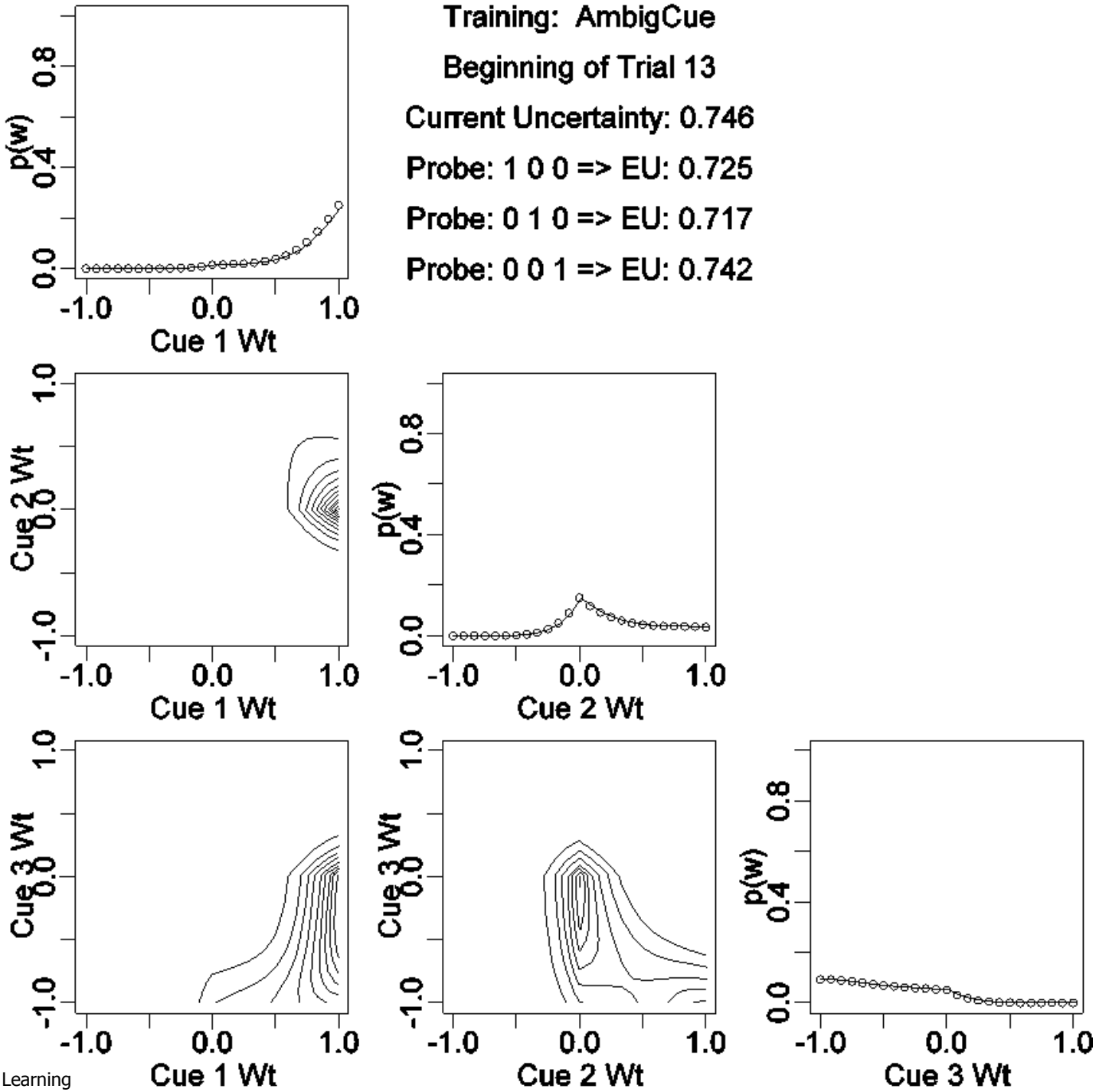

Training: AmbigCue
Beginning of Trial 8
Current Uncertainty: 0.81
Probe: 1 0 0 => EU: 0.786
Probe: 0 1 0 => EU: 0.782
Probe: 0 0 1 => EU: 0.799
p(w)
0.0
0.4
0.8
-1.0
0.0
1.0
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

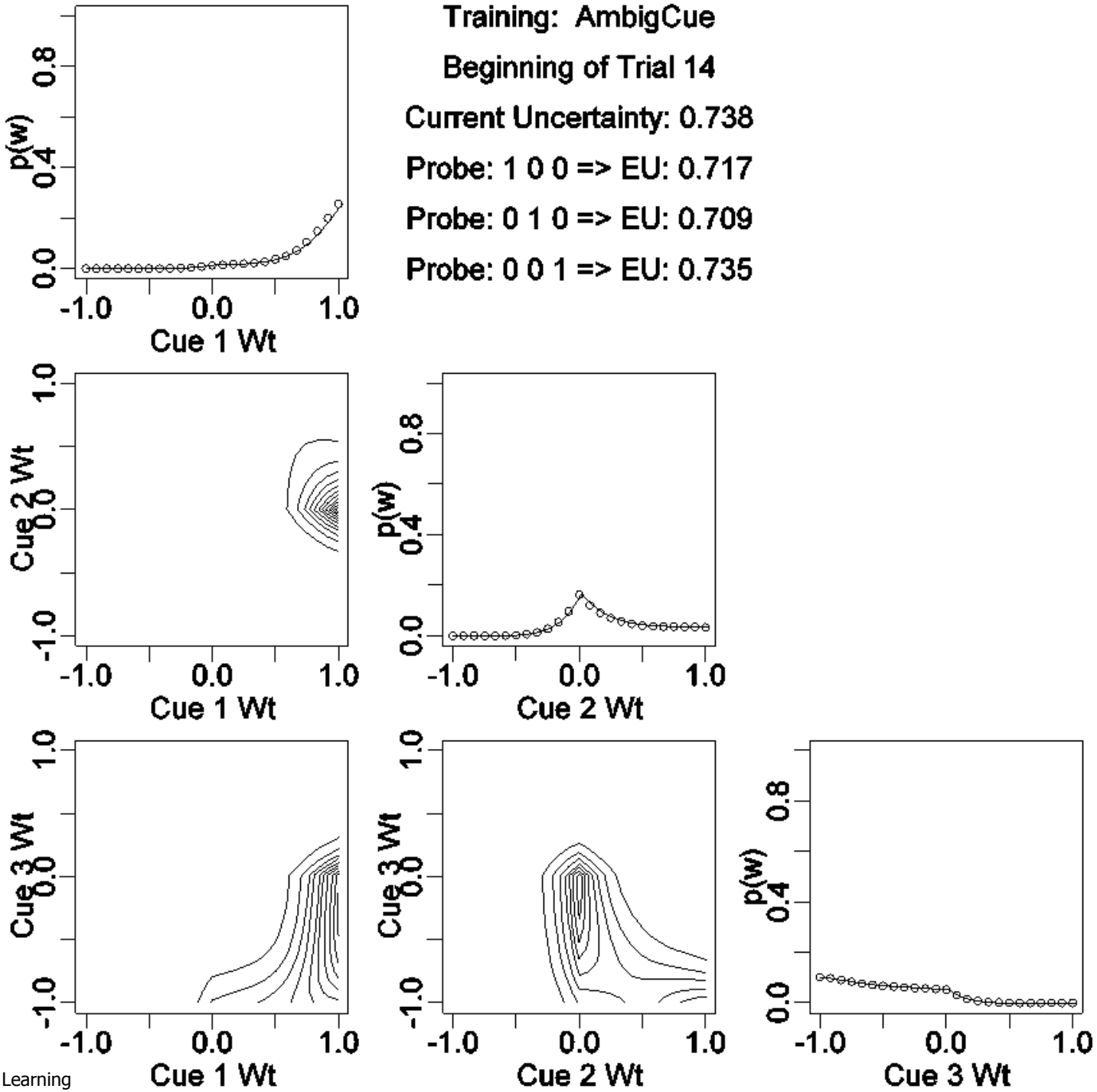

Training: AmbigCue
Beginning of Trial 9
Current Uncertainty: 0.794
Probe: 1 0 0 => EU: 0.771
Probe: 0 1 0 => EU: 0.765
Probe: 0 0 1 => EU: 0.786
p(w)
0.0
0.4
0.8
-1.0
0.0
1.0
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

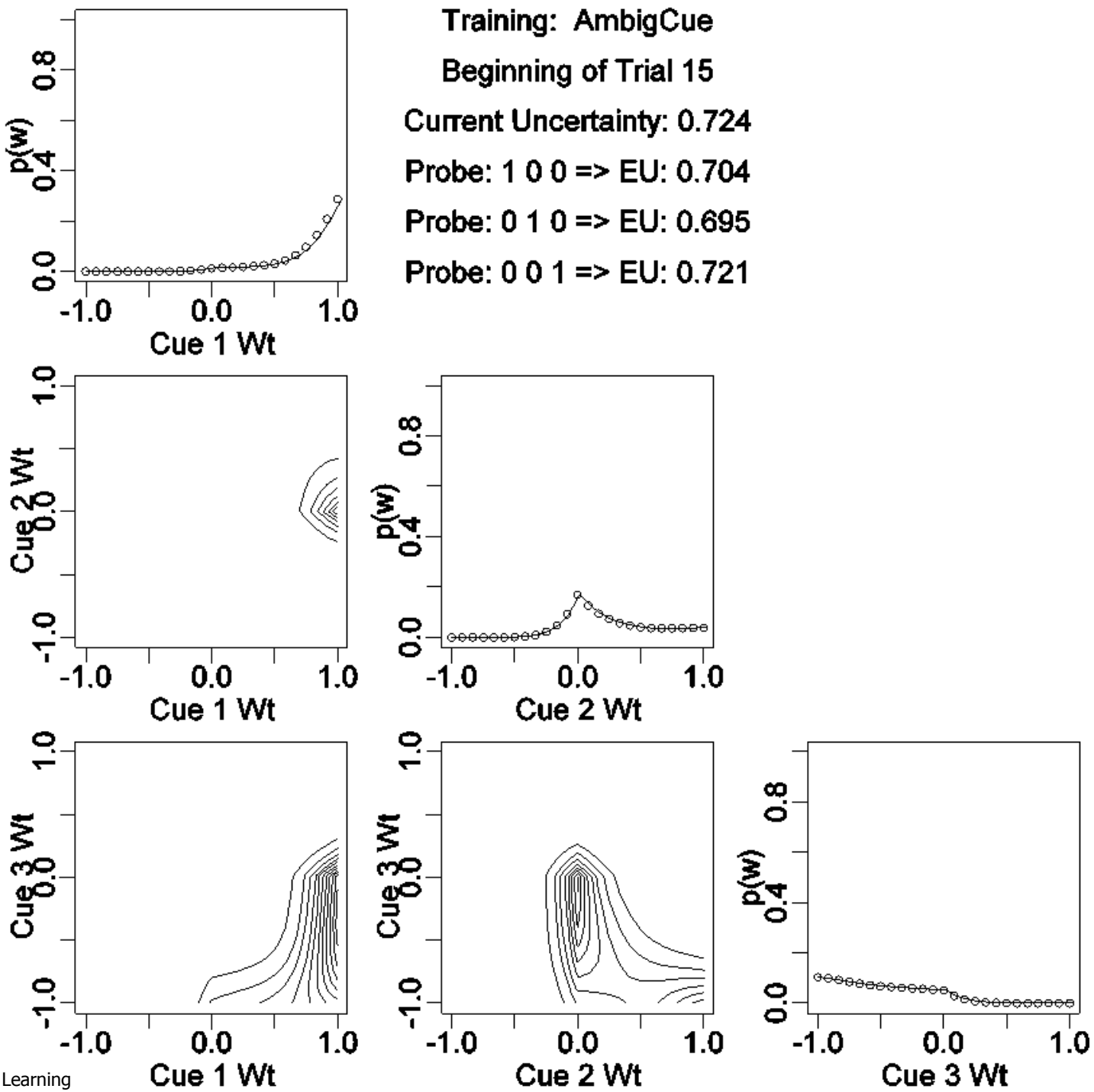

Training: AmbigCue
Beginning of Trial 10
Current Uncertainty: 0.781
Probe: 1 0 0 => EU: 0.758
Probe: 0 1 0 => EU: 0.753
Probe: 0 0 1 => EU: 0.775
p(w)
0.0
0.4
0.8
-1.0
0.0
1.0
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

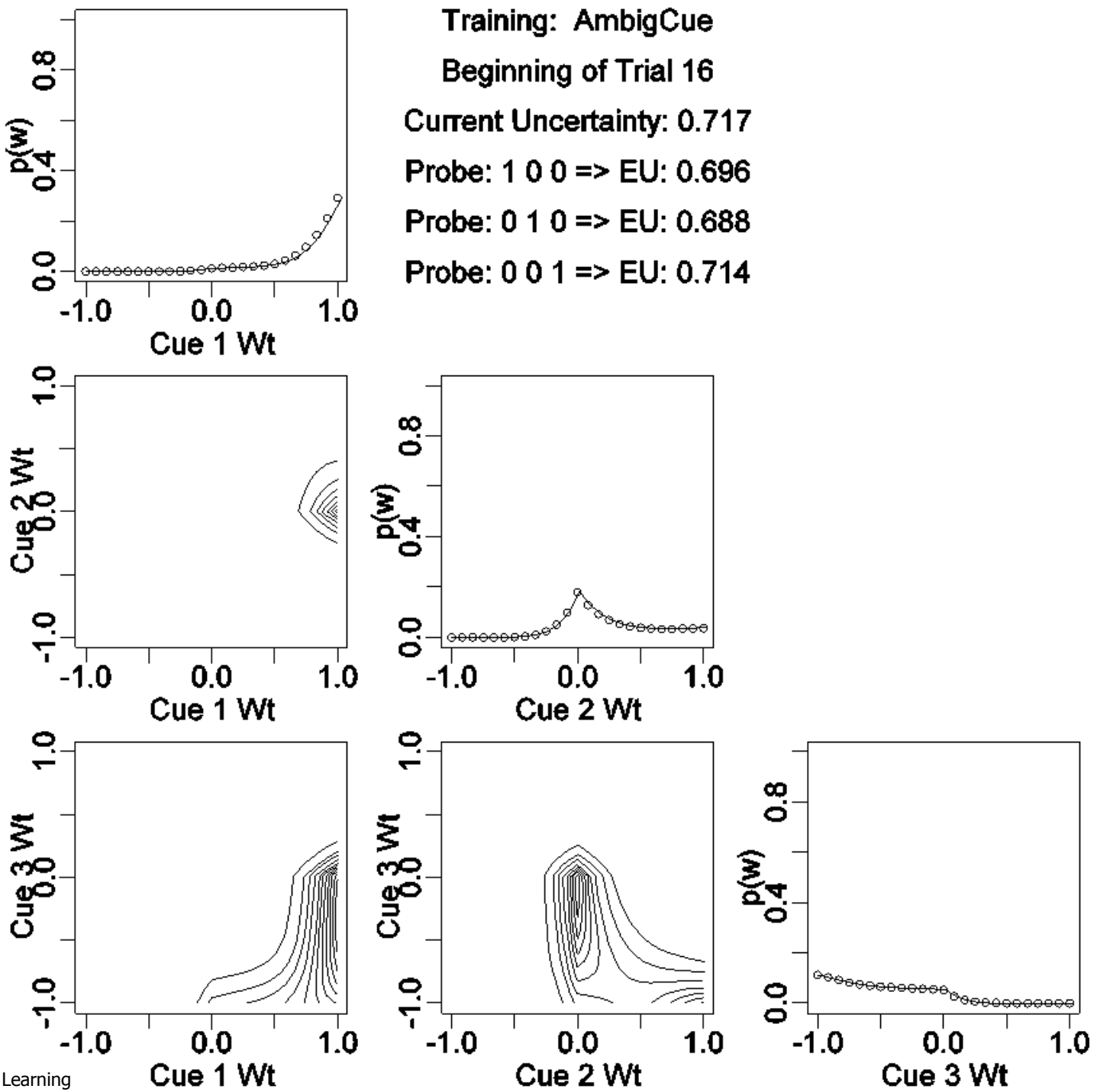

Training: AmbigCue
Beginning of Trial 11
Current Uncertainty: 0.77
Probe: 1 0 0 => EU: 0.748
Probe: 0 1 0 => EU: 0.742
Probe: 0 0 1 => EU: 0.765
p(w)
0.0
0.4
0.8
-1.0
0.0
1.0
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

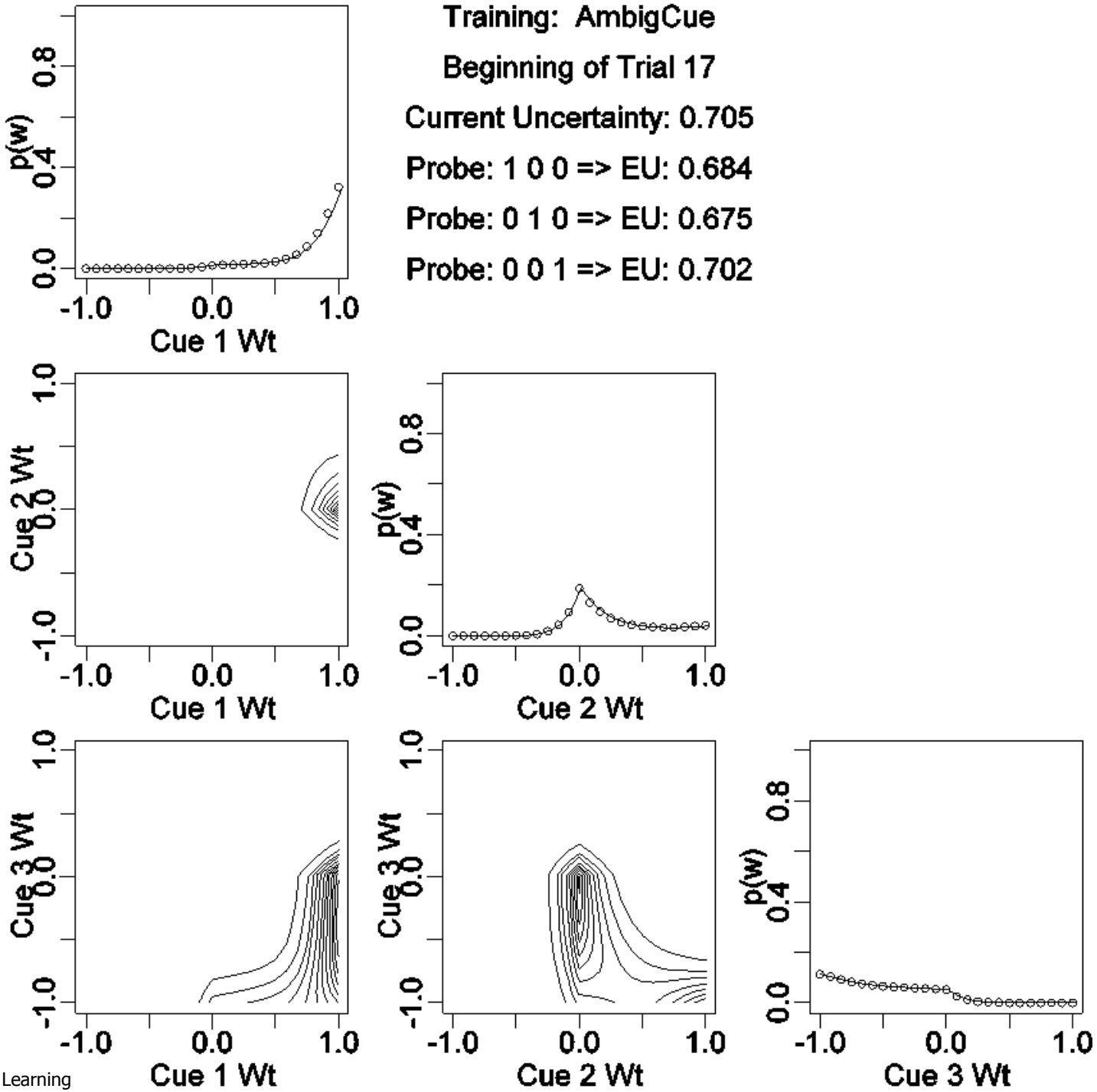

Training: AmbigCue
Beginning of Trial 12
Current Uncertainty: 0.761
Probe: 1 0 0 => EU: 0.74
Probe: 0 1 0 => EU: 0.733
Probe: 0 0 1 => EU: 0.757
p(w)
0.0
0.4
0.8
-1.0
0.0
1.0
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

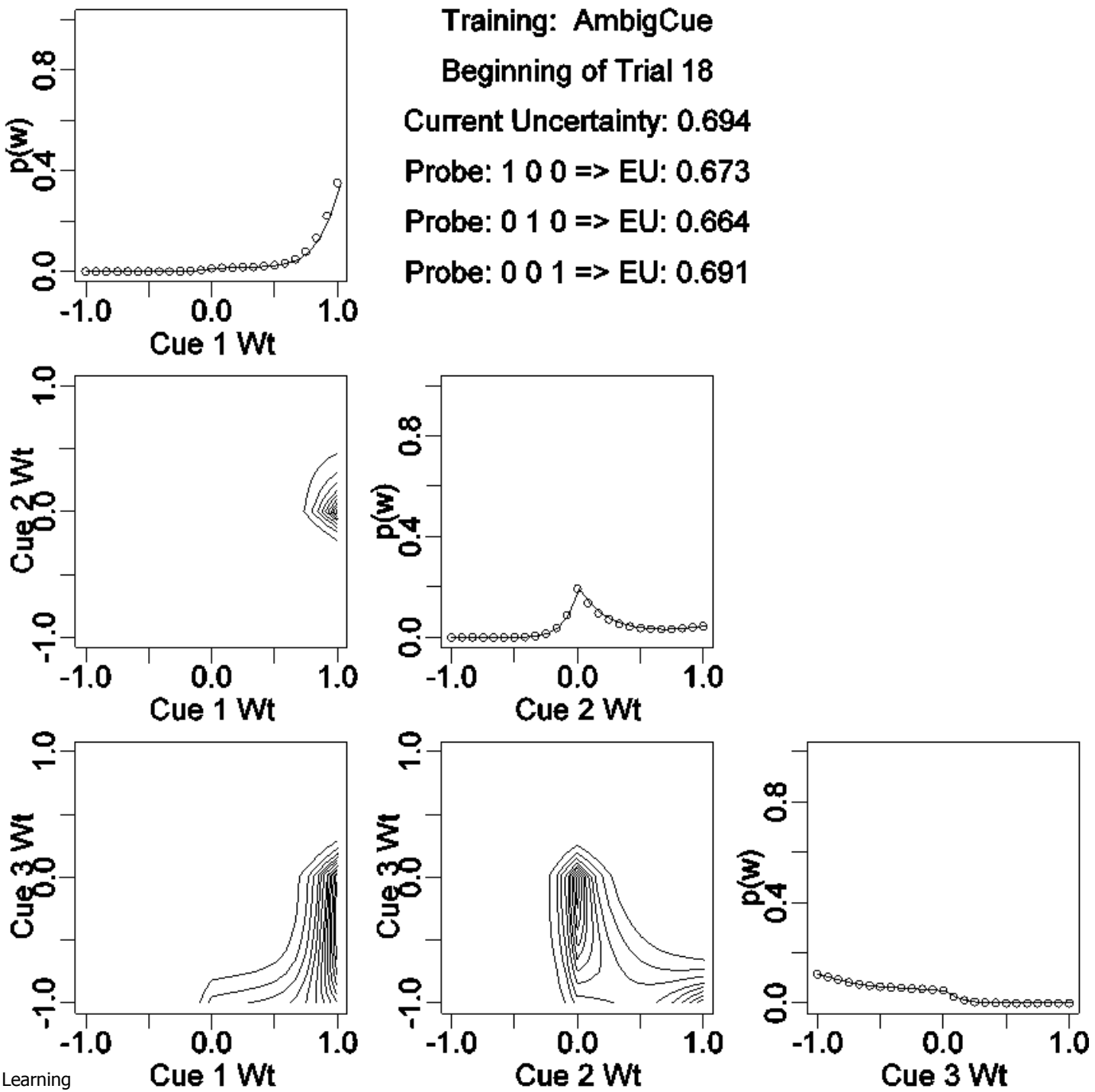

Training: AmbigCue
Beginning of Trial 13
Current Uncertainty: 0.746
Probe: 1 0 0 => EU: 0.725
Probe: 0 1 0 => EU: 0.717
Probe: 0 0 1 => EU: 0.742
p(w)
0.0
0.4
0.8
-1.0
0.0
1.0
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

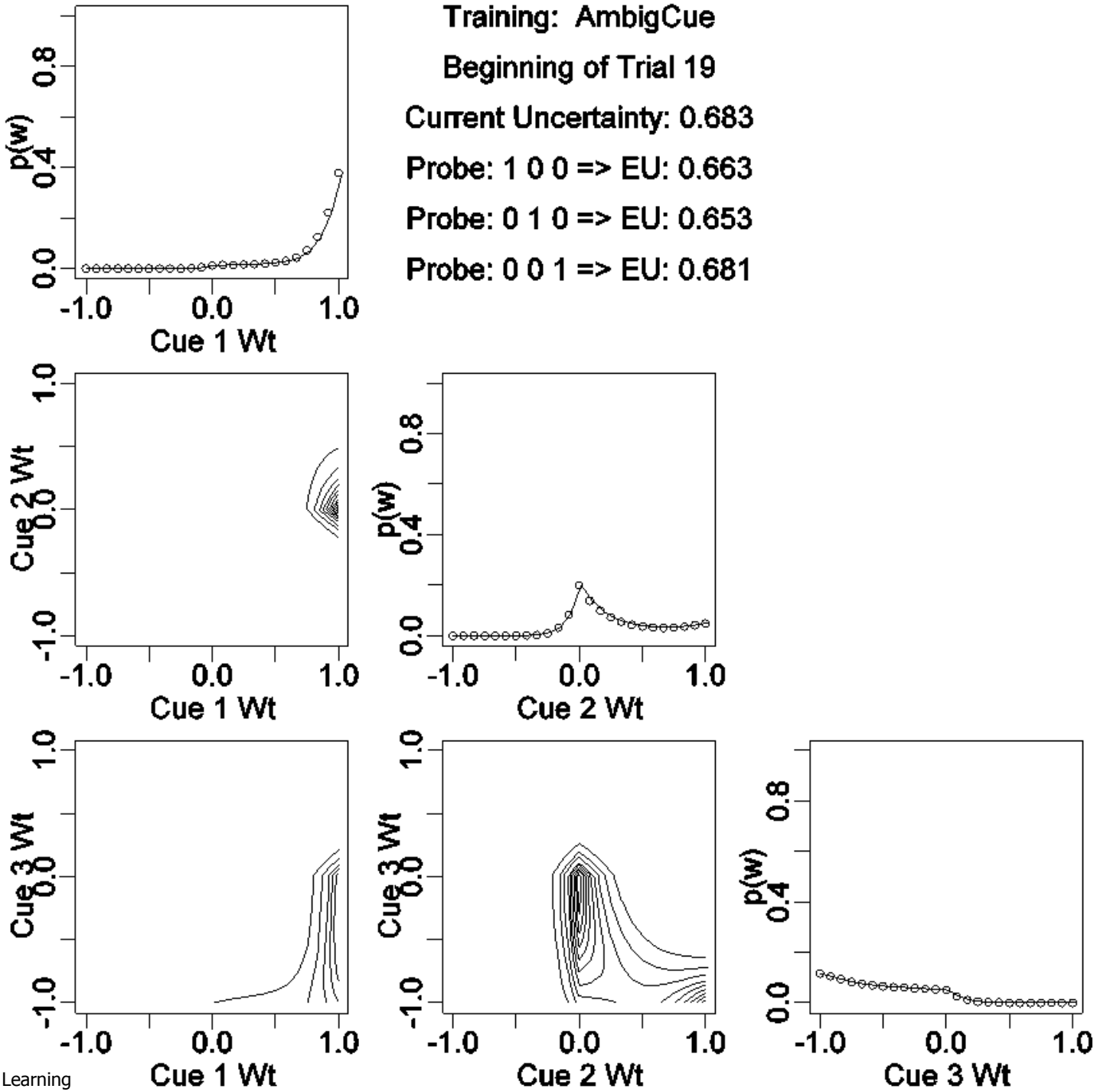

Training: AmbigCue
Beginning of Trial 14
Current Uncertainty: 0.738
Probe: 1 0 0 => EU: 0.717
Probe: 0 1 0 => EU: 0.709
Probe: 0 0 1 => EU: 0.735
p(w)
0.0
0.4
0.8
-1.0
0.0
1.0
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

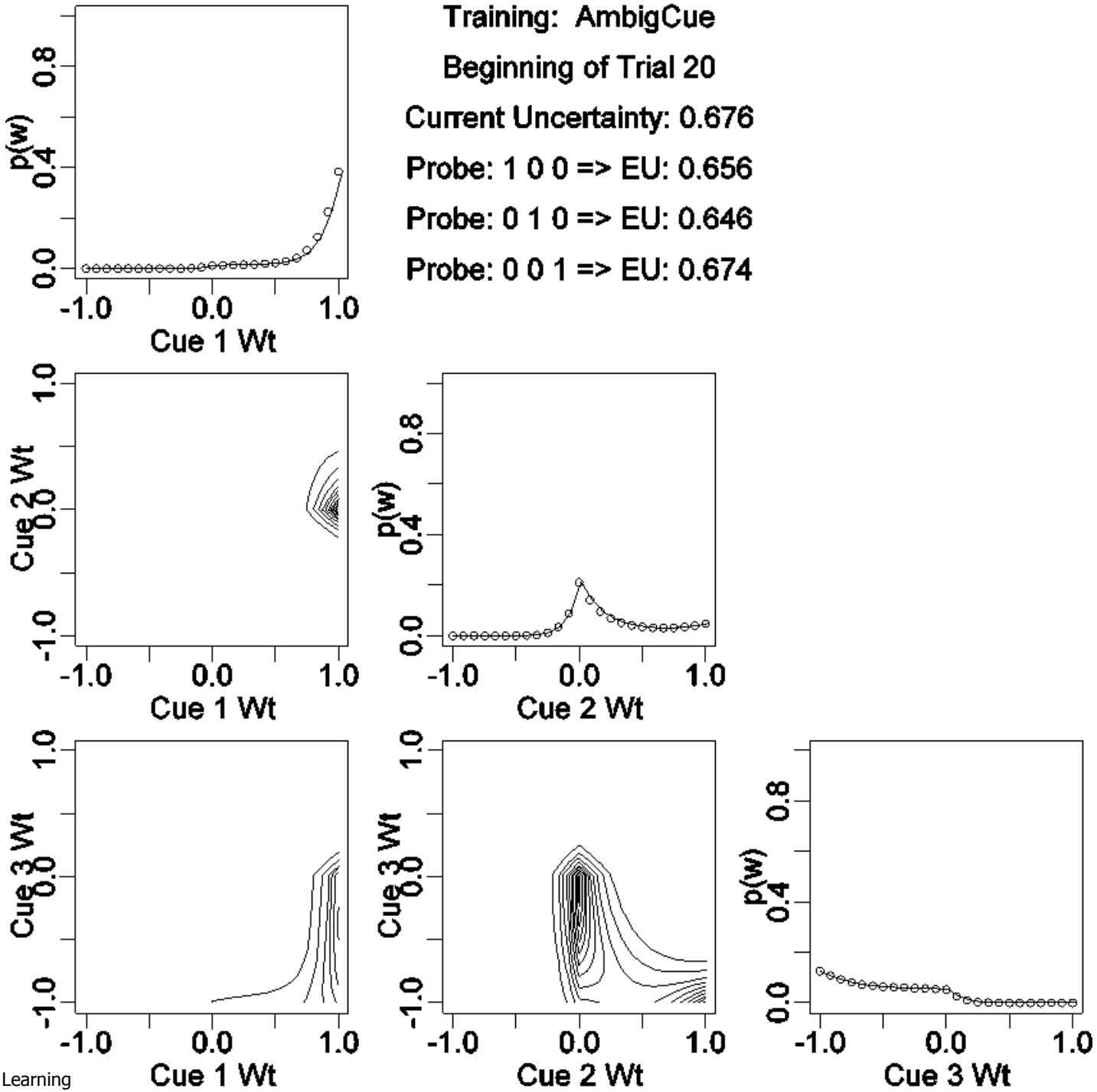

Training: AmbigCue
Beginning of Trial 15
Current Uncertainty: 0.724
Probe: 1 0 0 => EU: 0.704
Probe: 0 1 0 => EU: 0.695
Probe: 0 0 1 => EU: 0.721
p(w)
0.0
0.4
0.8
-1.0
0.0
1.0
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

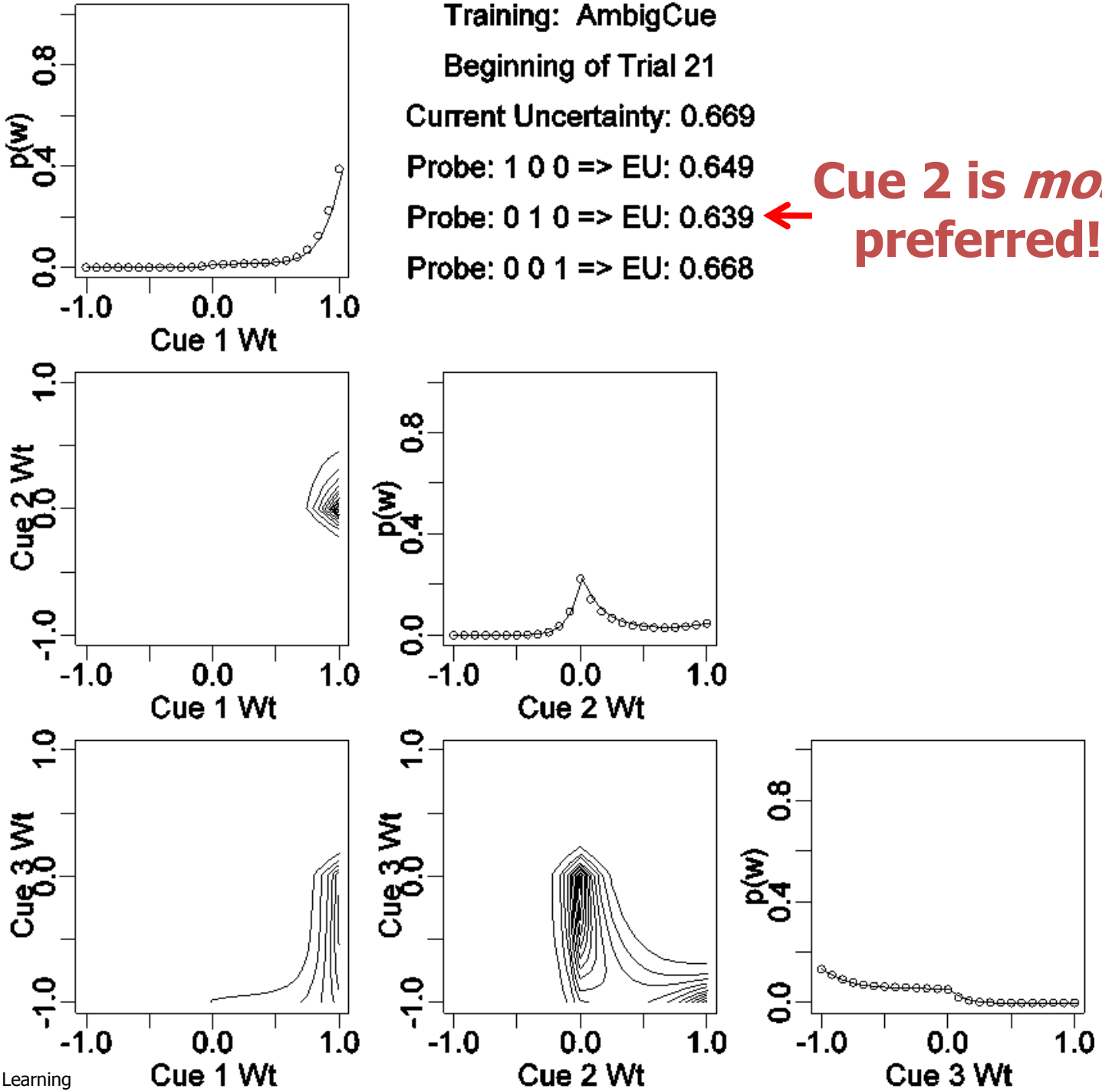

Training: AmbigCue
Beginning of Trial 16
Current Uncertainty: 0.717
Probe: 1 0 0 => EU: 0.696
Probe: 0 1 0 => EU: 0.688
Probe: 0 0 1 => EU: 0.714
p(w)
0.0
0.4
0.8
-1.0
0.0
1.0
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt


Training: AmbigCue
Beginning of Trial 17
Current Uncertainty: 0.705
Probe: 1 0 0 => EU: 0.684
Probe: 0 1 0 => EU: 0.675
Probe: 0 0 1 => EU: 0.702
p(w)
0.0
0.4
0.8
-1.0
0.0
1.0
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt


Training: AmbigCue
Beginning of Trial 18
Current Uncertainty: 0.694
Probe: 1 0 0 => EU: 0.673
Probe: 0 1 0 => EU: 0.664
Probe: 0 0 1 => EU: 0.691
p(w)
0.0
0.4
0.8
-1.0
0.0
1.0
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt


Training: AmbigCue
Beginning of Trial 19
Current Uncertainty: 0.683
Probe: 1 0 0 => EU: 0.663
Probe: 0 1 0 => EU: 0.653
Probe: 0 0 1 => EU: 0.681
p(w)
0.0
0.4
0.8
-1.0
0.0
1.0
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt

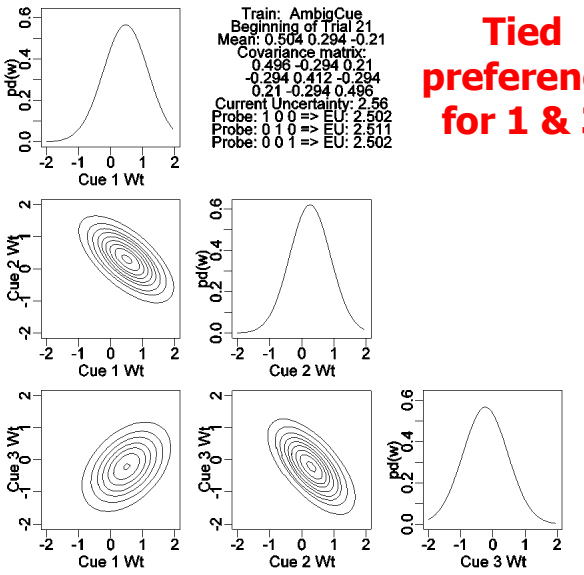

Training: AmbigCue
Beginning of Trial 20
Current Uncertainty: 0.676
Probe: 1 0 0 => EU: 0.656
Probe: 0 1 0 => EU: 0.646
Probe: 0 0 1 => EU: 0.674
p(w)
0.0
0.4
0.8
-1.0
0.0
1.0
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt


Training: AmbigCue
Beginning of Trial 21
Current Uncertainty: 0.669
Probe: 1 0 0 => EU: 0.649
Probe: 0 1 0 => EU: 0.639
Probe: 0 0 1 => EU: 0.668
Cue 2 is most preferred!
p(w)
Cue 1 Wt
Cue 2 Wt
Cue 3 Wt
-1.0
0.0
1.0
0.0
0.4
0.8

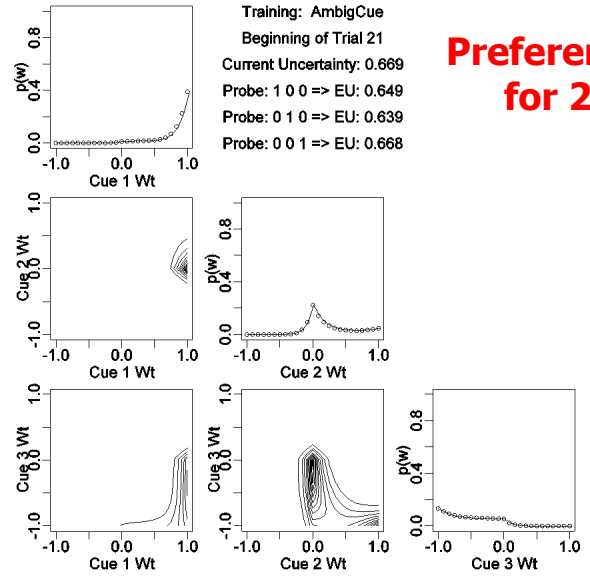
Kalman filter

Noisy logic gate

Minimize Expected Uncertainty

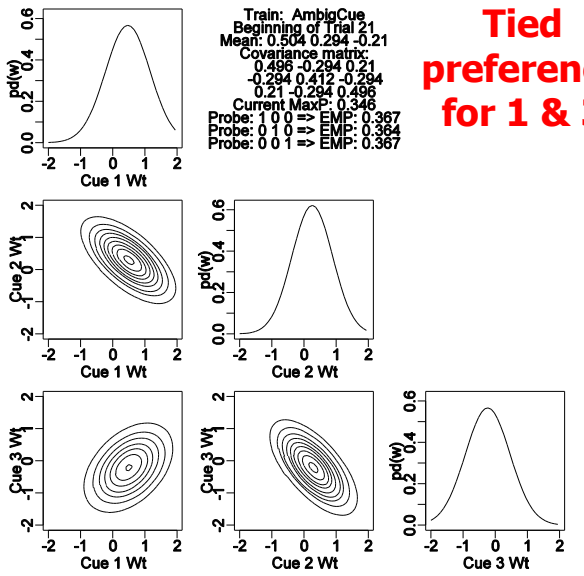


Tied preference for 1 & 3

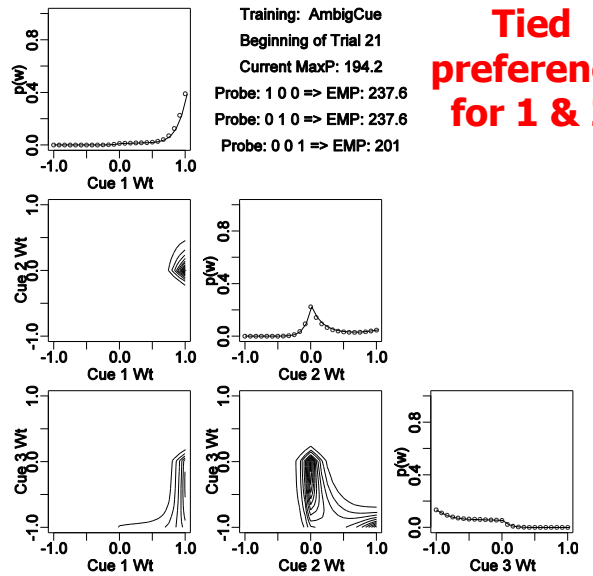


Preference for 2

Maximize Expected Max P



Tied preference for 1 & 3



Tied preference for 1 & 2

Summary of Active Learning Predictions and Match to Human Preference

		Bayesian model of (passive) associative learning					
		Kalman filter			Noisy logic gate		
Goal for active learning	**Minimize Expected Uncertainty**	(Backward) Blocking ☺	Blocking & Reduced Overshadow. ☹	Ambiguous Cue ☹	(Backward) Blocking ☺	Blocking & Reduced Overshadow. ☺	Ambiguous Cue ☺
	Maximize Expected Max P	(Backward) Blocking ☺	Blocking & Reduced Overshadow. ☹	Ambiguous Cue ☹	(Backward) Blocking ☺	Blocking & Reduced Overshadow. ☺	Ambiguous Cue ☹

The Future of (Associative) Learning Theory

		Bayesian model of (passive) associative learning		
		Kalman filter	**Noisy logic gate**	**New Models …**
Goal for active learning	**Minimize Expected Uncertainty**	New Training Structures	New Training Structures	New Training Structures
	Maximize Expected Max P	New Training Structures	New Training Structures	New Training Structures
	New Goals …	New Training Structures	New Training Structures	New Training Structures

The Future of (Associative) Learning Theory

		Bayesian model of (passive) associative learning		
		Kalman filter	**Noisy logic gate**	**New Models …**
Goal for active learning	**Minimize Expected Uncertainty**	New Training Structures	New Training S	
	Maximize Expected Max P	New Training Structures	New Training Structures	New Training Structures
	New Goals …	New Training Structures	New Training Structures	New Training Structures

Rich hierarchical models.
Other levels of analysis: neuron, functional module within individual, corporation.

The Future of (Associative) Learning Theory

		Bayesian model of (passive) associative learning		
		Kalman filter	**Noisy logic gate**	**New Models …**
Goal for active learning	**Minimize Expected Uncertainty**	New Training Structures	New Training S	
	Maximize Expected Max P	New Training Structures	New Training Structures	New Training Structures
	New Goals …			Training Structures

Rich hierarchical models.
Other levels of analysis: neuron, functional module within individual, corporation.

Different information objectives.
Incorporate utilities of expected outcomes.
Incorporate costs of different probes.
Consider multiple steps ahead.