Load and Attentional Bayes

based on: Peter Dayan, Gatsby (2009)

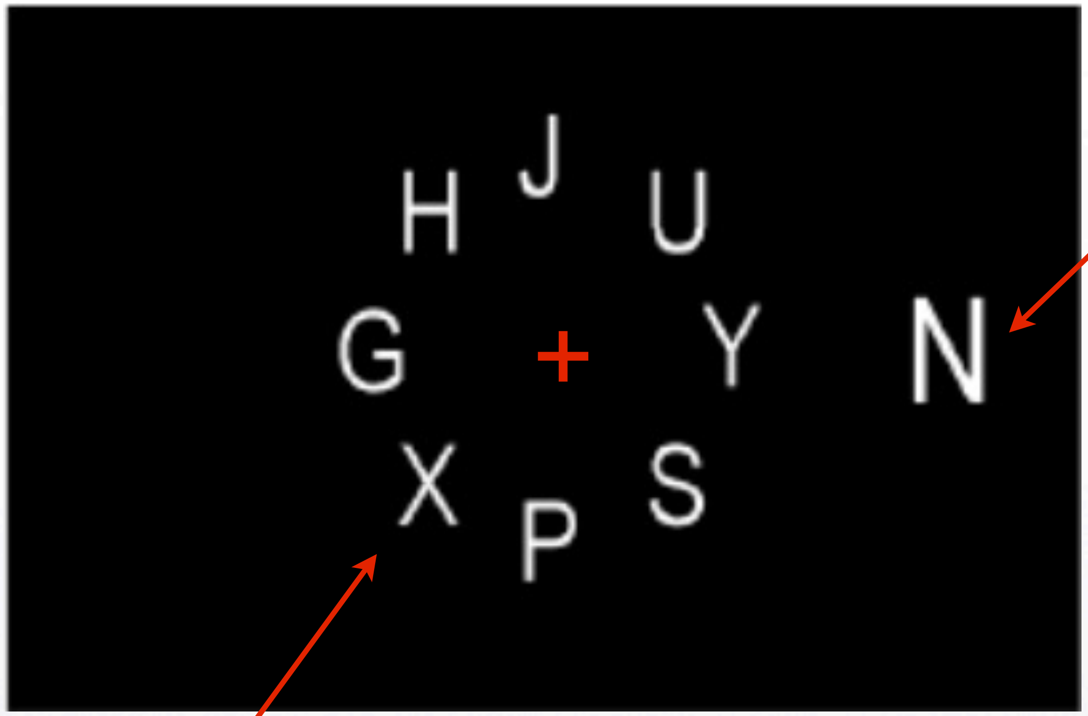
Zoya Bylinskii, May 3, 2014 - paper presentation for 9.S912

Attentional Load

Attentional load hypothesis:

- when little attention is required to solve a set task, inputs associated with distractor stimuli “leak through”, and cause disruption
- when the task is difficult, attention is totally occupied, leaving nothing left over (to attend to distractors)

Attentional Load Experiments
Lavie and de Fockert (2003)



distractor

- **compatible:** same identity as target
- **incompatible:** X if target is N, and vice versa
- **neutral:** some letter other than X or N

target

- could take on the value of X or N
- could occur in any of the 8 positions around the fixation point

Attentional Load Experiments
Lavie and de Fockert (2003)

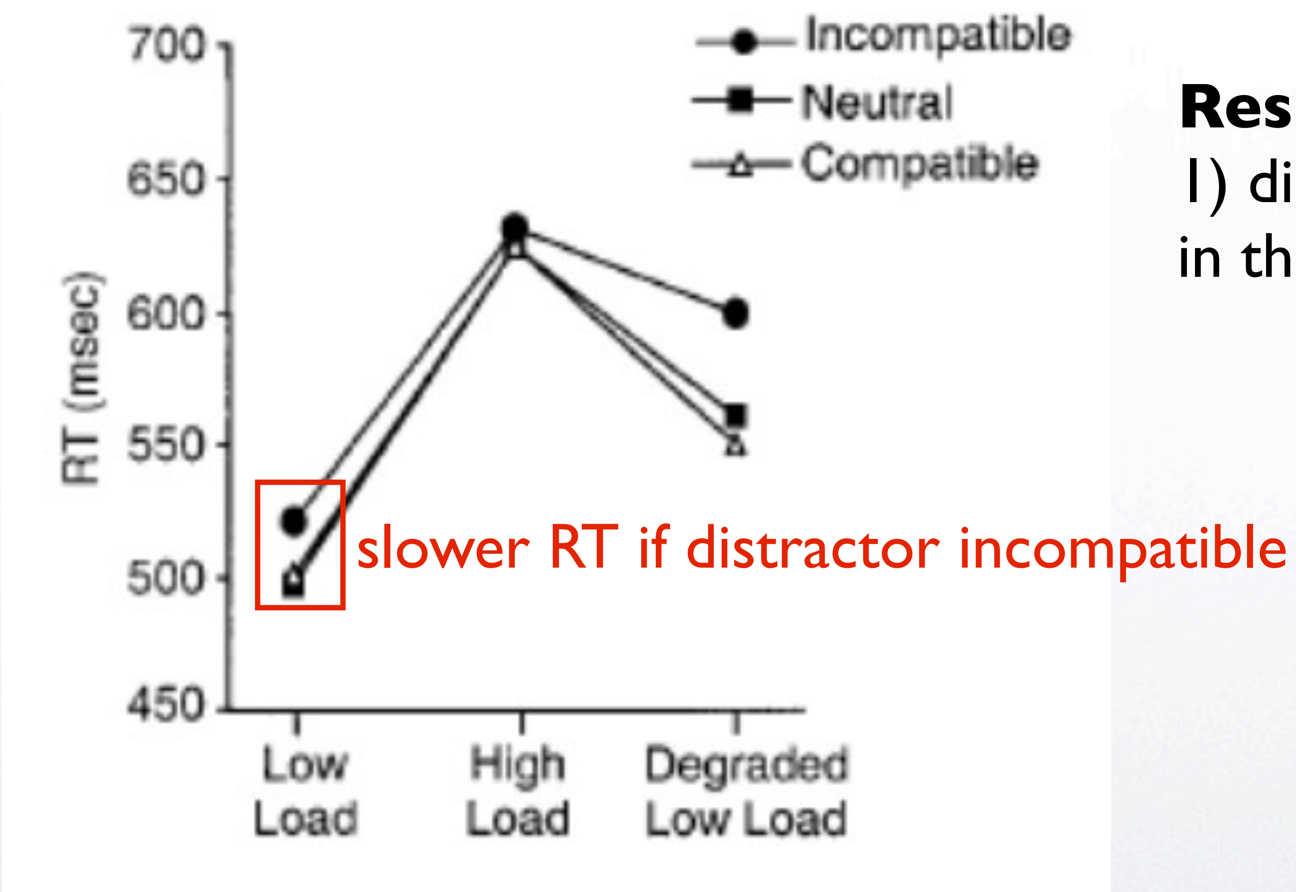
high-load

low-load

degraded low-load

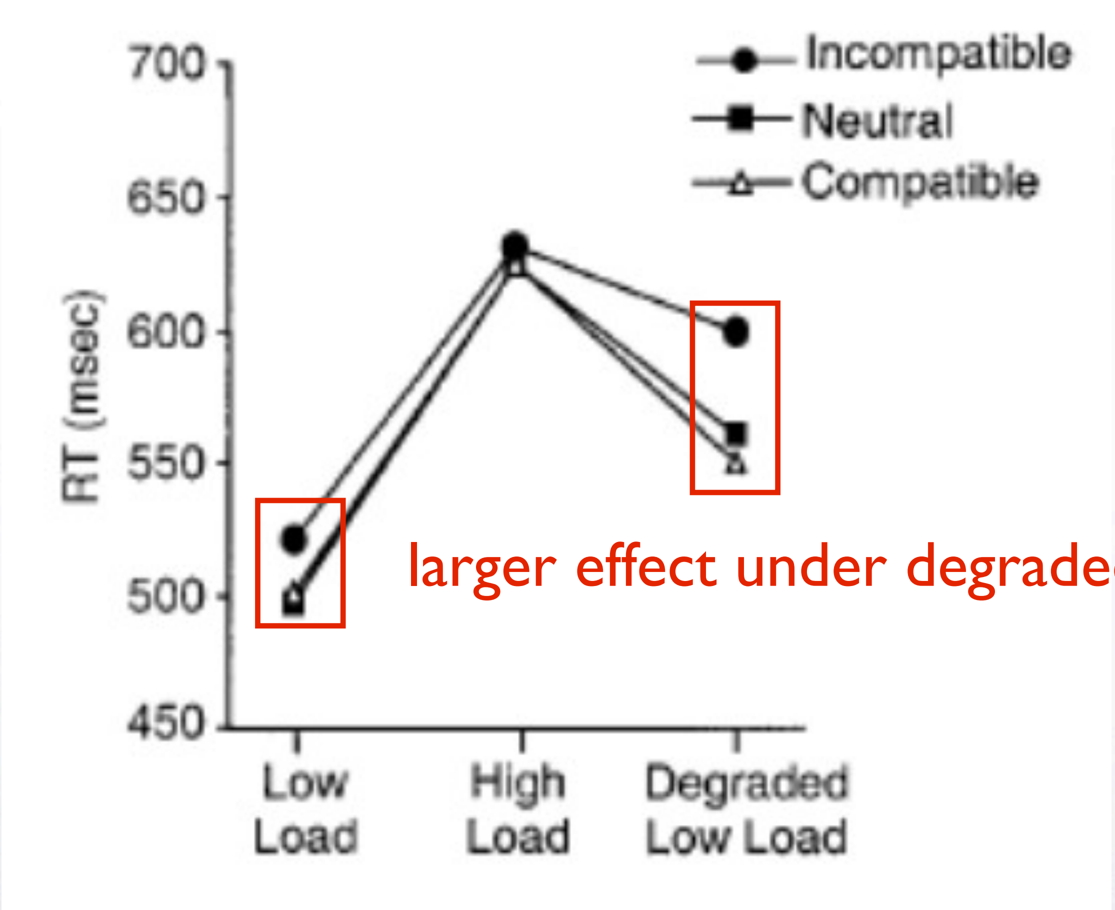
here, the difficulty of sensory processing is increased without changing attentional load

Attentional Load Experiments
Lavie and de Fockert (2003)



Results:
1) distractor had a significant effect only in the low-load case

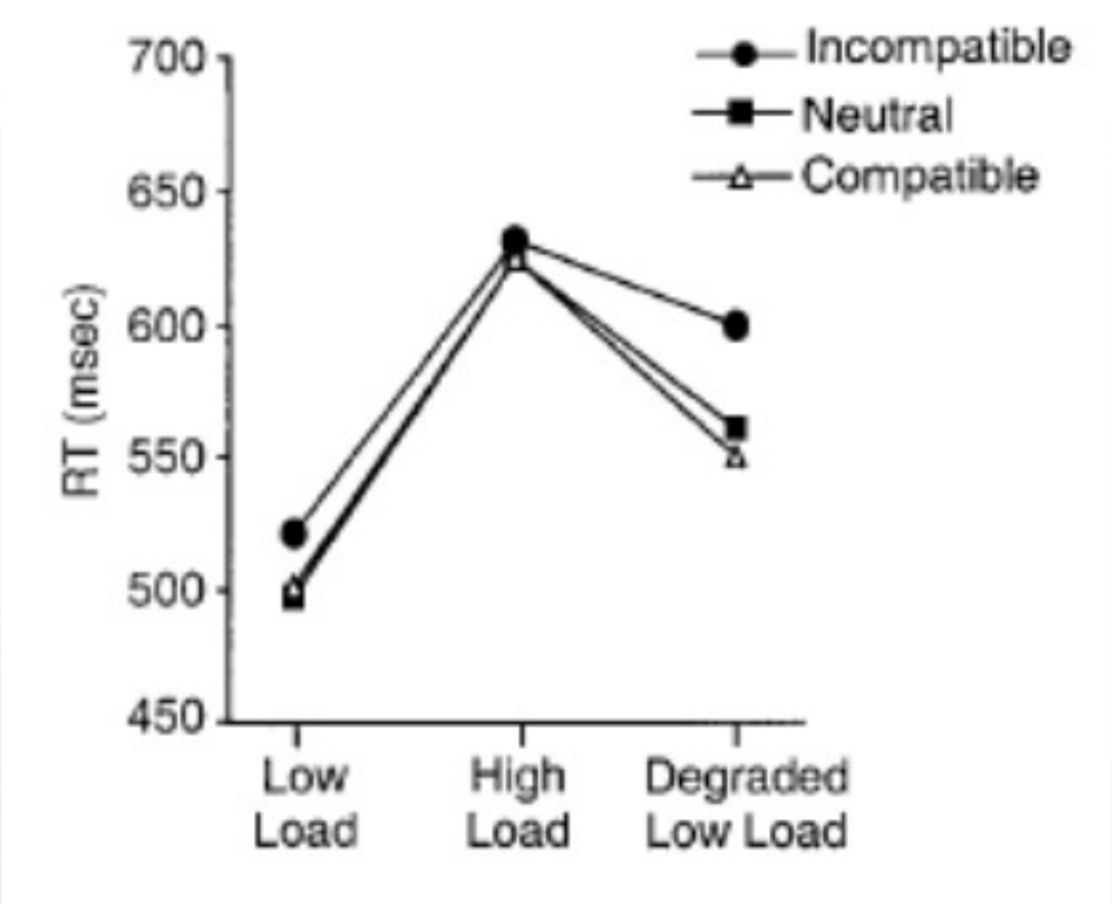
Attentional Load Experiments
Lavie and de Fockert (2003)



larger effect under degraded input

Results:
1) distractor had a significant effect only in the low-load case
2) greater distractor impact under degraded condition (even though less resources required than for high-load case)

Attentional Load Experiments
Lavie and de Fockert (2003)



Results:
1) distractor had a significant effect only in the low-load case
2) greater distractor impact under degraded condition (even though less resources required than for high-load case)
3) compatible distractors did not help, incompatible ones were harmful

Attentional Load Experiments
Lavie and de Fockert (2003)

Why does the distractor corrupt processing of the target in the easy, low-load case but not in the difficult, high-load case - even though less attentional resources are required for the former?

This paper
Peter Dayan (2009)

Why does the distractor corrupt processing of the target in the easy, low-load case but not in the difficult, high-load case - even though less attentional resources are required for the former?

Visual system has receptive fields of different sizes:
- smaller, spatially precise ones are confined to the target
- larger, spatially extended ones, include both target and distractor

High-load case: proximal stimuli add so much extra noise to large receptive fields, that only the smallest receptive fields contain useful information
Low-load case: large receptive fields also include the target, and so distractor will exert an influence

This paper
Peter Dayan (2009)

Why does the distractor corrupt processing of the target in the easy, low-load case but not in the difficult, high-load case - even though less attentional resources are required for the former?

Visual system has receptive fields of different sizes:
- smaller, spatially precise ones are confined to the target
- larger, spatially extended ones, include both target and distractor

High-load case: proximal stimuli add so much extra noise to large receptive fields, that only the smallest receptive fields contain useful information
Low-load case: large receptive fields also include the target, and so distractor will exert an influence

Since processing combines information from all receptive fields, this differential processing occurs due to inference and marginalization (which eliminates impact from useless/confusion units) without explicit attentional control required.

Attentional Load Experiments
Peter Dayan (2009), based on Eriksen task (1974)

n t n d

target

- could take on the value +1, -1
- subjects have to report its sign

degraded condition: +0.3, -0.3

Attentional Load Experiments
Peter Dayan (2009), based on Eriksen task (1974)

n t n d

non-target flankers

- **low-load case:** both have value 0
- **high-load case:** one has value -1, other has value +1

Attentional Load Experiments
Peter Dayan (2009), based on Eriksen task (1974)

n t n d

distractor

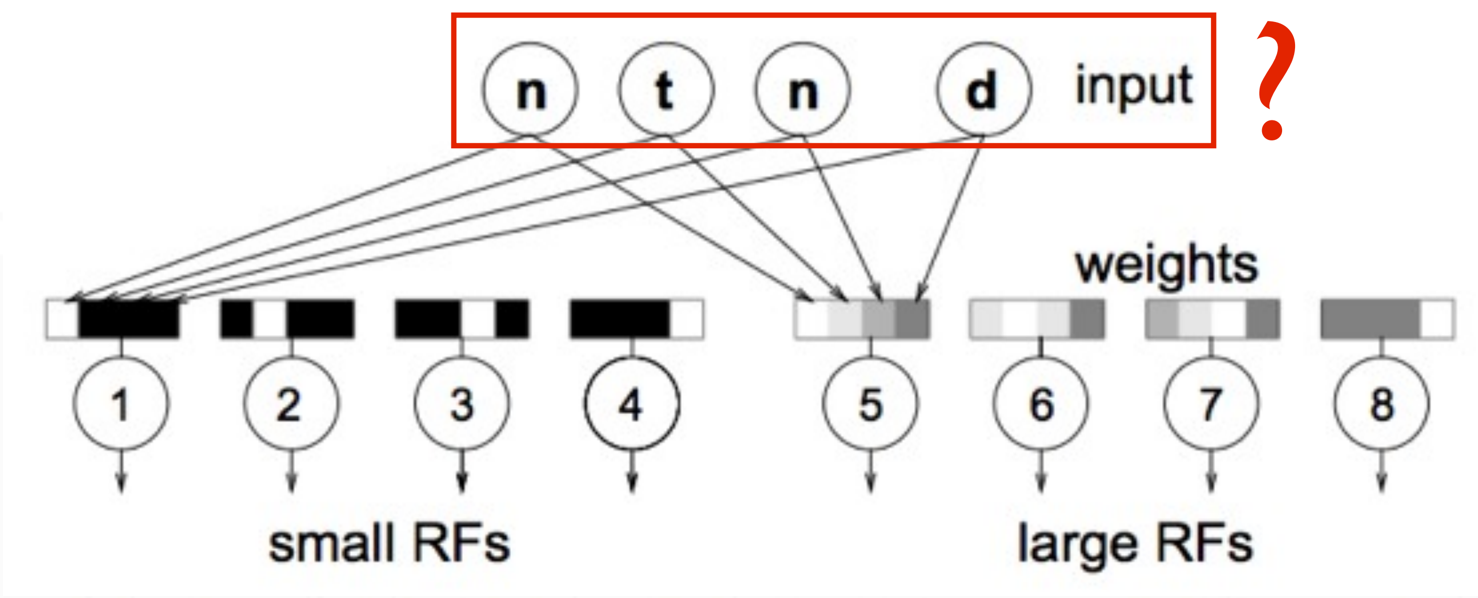
- **compatible:** same sign as target
- **incompatible:** different sign from target
- **neutral:** value 0

Attentional Load Experiments

Peter Dayan (2009), based on Eriksen task (1974)

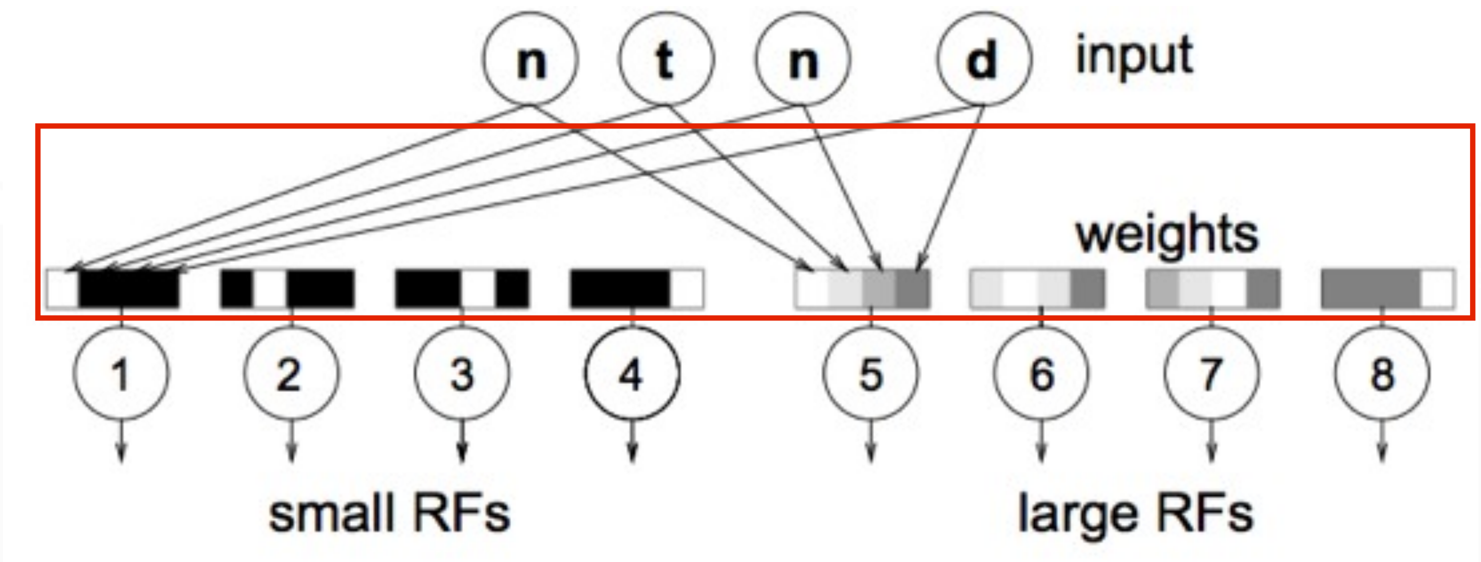
	high-load	low-load	degraded low-load
compatible	-1 +1 +1 +1	0 +1 0 +1	0 +0.3 0 +1
incompatible	-1 +1 +1 -1	0 +1 0 -1	0 +0.3 0 -1
neutral	-1 +1 +1 0	0 +1 0 0	0 +0.3 0 0

The Generative Model



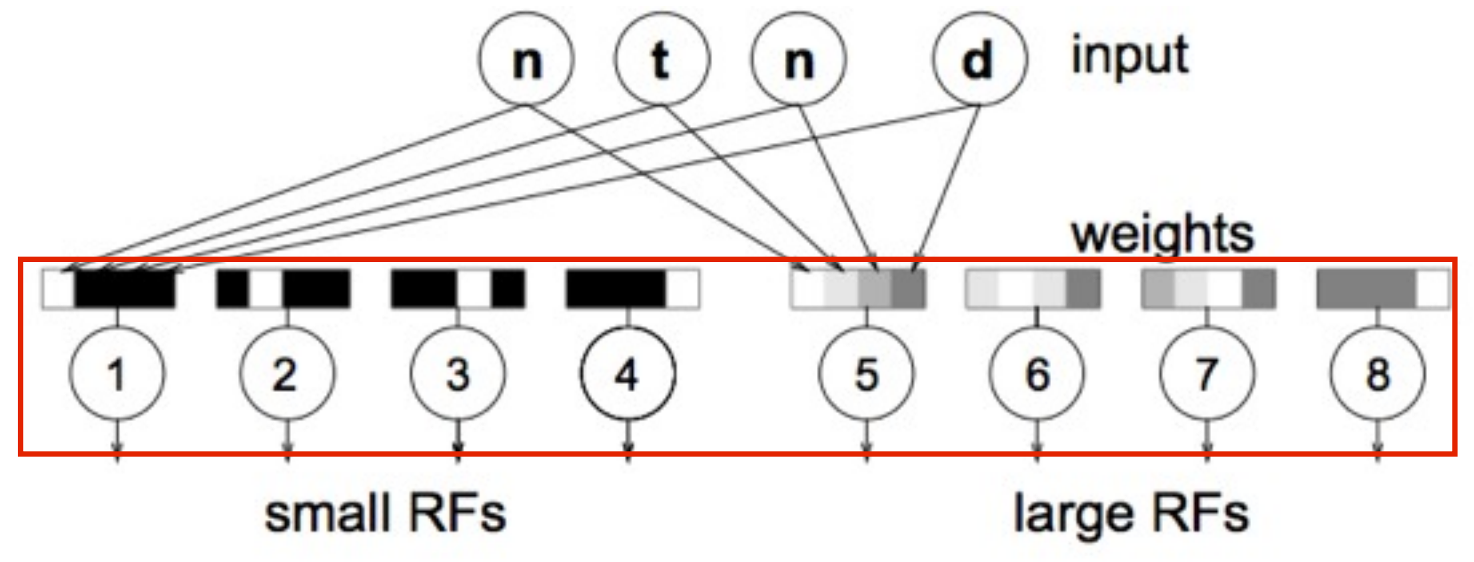
- hidden input
- input is mapped and mixed through various receptive fields
- result is noisy, and is used as data by a Bayesian recognition model
- Bayesian model calculates the posterior probability of the hidden settings | data
- marginalizing out all hidden settings apart from target, report sign of target

The Generative Model



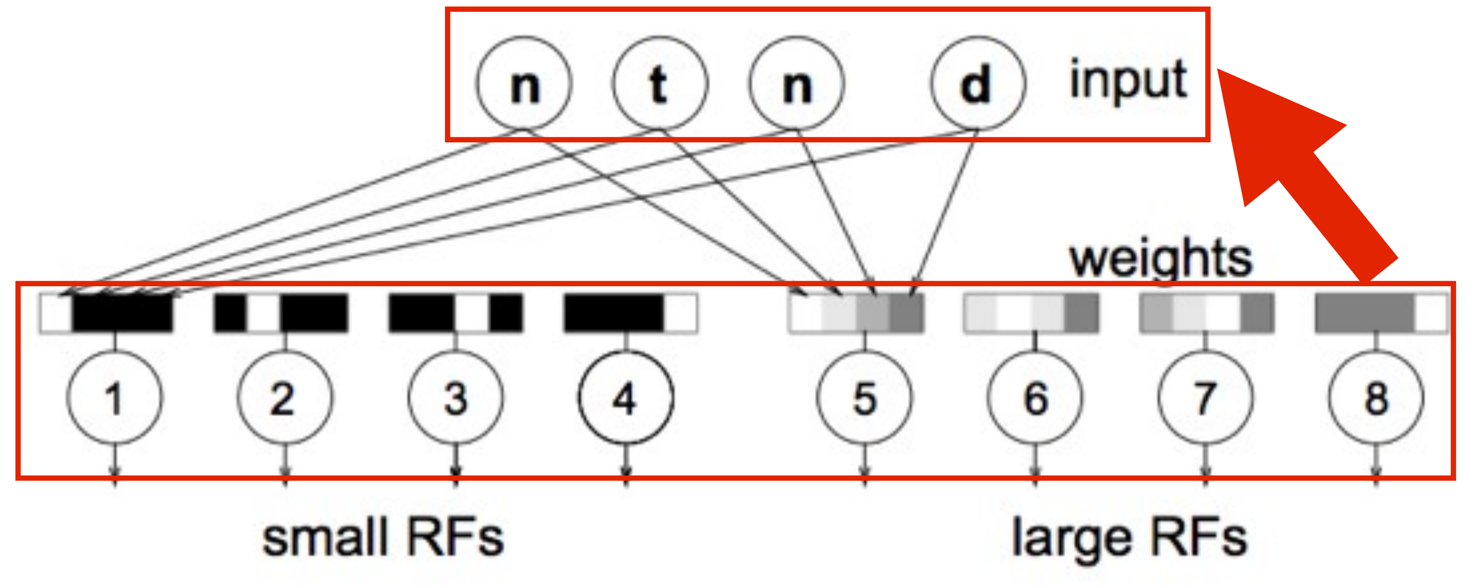
- hidden input
- input is mapped and mixed through various receptive fields
- result is noisy, and is used as data by a Bayesian recognition model
- Bayesian model calculates the posterior probability of the hidden settings | data
- marginalizing out all hidden settings apart from target, report sign of target

The Generative Model



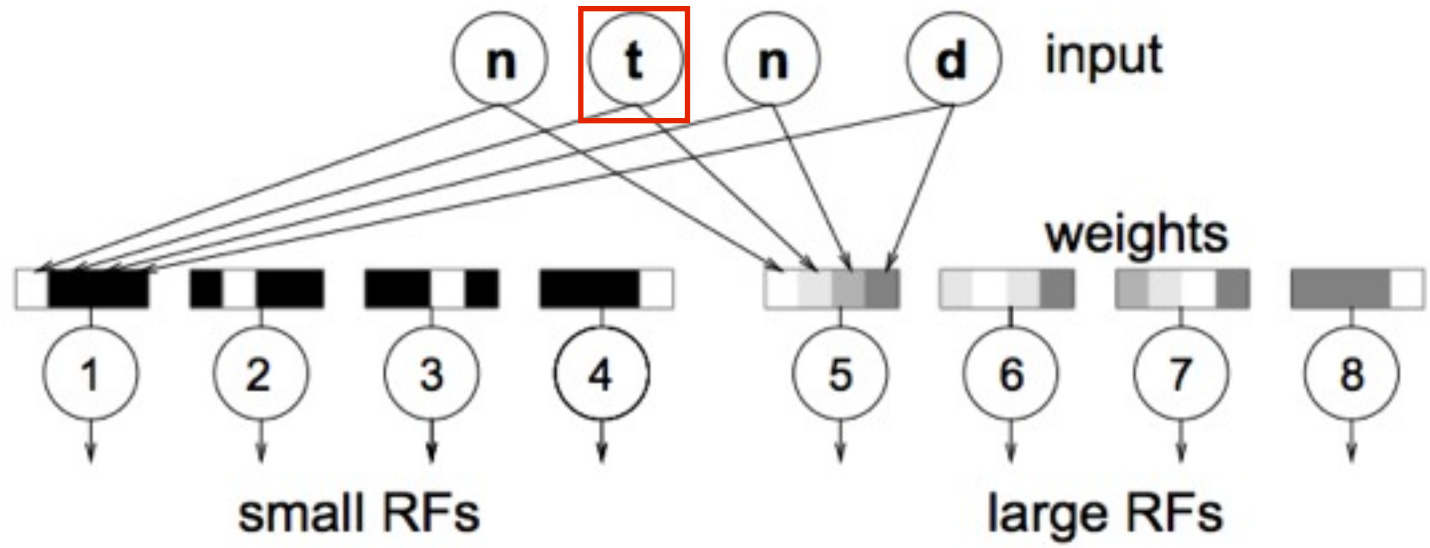
- hidden input
- input is mapped and mixed through various receptive fields
- result is noisy, and is used as data by a Bayesian recognition model
- Bayesian model calculates the posterior probability of the hidden settings | data
- marginalizing out all hidden settings apart from target, report sign of target

The Generative Model



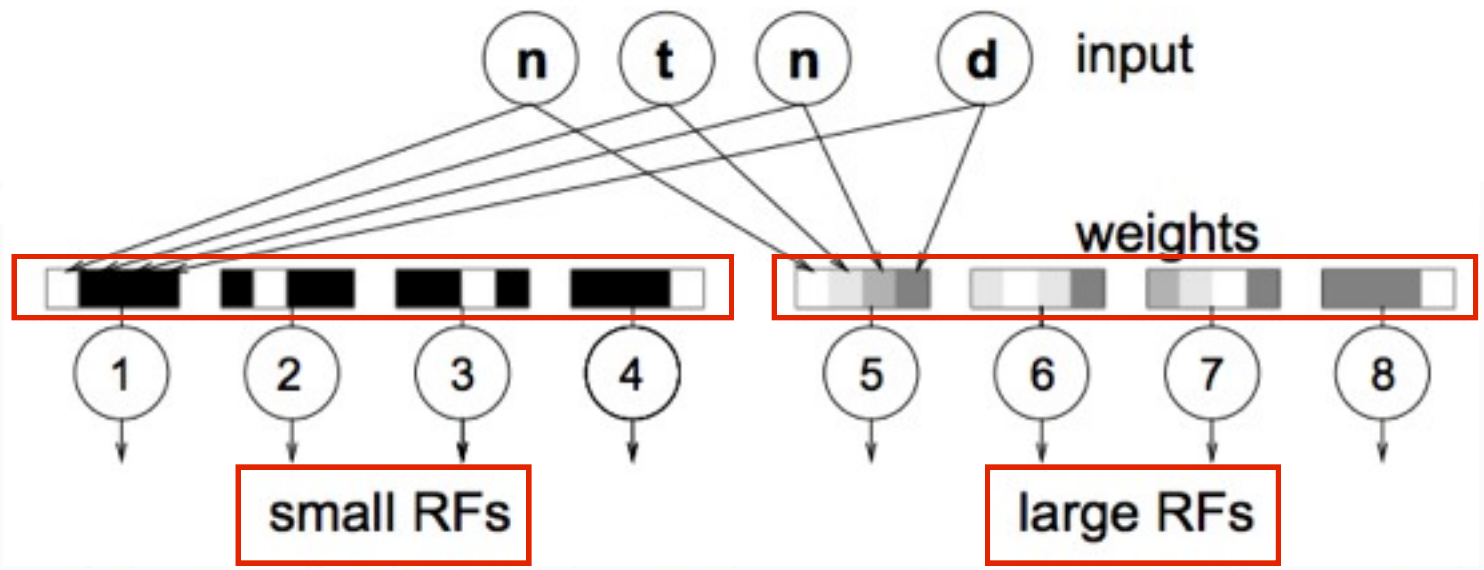
- hidden input
- input is mapped and mixed through various receptive fields
- result is noisy, and is used as data by a Bayesian recognition model
- Bayesian model calculates the posterior probability of the hidden settings | data
- marginalizing out all hidden settings apart from target, report sign of target

The Generative Model



- hidden input
- input is mapped and mixed through various receptive fields
- result is noisy, and is used as data by a Bayesian recognition model
- Bayesian model calculates the posterior probability of the hidden settings | data
- marginalizing out all hidden settings apart from target, report sign of target

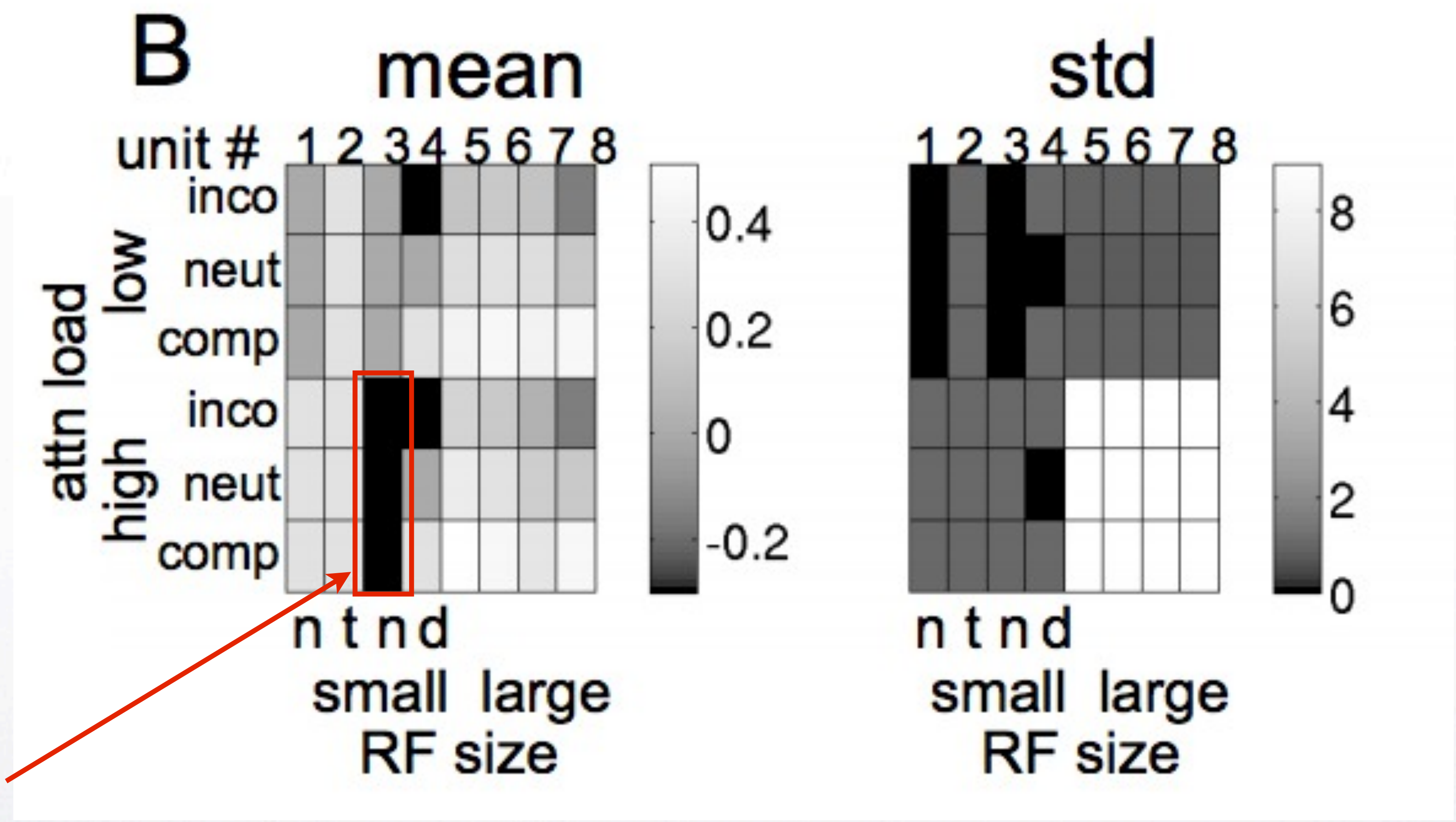
The Generative Model



- local to a single input character

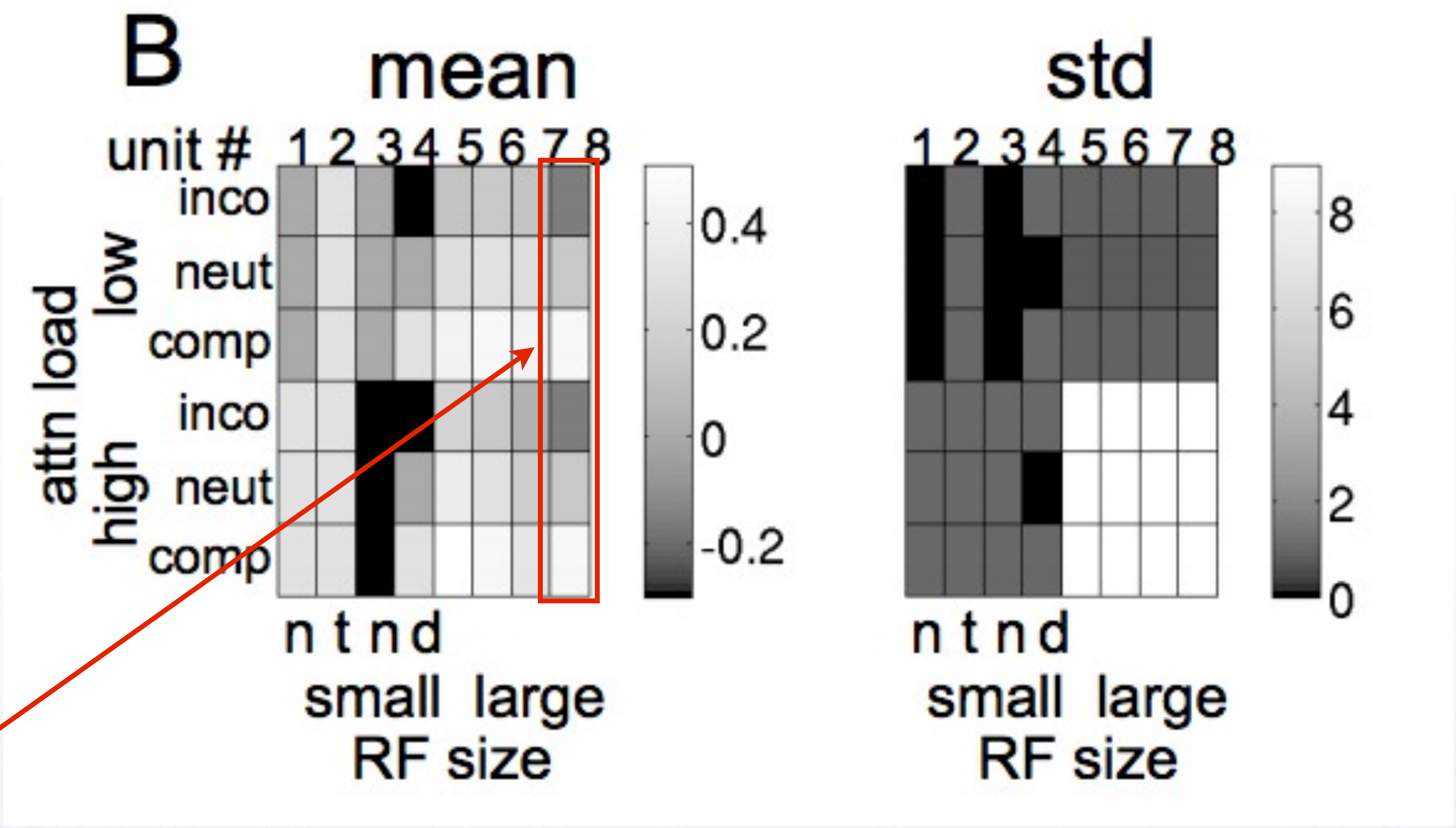
- include entire input
- due to distance, differing weights for non-targets and distractor

The Generative Model



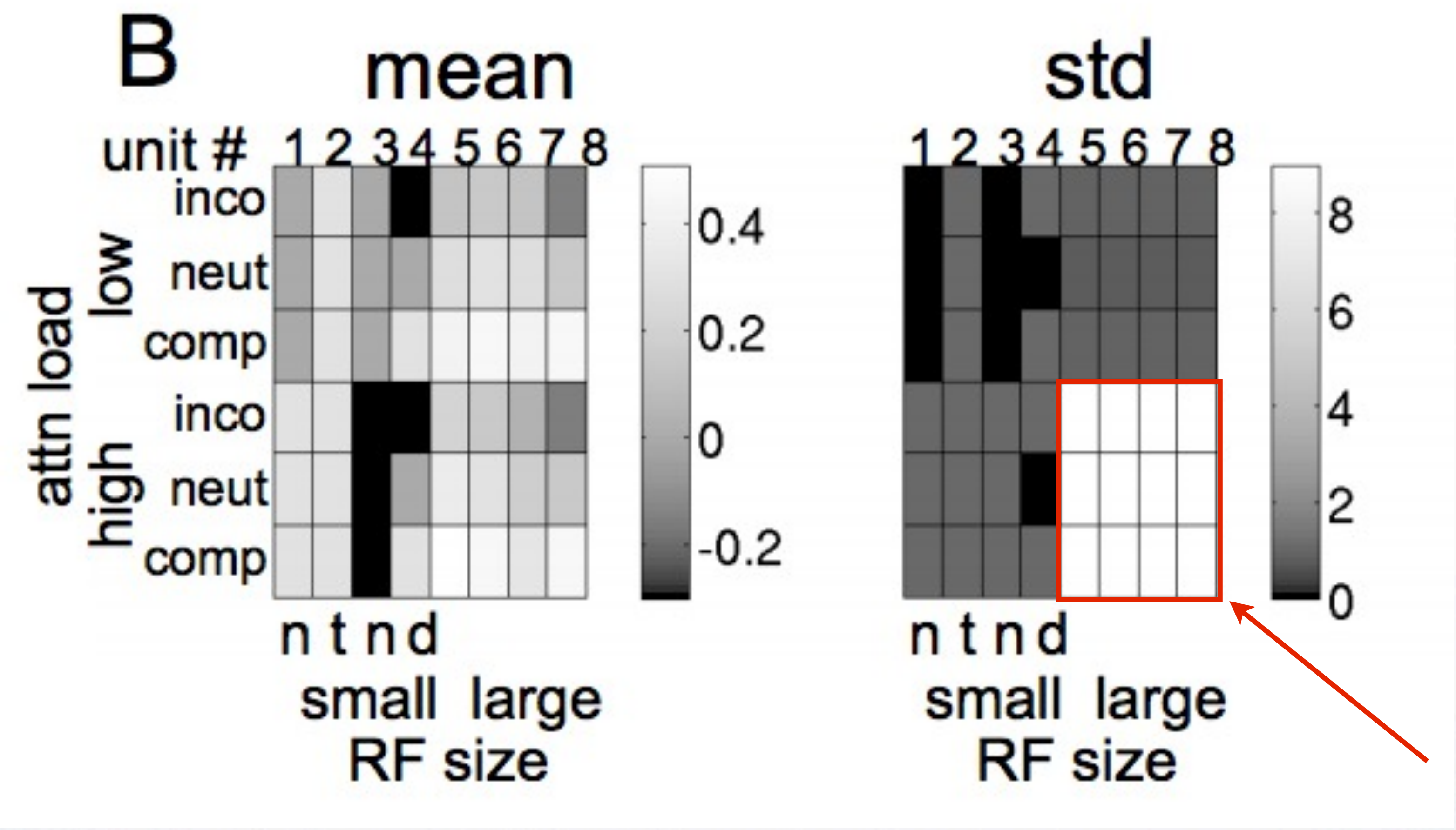
lack of bias from non-targets under high load for small RFs

The Generative Model



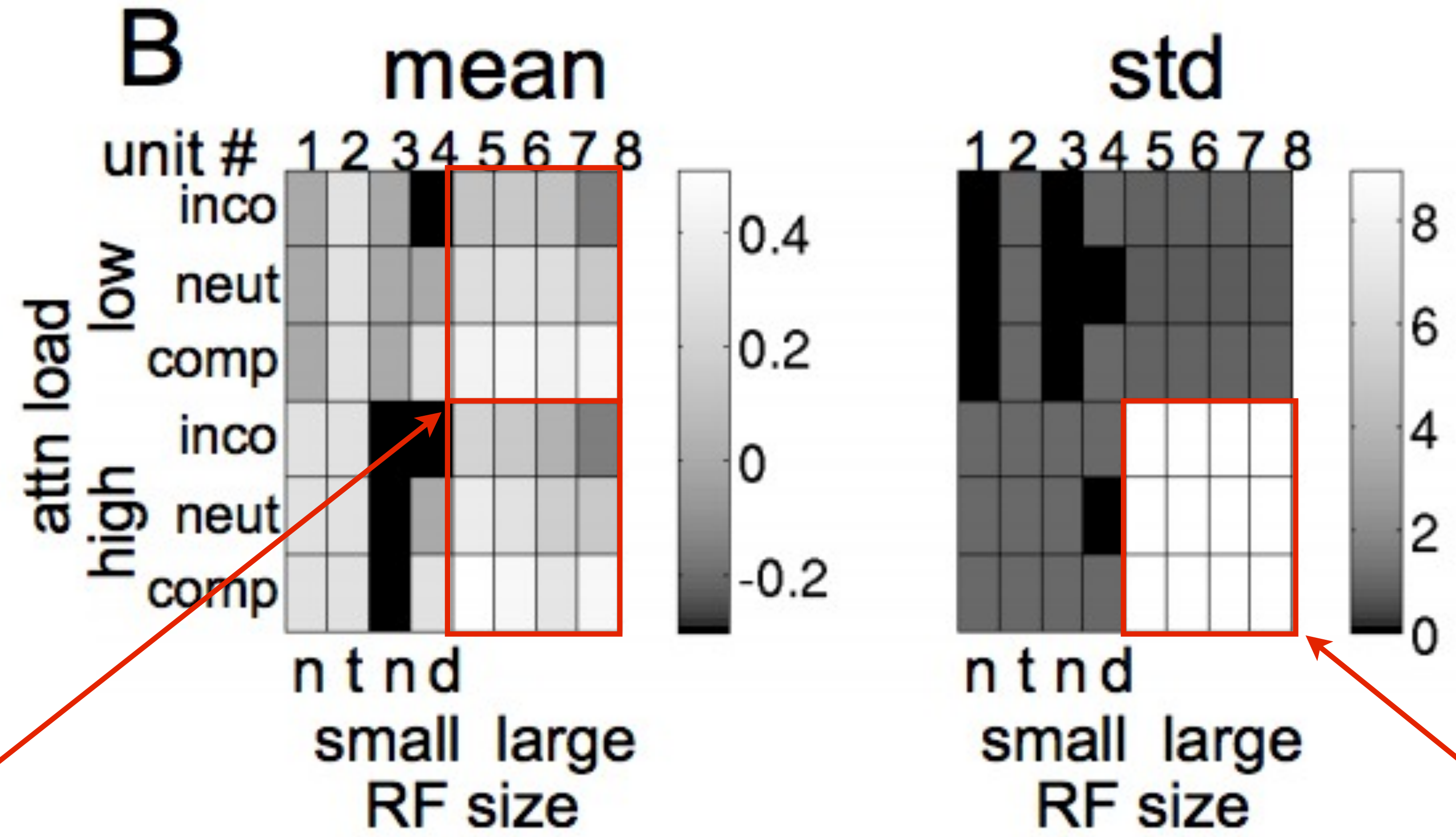
bias from
distractor
for large RFs

The Generative Model



modeling assumption: signal-dependent noise; captures uselessness of large RFs under high load

The Generative Model



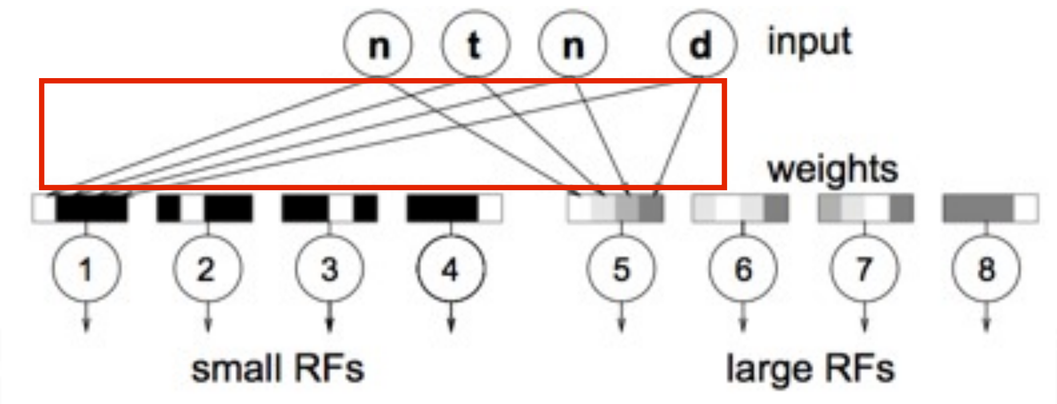
mean values of large RFs remain unaffected across different conditions

modeling assumption: signal-dependent noise; captures uselessness of large RFs under high load

The Generative Model

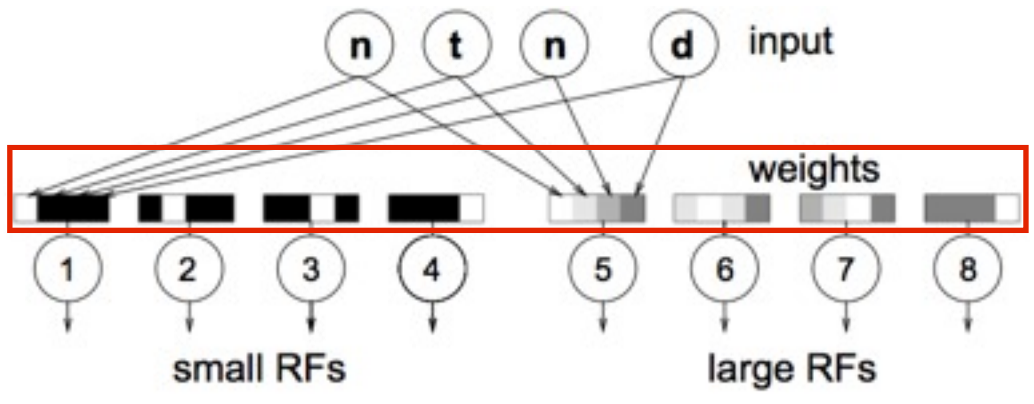
Model specified by:

- structures (units) and connections
- weights of connections
- mean and std values of units



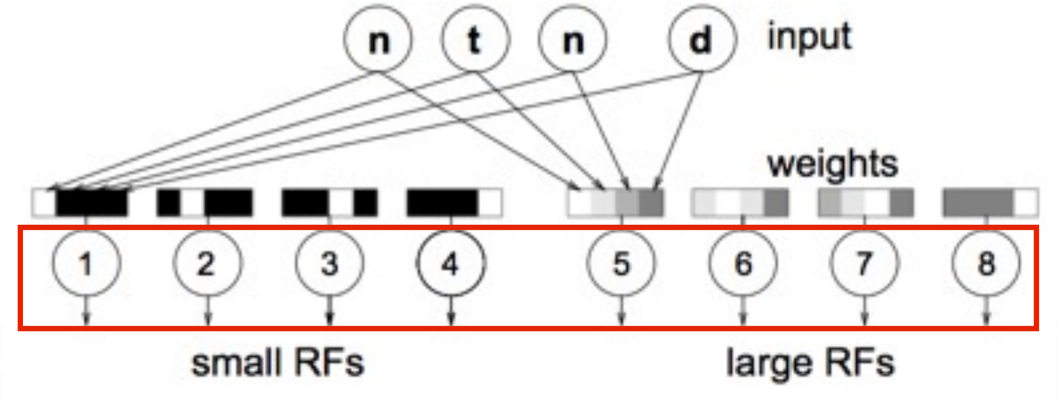
The Generative Model

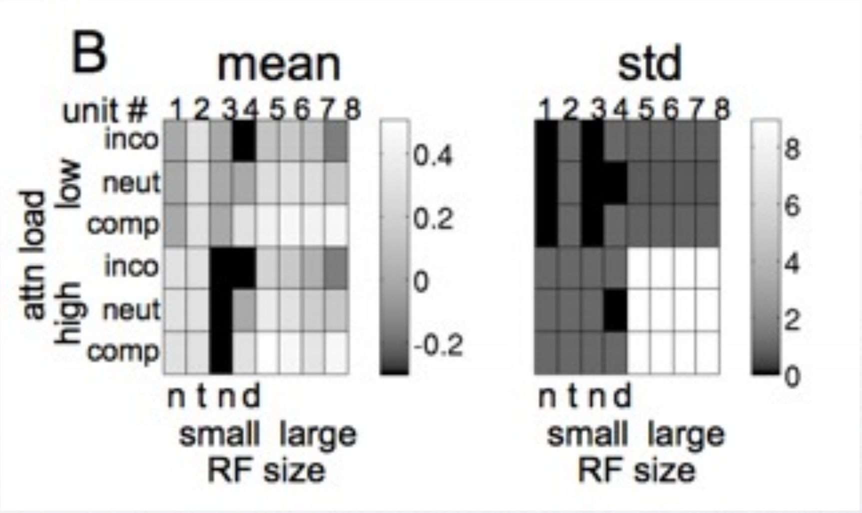
Model specified by:
- structures (units) and connections
- weights of connections
- mean and std values of units



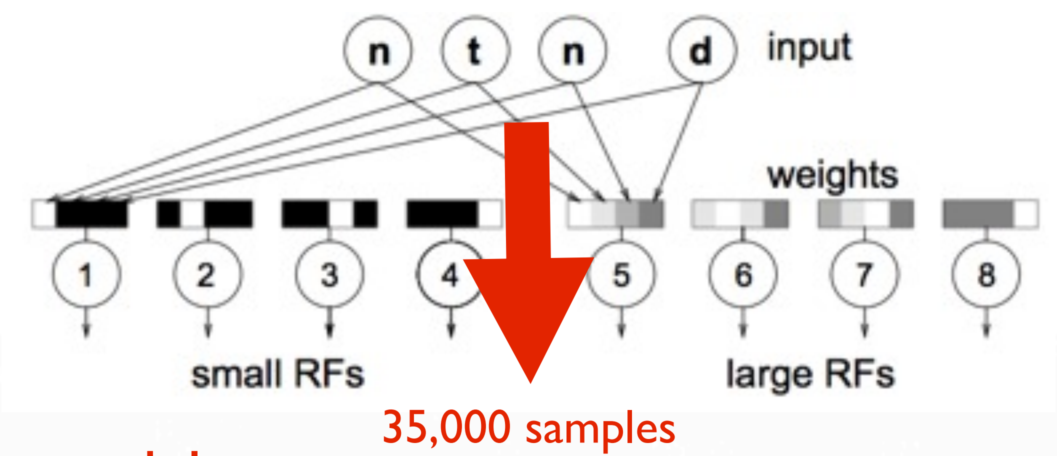
The Generative Model

Model specified by:
- structures (units) and connections
- weights of connections
- mean and std values of units





The Generative Model



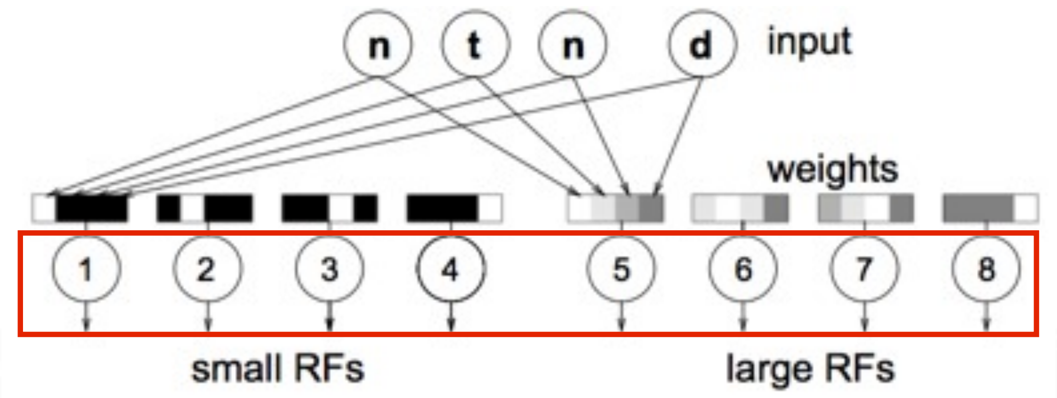
Model specified by:
- structures (units) and connections
- weights of connections
- mean and std values of units

- at each time step, a new sample is generated by model
- Gaussian noise corrupts each of the observed units

- recognition model accumulates evidence until > 0.9 confident about sign of target
- probability of 0.01 of stopping accumulation early and reporting sign with higher probability (to simulate making early, possibly erroneous responses)

The Generative Model

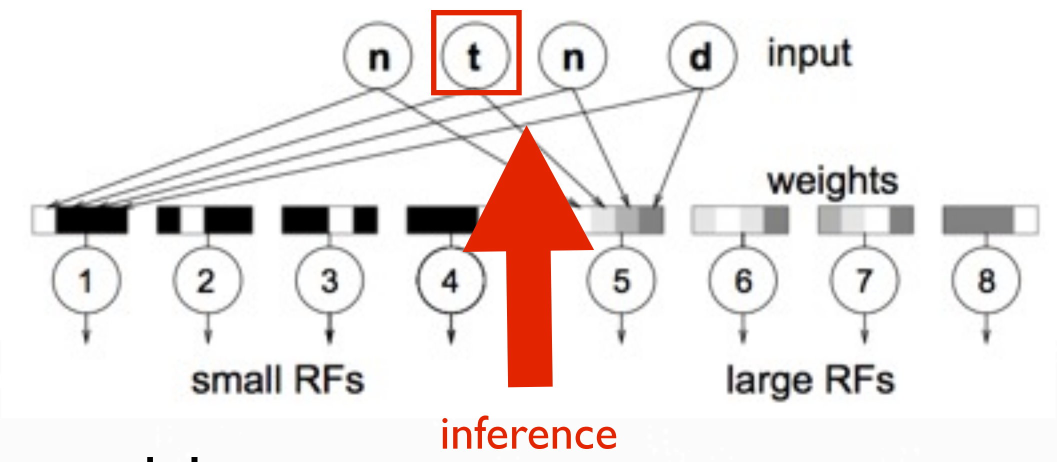
Model specified by:
- structures (units) and connections
- weights of connections
- mean and std values of units



- at each time step, a new sample is generated by model
- Gaussian noise corrupts each of the observed units

- recognition model accumulates evidence until > 0.9 confident about sign of target
- probability of 0.01 of stopping accumulation early and reporting sign with higher probability (to simulate making early, possibly erroneous responses)

The Generative Model



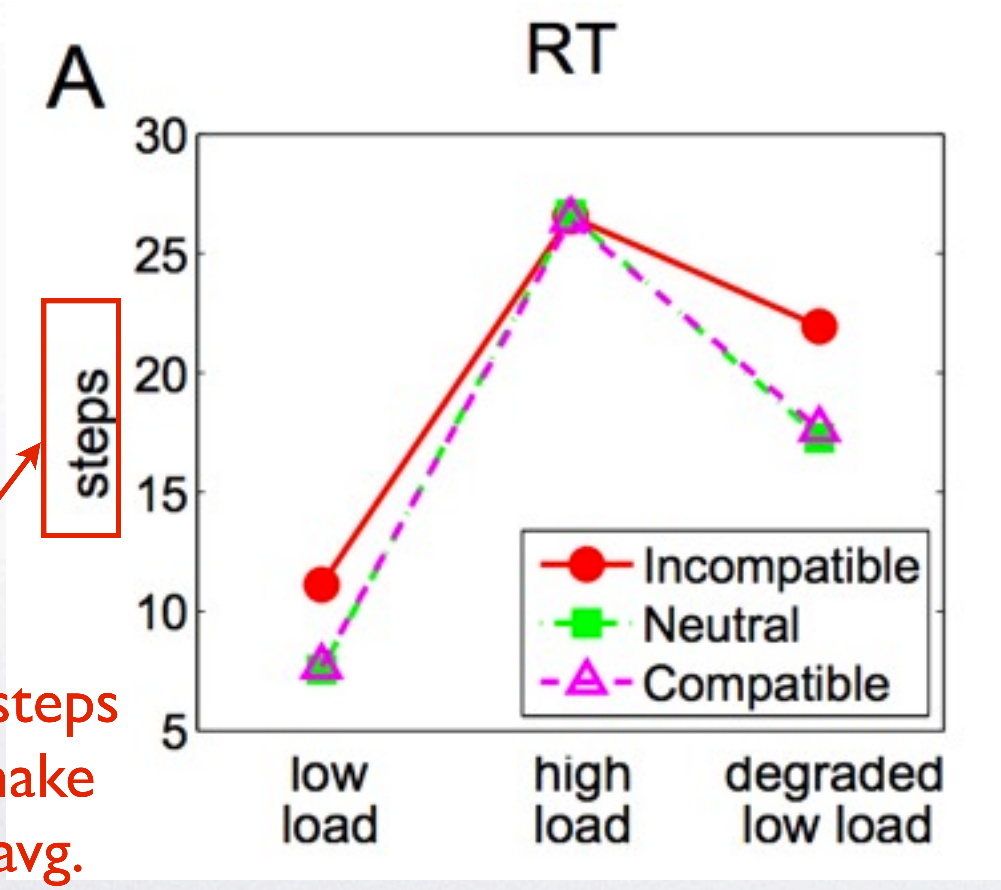
Model specified by:
- structures (units) and connections
- weights of connections
- mean and std values of units

- at each time step, a new sample is generated by model
- Gaussian noise corrupts each of the observed units

- recognition model accumulates evidence until > 0.9 confident about sign of target
- probability of 0.01 of stopping accumulation early and reporting sign with higher probability (to simulate making early, possibly erroneous responses)

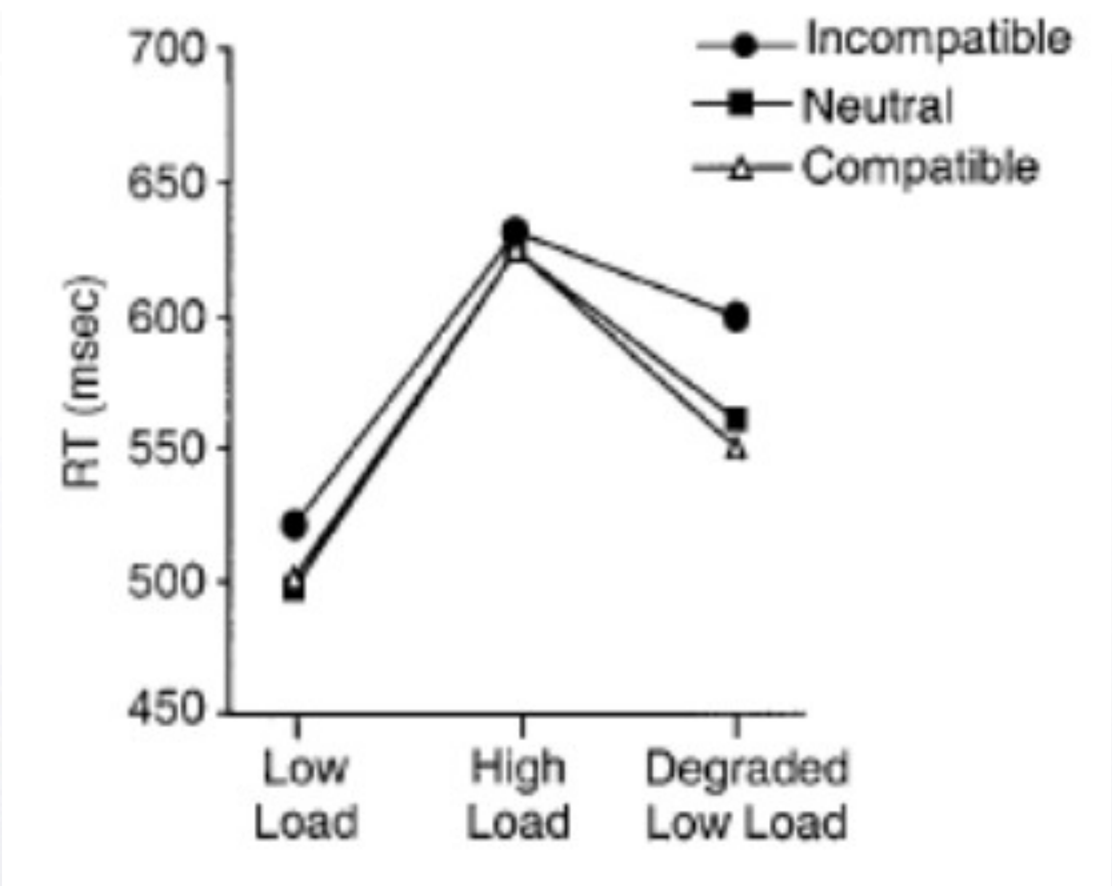
The Generative Model

computational prediction:

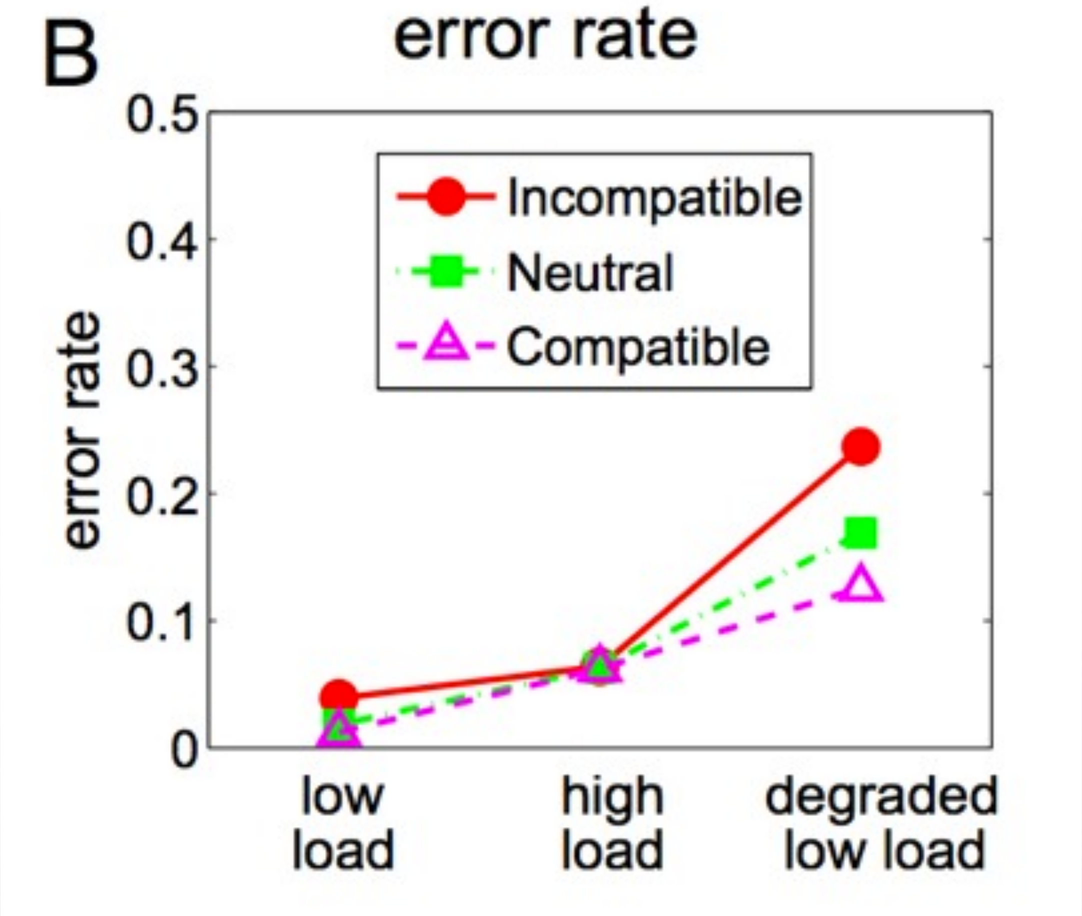


number of steps until can make decision (avg. over samples)

human performance:



The Generative Model



Errors are relatively rare, so computational data matches human data.

Pattern of RTs not accounted for by a tradeoff between speed and accuracy

What has the model been able to explain?

- under low-load, the lack of non-targets means that inputs based on large RFs are usually informative about the target, and thus play a role in inference
- in these cases, distractors also part of input, and thus affect RTs when incompatible
- under high-load, non-targets are closer to target and exert influence over noise corrupting the large RFs
- in these cases, the smaller RFs are relied upon instead, and are not affected by the distractor
- when input is degraded, information from large RFs is used to make inferences about the target; therefore distractor has a large influence on RTs in this case
- compatible distractor is less helpful than incompatible is harmful, due to (a) ceiling effects and (b) signal-dependent noise that reduces large RF informativeness

0 +1 0 -1

large RFs include distractor

What has the model been able to explain?

- under low-load, the lack of non-targets means that inputs based on large RFs are usually informative about the target, and thus play a role in inference
- in these cases, distractors also part of input, and thus affect RTs when incompatible
- under high-load, non-targets are closer to target and exert influence over noise corrupting the large RFs
- in these cases, the smaller RFs are relied upon instead, and are not affected by the distractor
- when input is degraded, information from large RFs is used to make inferences about the target; therefore distractor has a large influence on RTs in this case
- compatible distractor is less helpful than incompatible is harmful, due to (a) ceiling effects and (b) signal-dependent noise that reduces large RF informativeness

-| +| +| -|

small RFs mostly relied upon

What has the model been able to explain?

- under low-load, the lack of non-targets means that inputs based on large RFs are usually informative about the target, and thus play a role in inference
- in these cases, distractors also part of input, and thus affect RTs when incompatible
- under high-load, non-targets are closer to target and exert influence over noise corrupting the large RFs
- in these cases, the smaller RFs are relied upon instead, and are not affected by the distractor
- when input is degraded, information from large RFs is used to make inferences about the target; therefore distractor has a large influence on RTs in this case
- compatible distractor is less helpful than incompatible is harmful, due to (a) ceiling effects and (b) signal-dependent noise that reduces large RF informativeness

0 +0.3 0 -1

large RFs required

What has the model been able to explain?

- under low-load, the lack of non-targets means that inputs based on large RFs are usually informative about the target, and thus play a role in inference
- in these cases, distractors also part of input, and thus affect RTs when incompatible
- under high-load, non-targets are closer to target and exert influence over noise corrupting the large RFs
- in these cases, the smaller RFs are relied upon instead, and are not affected by the distractor
- when input is degraded, information from large RFs is used to make inferences about the target; therefore distractor has a large influence on RTs in this case
- compatible distractor is less helpful than incompatible is harmful, due to (a) ceiling effects and (b) signal-dependent noise that reduces large RF informativeness

0	+1	0	+1

0	+1	0	-1

Attentional Load

Attentional load hypothesis:

- when little attention is required to solve a set task, inputs associated with distractor stimuli "leak through", and cause disruption
- when the task is difficult, attention is totally occupied, leaving nothing left over (to attend to distractors)

Attenuation theory:

- can replicate same data, given a different underlying idea
- version of attenuation theory
- RF units that are not very informative (e.g. provide noisy information), are attenuated, and do not have as large of an influence on final decision

Future Work

To extend experimental circumstances to match Lavie and de Fockert (2003) task:
- exact location of target in stimulus array unknown
- more complex collection of letter-based RFs
- confusion matrix associated with perceptual similarities of letters

Multiple sizes of RFs should still be sufficient to explain most of observations.