UNIVERSITY OF
OXFORD

Why we are all bad scientists?

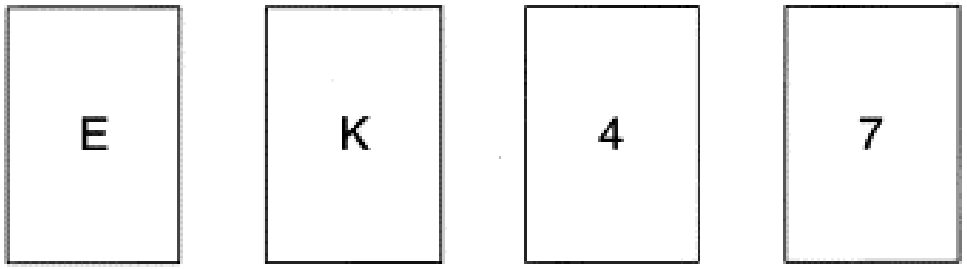
RL, confirmation bias, and lots more...

Student Juan del Ojo Balaguer

Supervisor Dr Christopher Summerfield

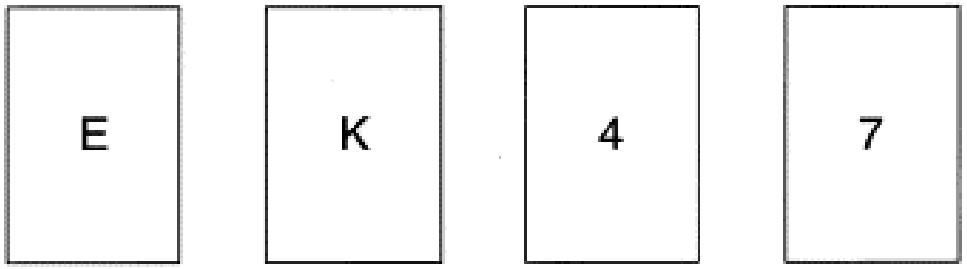
Wason Rule Discovery Test

Conditional Rule: "If there is a vowel on one side of the card, then there is an even number on the other side of the card."



Wason Rule Discovery Test

Conditional Rule: "If there is a vowel on one side of the card, then there is an even number on the other side of the card."



We should learn as much from confirmatory than from disconfirmatory evidence!

"`Vowel → Even`", then "`~Even → ~Vowel`" !

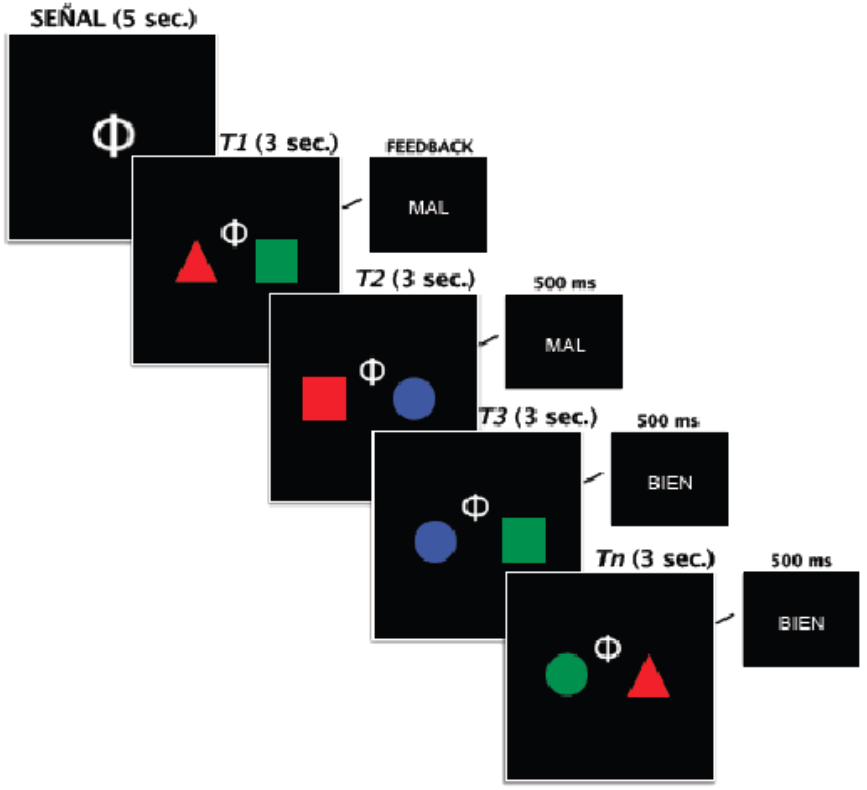
Confirmation Bias

So, three types of confirmation bias:

- Biased search for information
  - **e.g.** making experiments that confirm our hypothesis
- Biased interpretation
  - **e.g.** understanding ambiguous results as if they were completely supporting our hypothesis
- Biased memory
  - **e.g.** not paying attention to results that deny or reject our hypothesis

Intuitively, we may even believe that we (usually) learn from surprising information. So we should mainly learn from disconfirmatory information!

The task



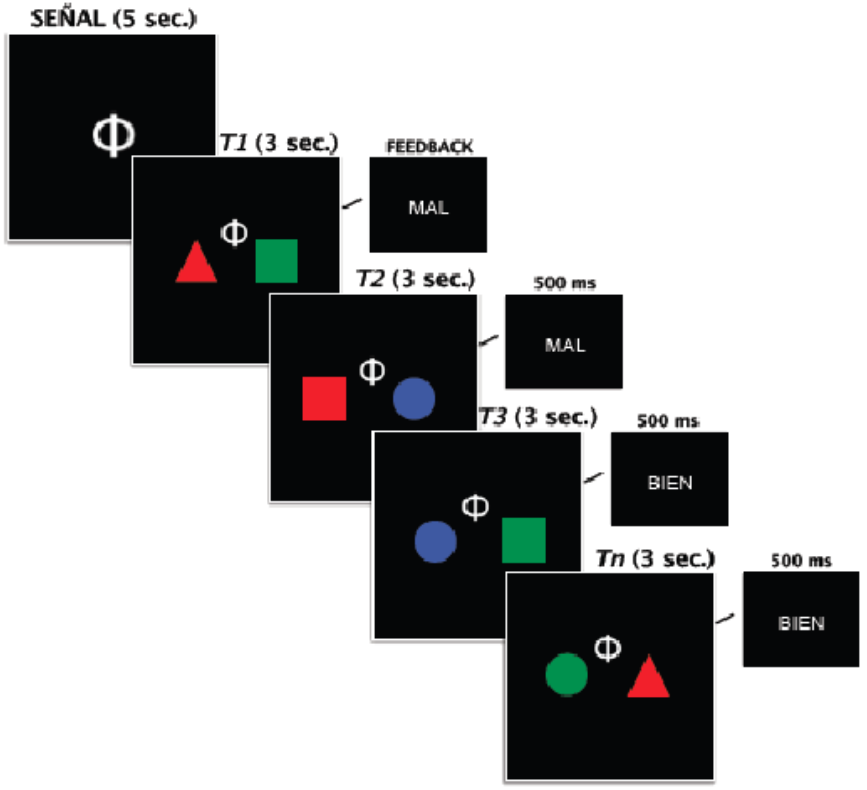
(Almost) not a single squircle!

Inductive rule task

Two dimensions:
- Color / Shape

Three features:
- Triangle / Square / Circle
- Red / Green / Blue

The task



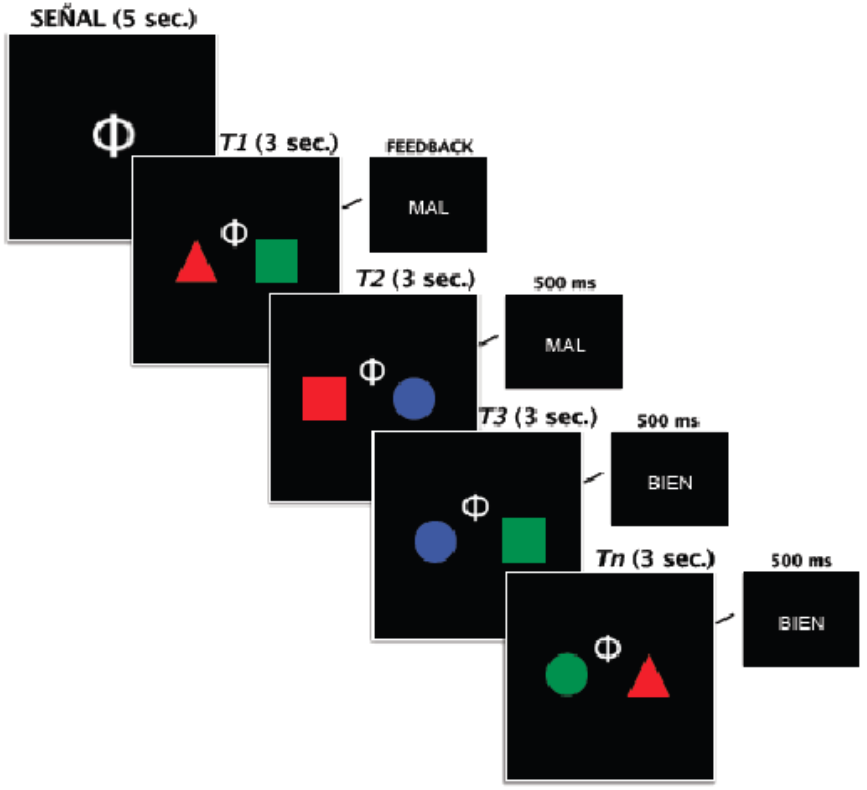
Two disjunctive simple rules

Example:
- Left side: BLUE
- Right side: TRIANGLE

Target:
BLUE (left)
or
TRIANGLE (right)

Non-Target:
otherwise

The task

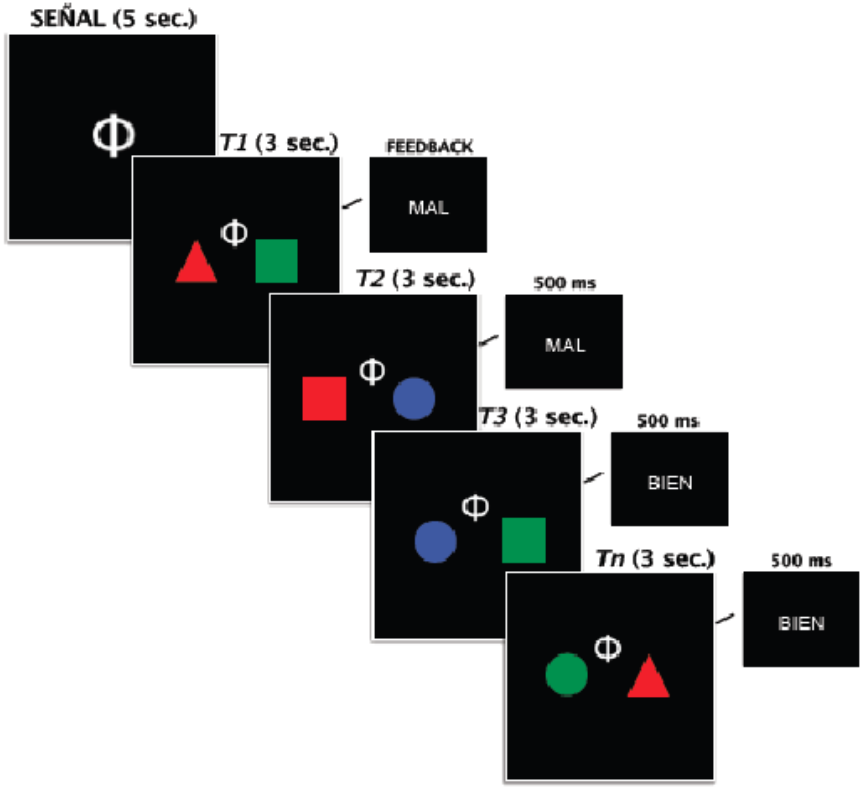


Cues give information about the relevant dimensions on each side.

Two conditions:

- Familiar cues (easy)
- Novel cues (crazy)

The task



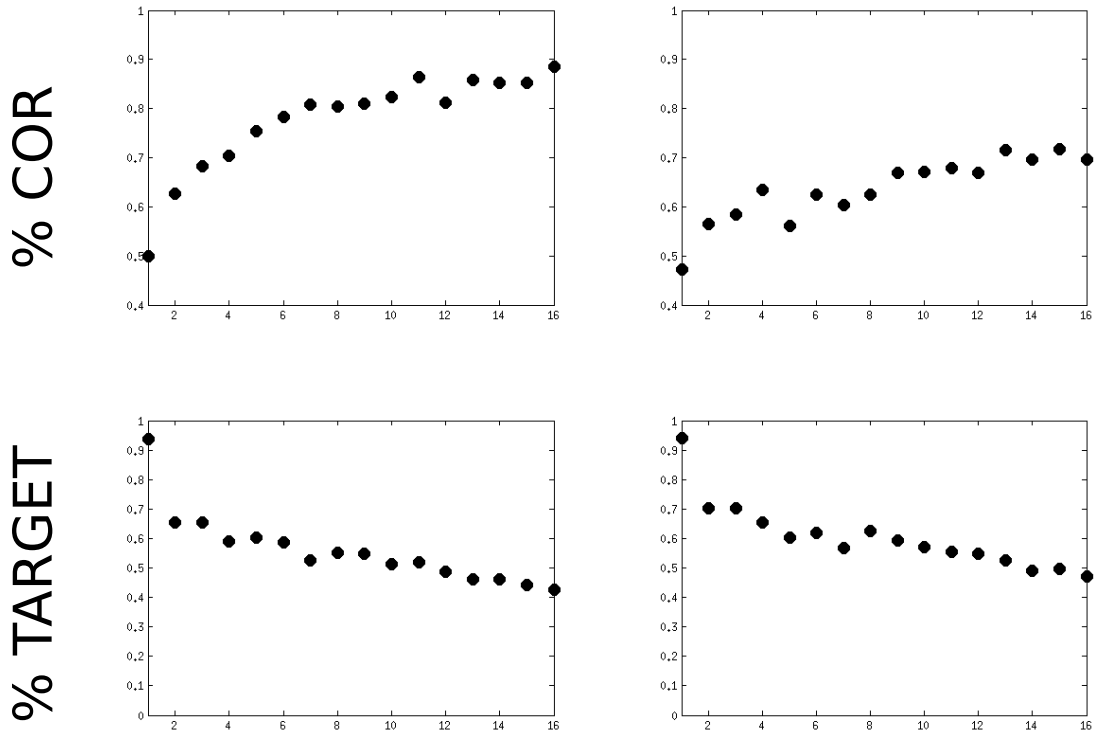
Research question:

Biased learning between

Target trials
and
Non-Target trials

→ is not really a confirmation bias.

Human Results



Models

“Confirmer” model

→ buffer with one number for each (side/dim/feat)
→ we say T if the value is greater than 0 for any side

[hack] = in the first trial we increment the features of the the stimulus … so we always start choosing T.

→ we increment values of each feature when is T
→ we reset values of each feature when is N

Red	Green	Blue	△	□	○
Red	Green	Blue	△	□	○

Models

“Disconfirmer” model

→ buffer with one number for each (side/dim/feat)
→ we say T if the value is 0 for any side

→ we decrement values of each feature when is N
→ we reset values of each feature when is T

Red	Green	Blue	△	□	○
Red	Green	Blue	△	□	○

Models

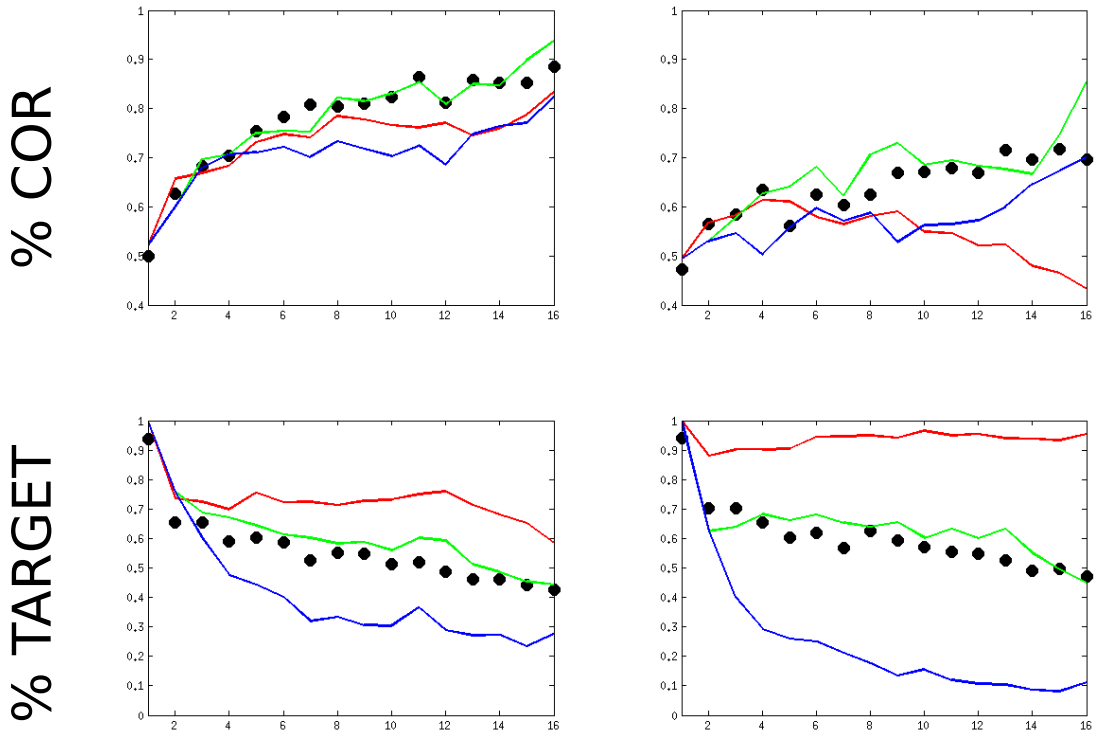
“Hybrid” (RL) model

→ buffer with one number for each (side/dim/feat)
→ we say T if the value is the max for any side

→ we increment values of each feature when is T
decrement when is N

Red	Green	Blue	△	□	○
Red	Green	Blue	△	□	○

Models



Confirmer

Hybrid

Disconfirmer

Models

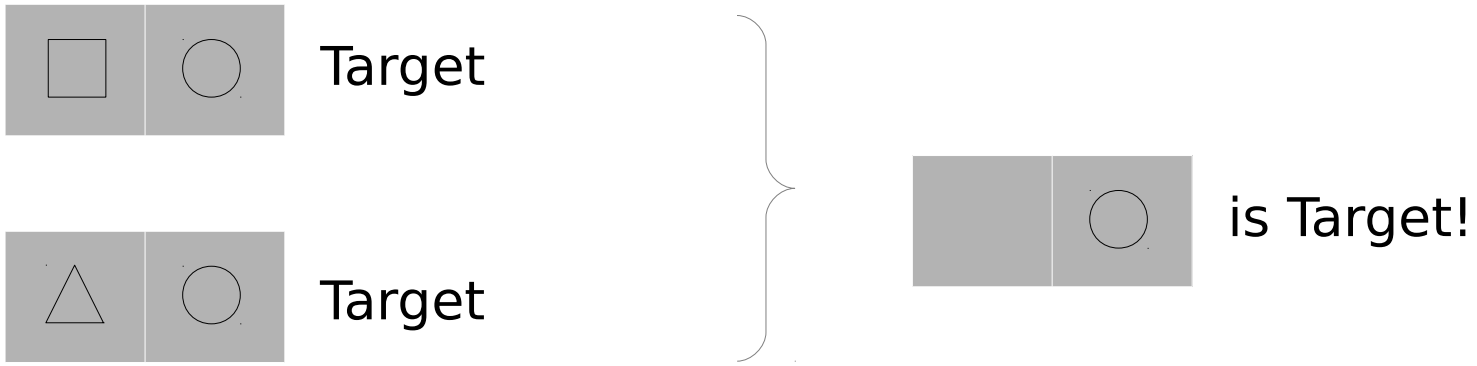
“Hypothesis” model

Thoughts:

Models

“Hypothesis” model

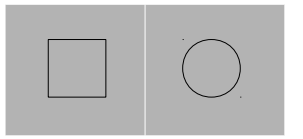
Thoughts:



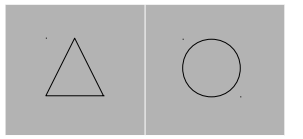
Models

“Hypothesis” model

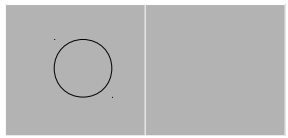
Thoughts:



Non-Target



Non-Target



is Target!

Models

“Hypothesis” model

Conclusions:

- we can learn deterministically by combining trials
- we may learn the same from T and NT trials
- if we understand how the task works, we may use logic to over-perform probabilistic (RL) strategies.

but!

- we need to combine trials in an efficient way
- the third combination is the most expensive in memory.

→ we can ignore the third combination

Models

“Hypothesis” model

1\ We use a variable to remember a Target trial.
→ is the 'hypothesis'

2\ We wait until we have another trial that partially matches our hypothesis
→ we deduce the rule
→ we store this into a second variable ('knowledge')

3\ Using the same logic, plus our new knowledge,
we can work out what is the rule for the other side!

Models

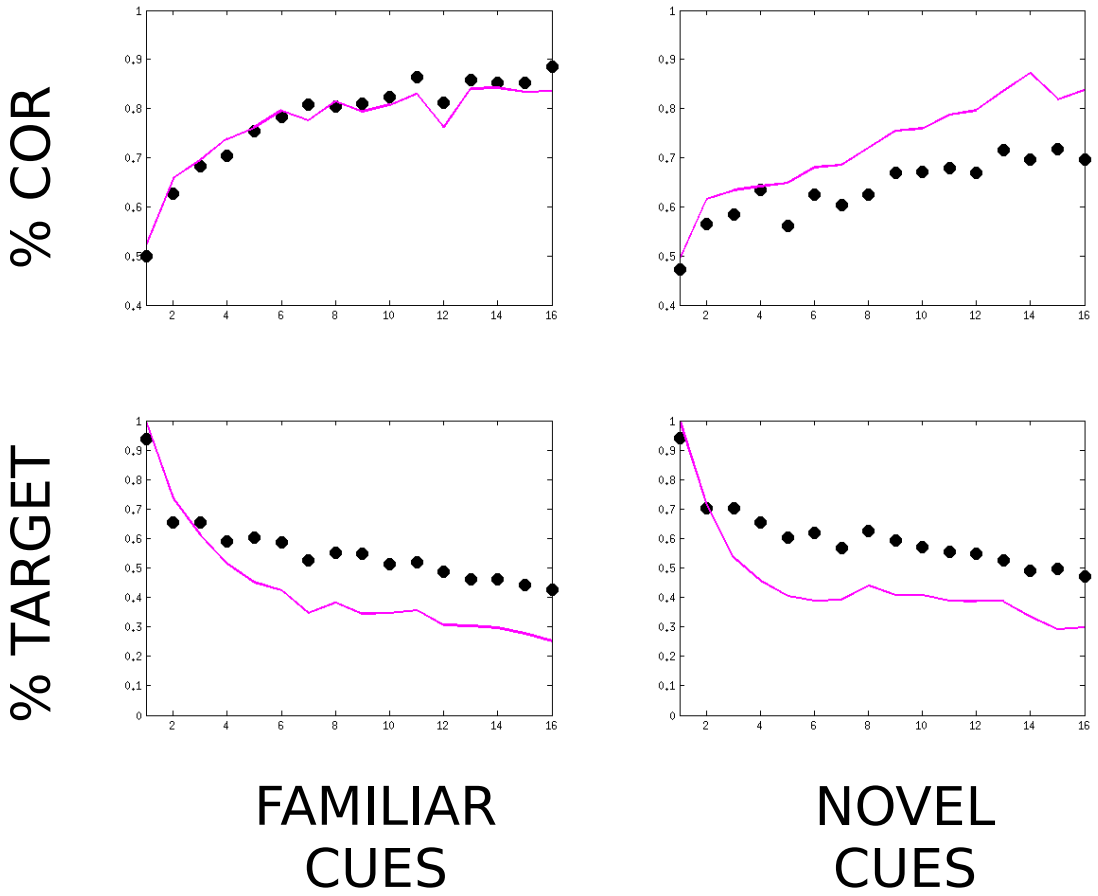
“Hypothesis” model

Interesting explanation for human results:
- They say Target on first trials
- When they have no information
→ Because they are waiting for a T trial!

Interesting bias in the model
Order of trials should not be important.
Model show a bias:
- learns from “T then NT”
- doesn't learn from “NT then T”

Generalization for novel cues (too long to explain)

Models



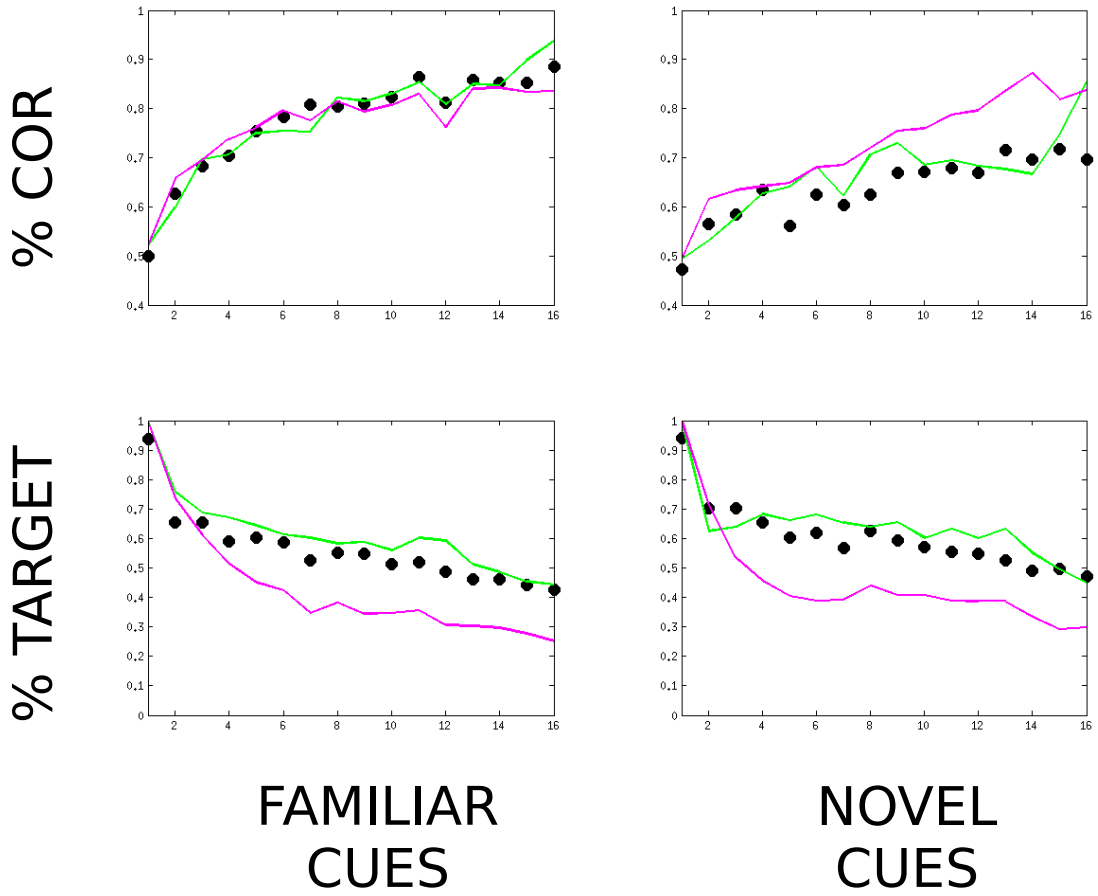
Comparison between models

Humans

Hybrid (RL)

Hypothesis

Comparison between models



Comparison between models

ERRORS

Flag fitting!

Hypothesis model

- Familiar Cue Confident Error = .1132
- Familiar Cue Not-confident Error = .2924
- Novel Cue Confident Error = .2296
- Novel Cue Not-confident Error = .3601

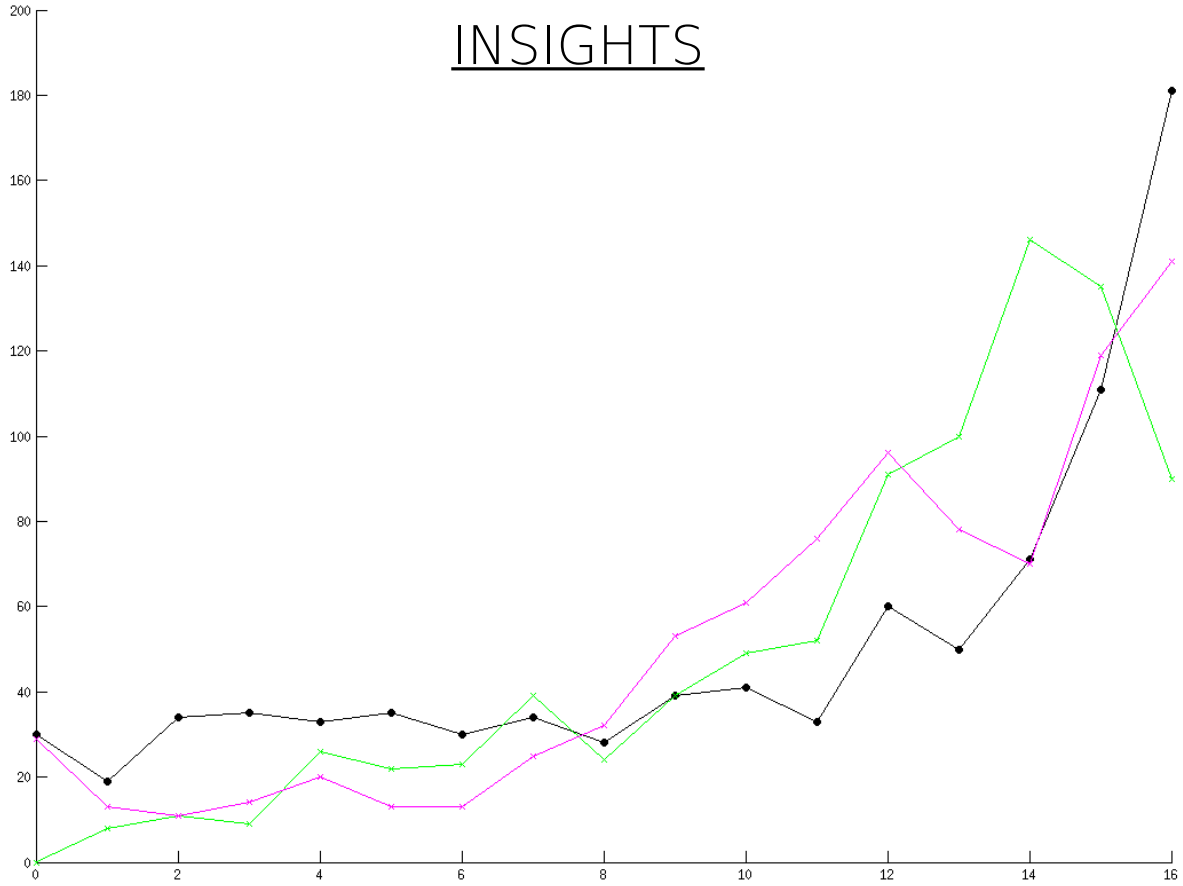
Comparison between models

ERRORS

Hybrid: Mean(abs(error)) = 0.2990

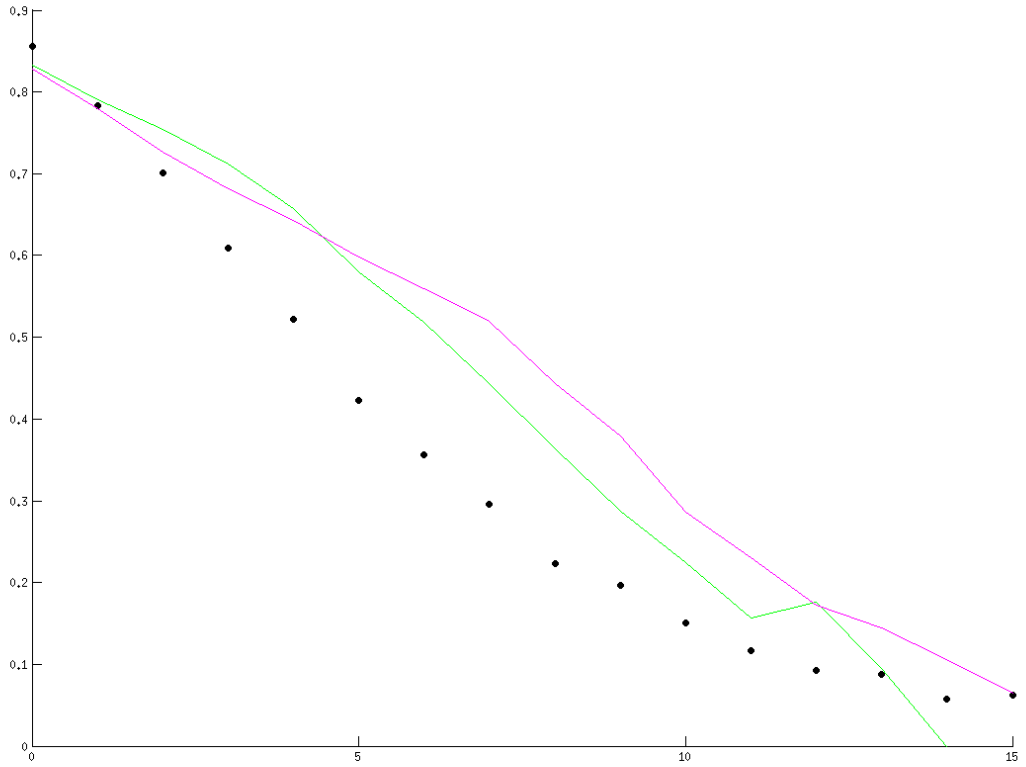
Hypothesis: Mean(abs(error)) = 0.2966

Comparison between models



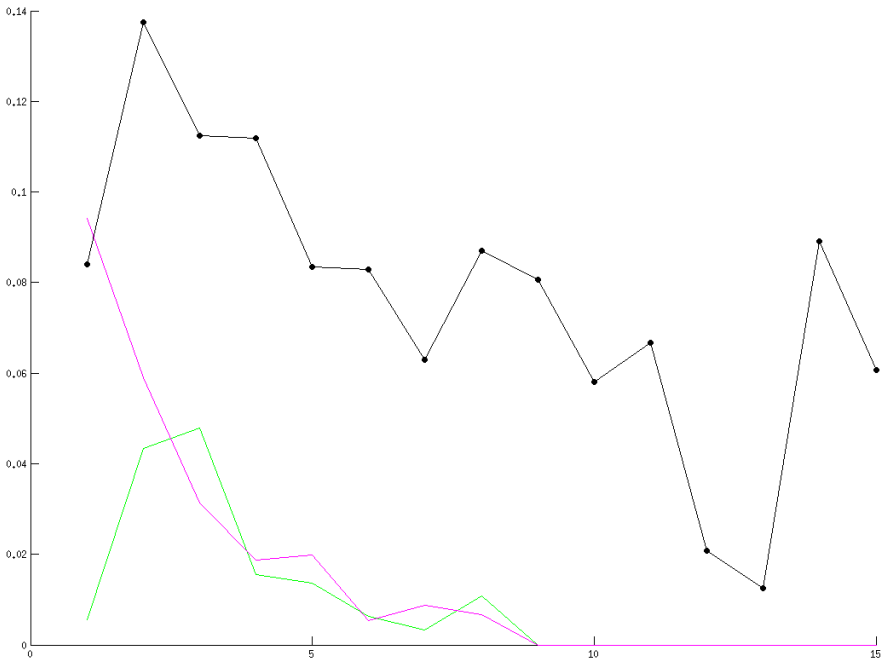
Comparison between models

INCORRECT TRIAL AFTER SOME CORRECT ONES



Comparison between models

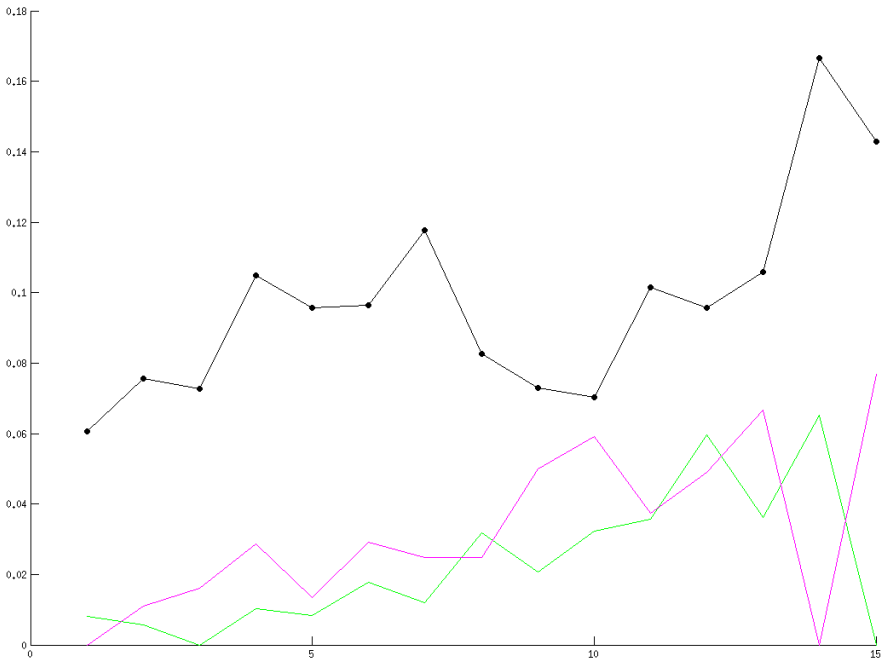
LEAKAGE



trials

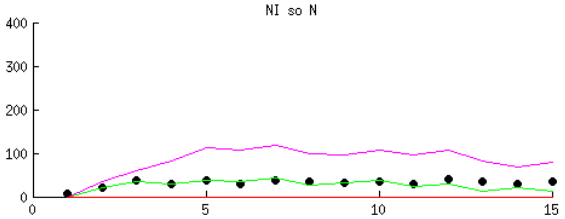
Comparison between models

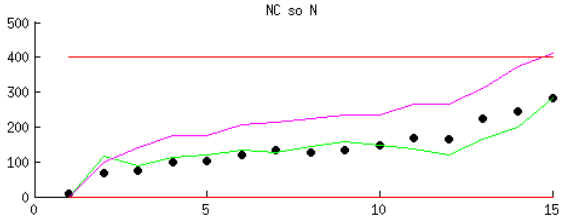
LEAKAGE

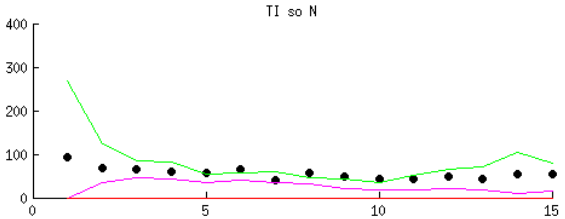


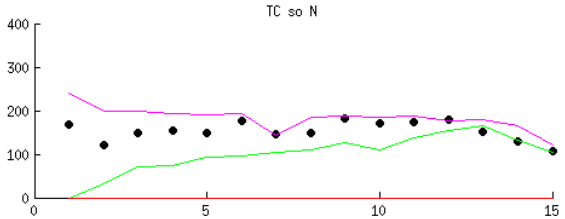
trial distances

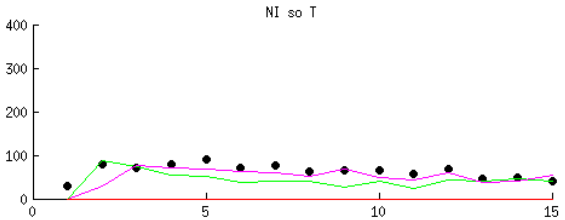
CONDITIONAL PROBABILITIES (length)

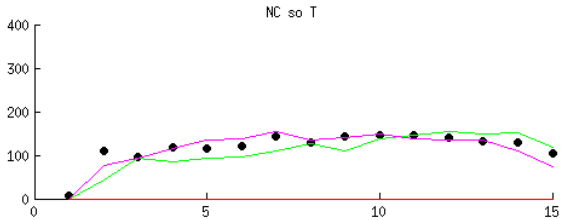


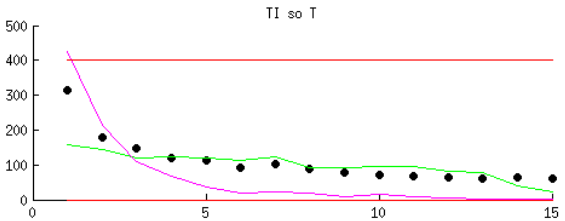


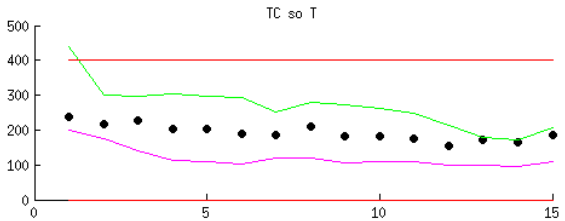










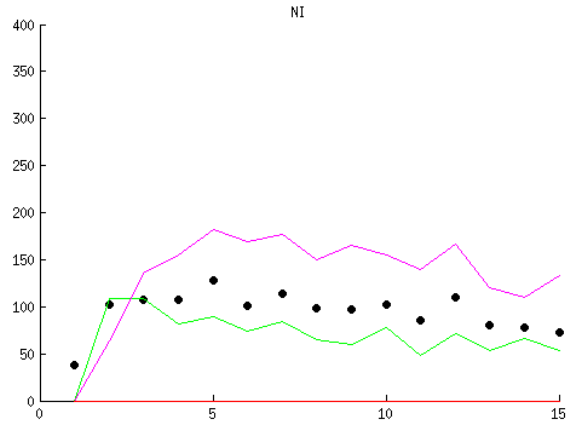


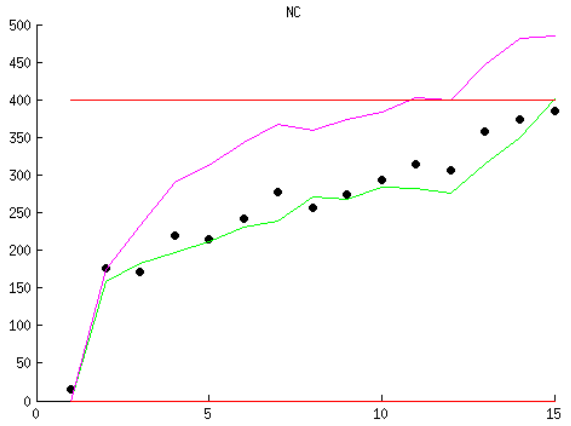
Humans

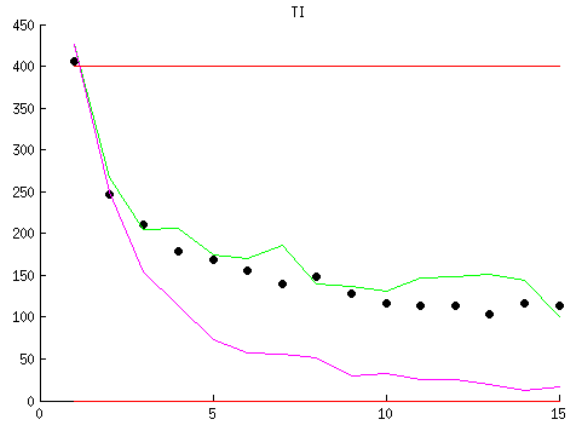
Hybrid

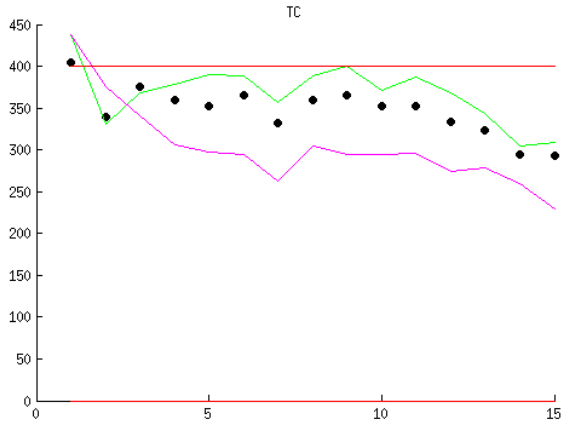
Hypothesis

CONDITIONAL PROBABILITIES (sum)







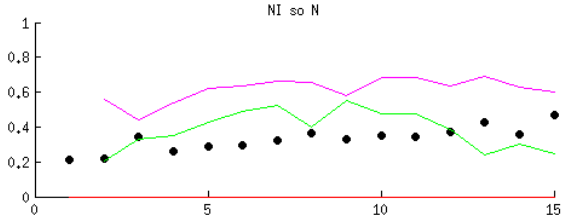


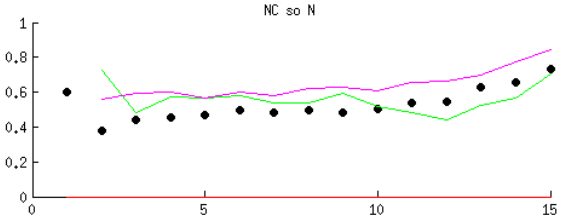
Humans

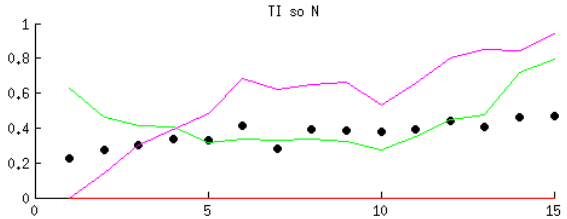
Hybrid

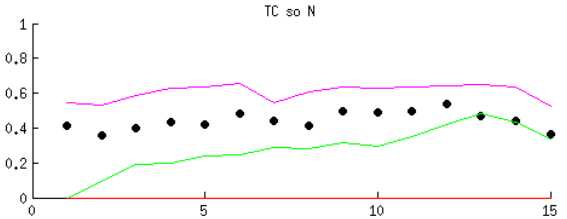
Hypothesis

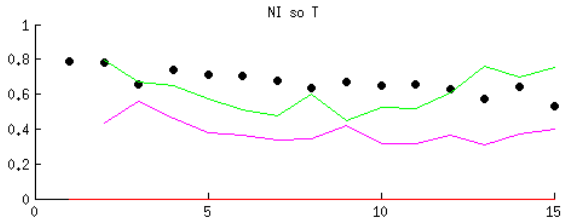
CONDITIONAL PROBABILITIES (prob)

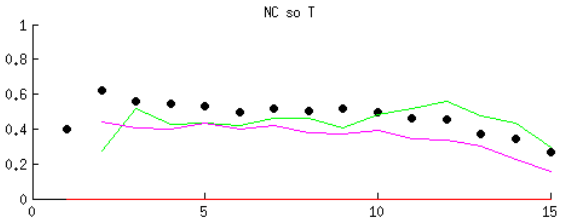


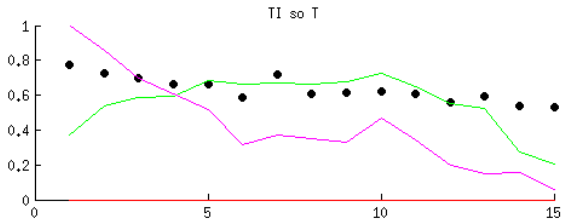


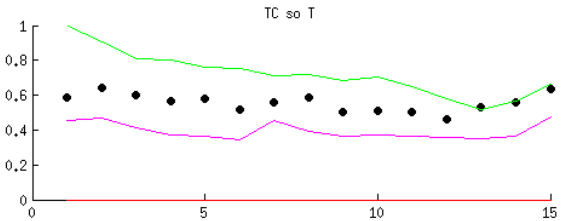












Humans

Hybrid

Hypothesis

What's next

Still need to find where they make divergent predictions!

fMRI:

- incremental value of features
- one trial learn

Behavioral:

- order of trials is not relevant
- order of trials is relevant

Thank you!