Causal learning in an imperfect world

Reliability, background noise and memory

David Lagnado

Neil Bramley

Contents

Introduction to causal learning research

The role of causal beliefs

Causality is key to *prediction, explanation* and *attribution*:

From relatively simple/predictable situations:

- Prediction: What will happen if I hit the ball like this?
- Explanation: How did he pot that?
- Attribution: Is it my fault I lost?

To highly complex uncertain situations:

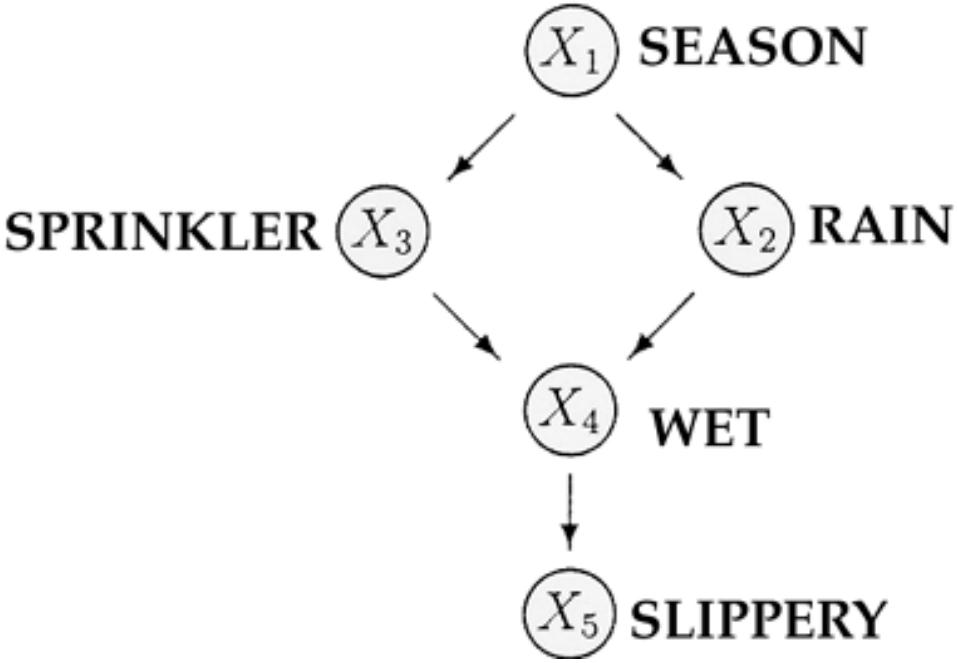
- What would happen if Scotland became independent?
- How does smoking lead to lung cancer?
- Who/what was responsible for the 2006 financial crash?

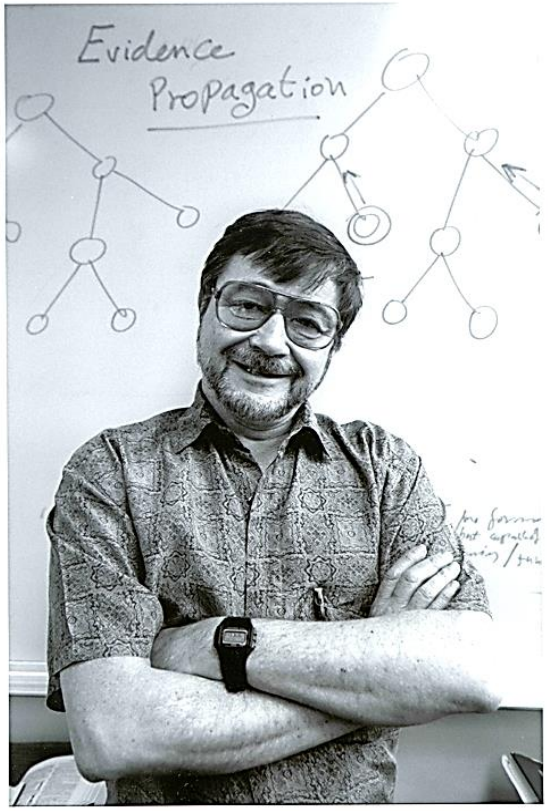
Structure vs. strength

- Traditionally psychology studies worried about how people judged the *strength* of causal connections (e.g. Cheng, 1997; Shanks, 2004)
- But, often identifying whether there is a causal relationship between A and B at all is of primary importance, before worrying about how strong/reliable that relationship is.
- Important to distinguish causal *structure* judgments from causal *strength* judgments
- Here we are interested in how people learn about causal *structure*

Causal models

- Causal Bayesian networks, are a useful framework for understanding causality
- Popularised by Judea Pearl (2000)



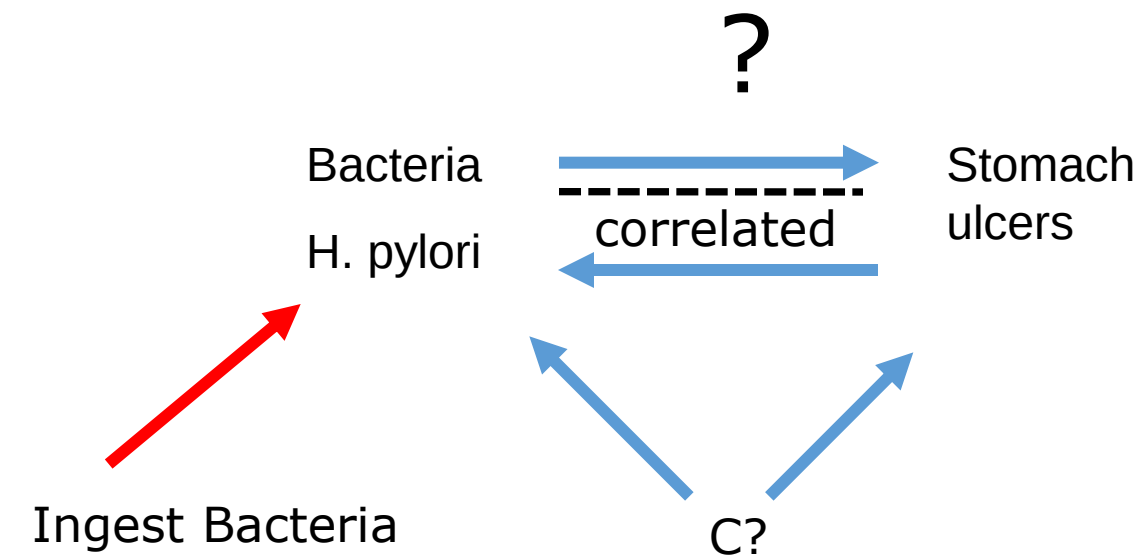


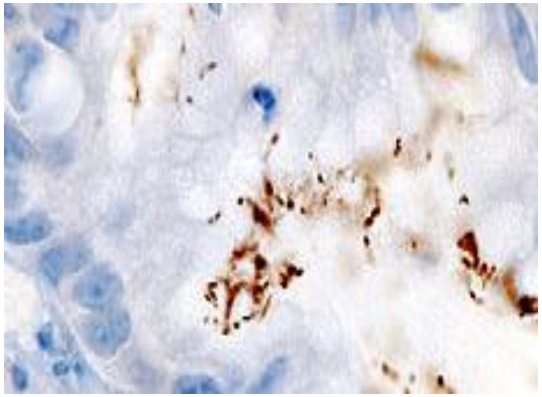
Learning the right causal model

Many factors help us to identify causal structure:

- Temporal cues
  - i.e. What order do events tend to occur?
  - How long/variable is the delay between thing A happening and thing B happening
- Spatial cues
  - Are A and B physically close to one another
  - Can we understand the mechanism connecting thing A to thing B?
- Statistical cues
  - How often do we see A and B happen together, vs. how often just A and how often just B etc.
  - Are their occurrences statistically significantly correlated?
- **Our own manipulations or interventions**
  - **What happens if I make event A occur vs. what happens if I make event B occur?**

Cause of stomach ulcers?

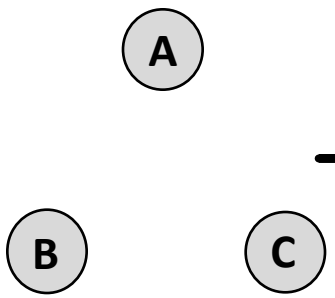


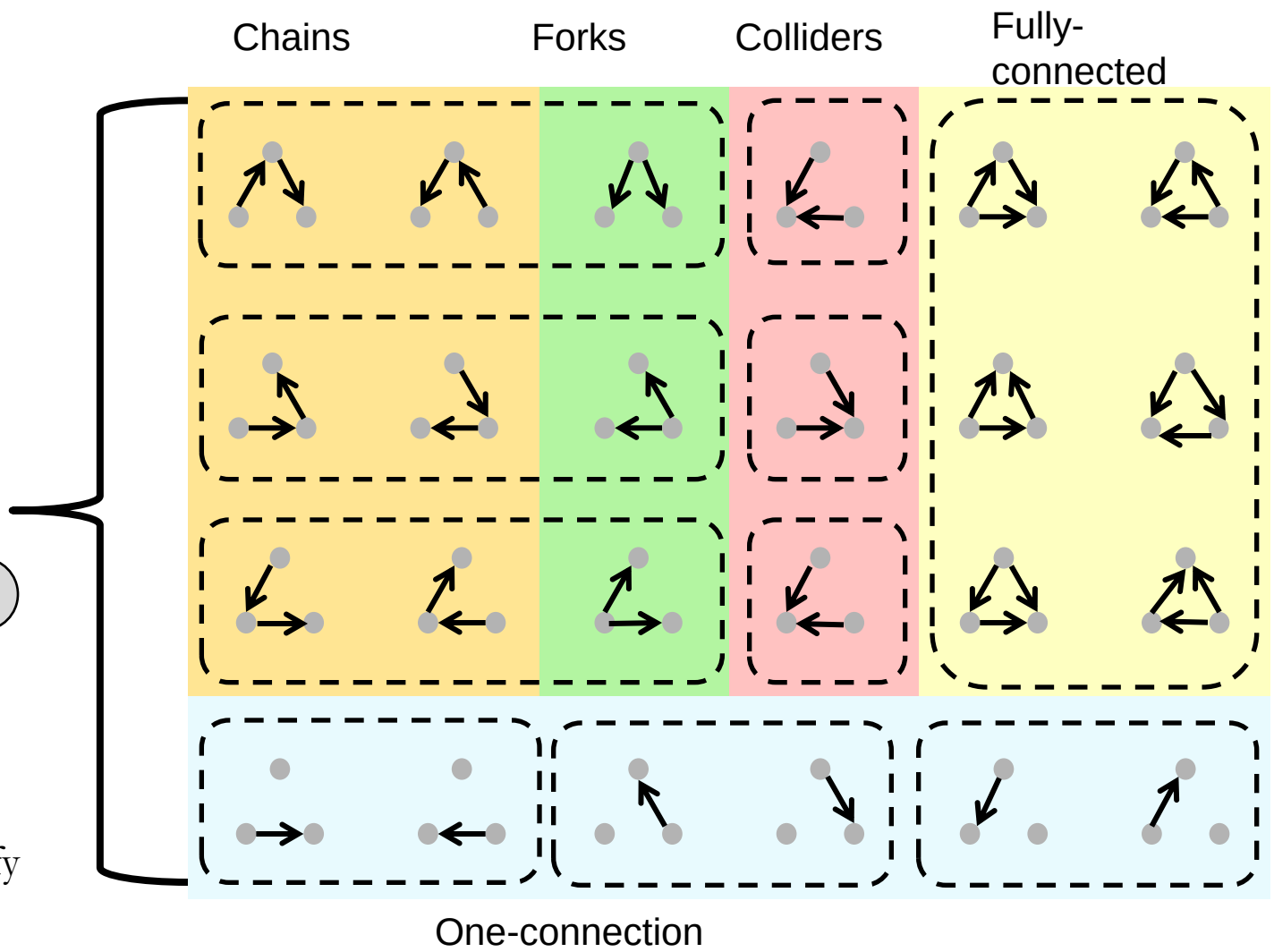


Barry Marshall
Nobel prizewinner

- Intervention to ingest bacteria leads to ulcers
- Therefore Bacteria causes Ulcers

Causal representation and observationally equivalent networks



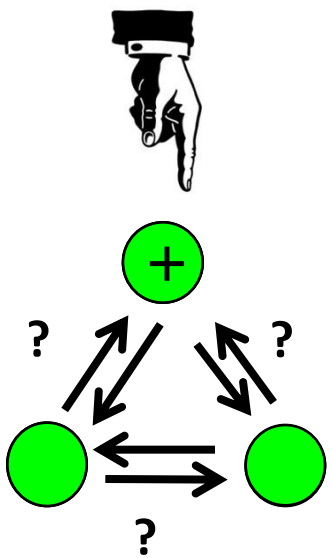


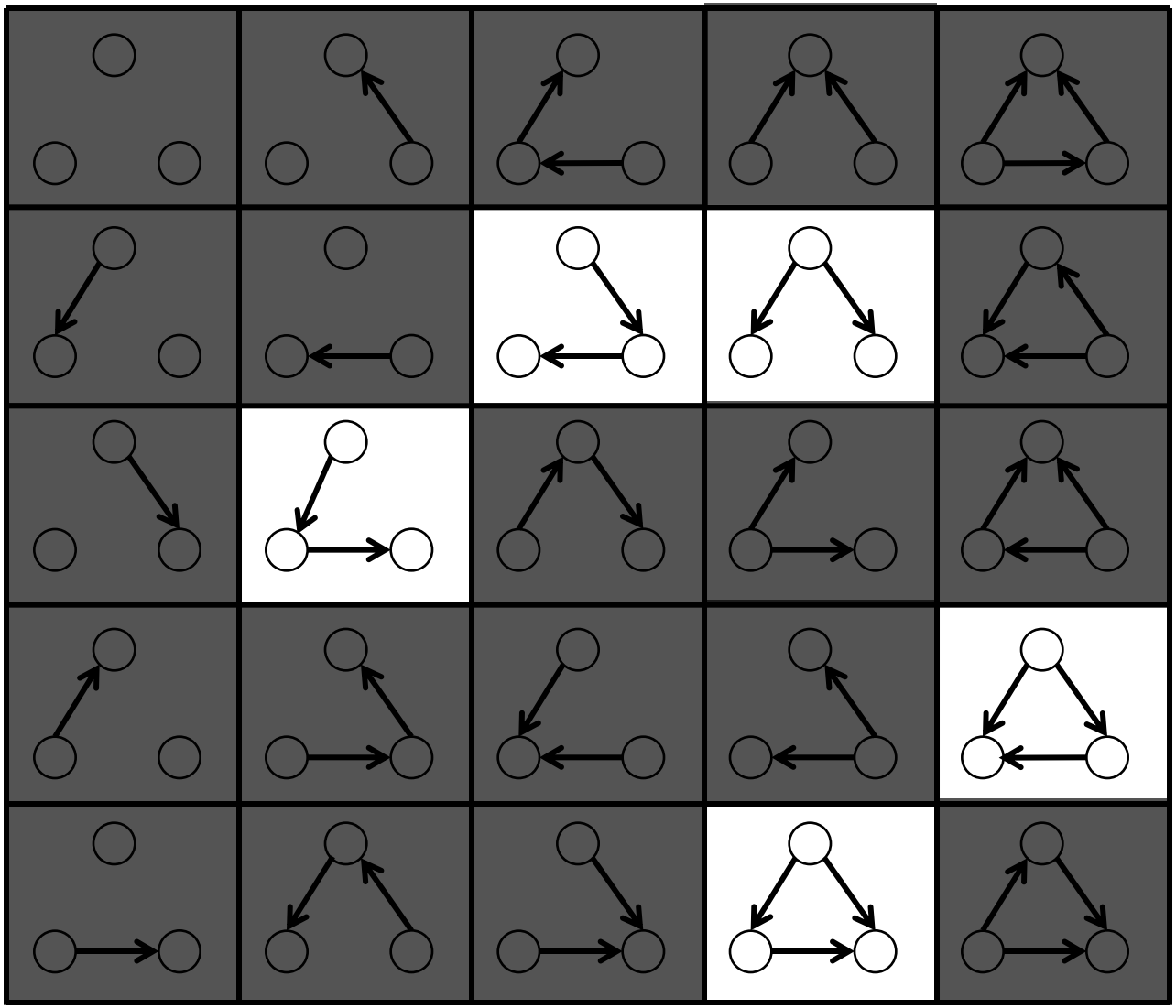
Need interventions (or at least some other cue) to identify unique possible causal structure

Suppose there are 3 binary variables (can be either on or off)

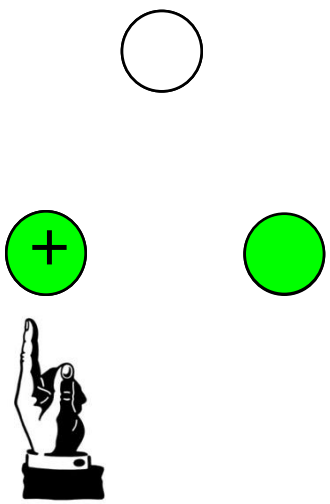
Select interventions (switching variables on or off) to uncover correct causal structure

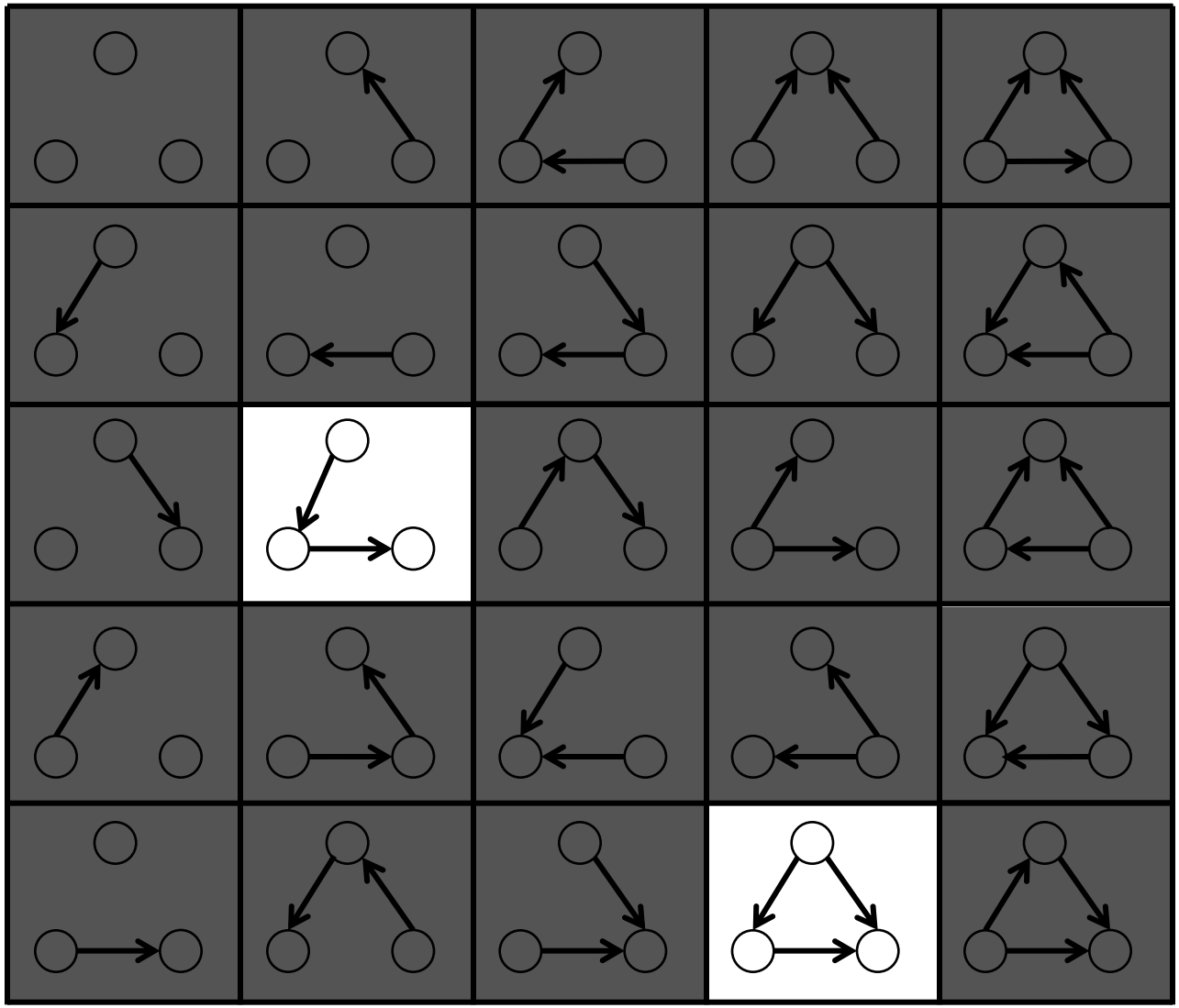
Intervention 1





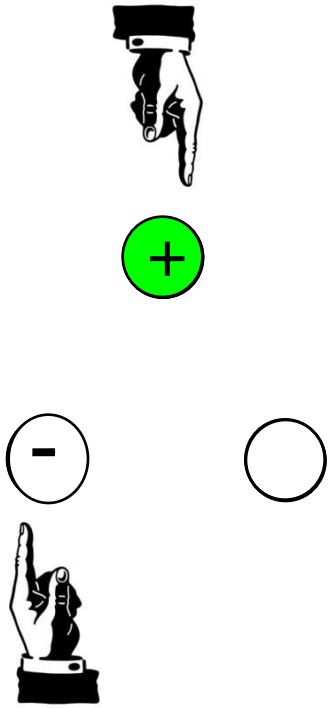
Intervention 2

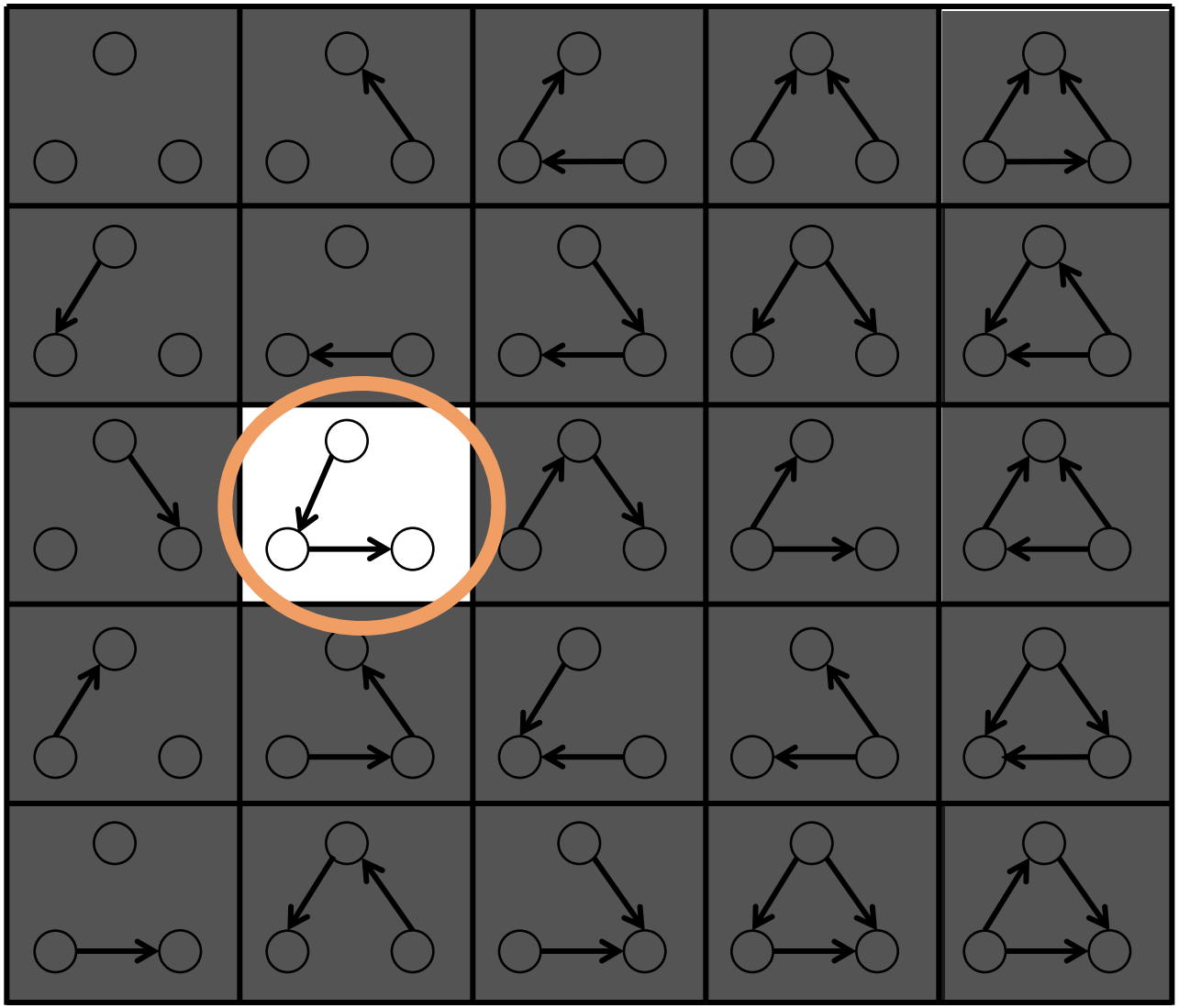




How do we distinguish between final two models?

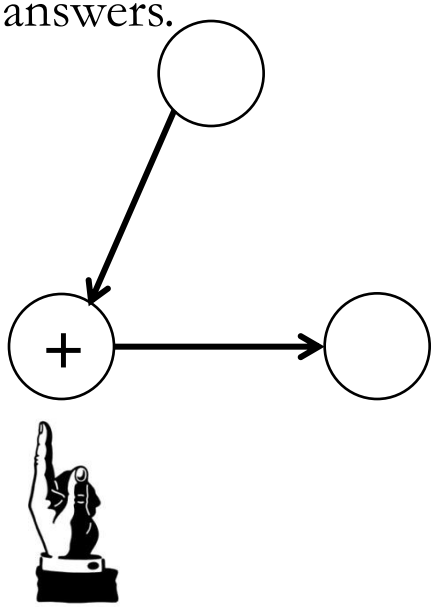
Need “controlled” intervention



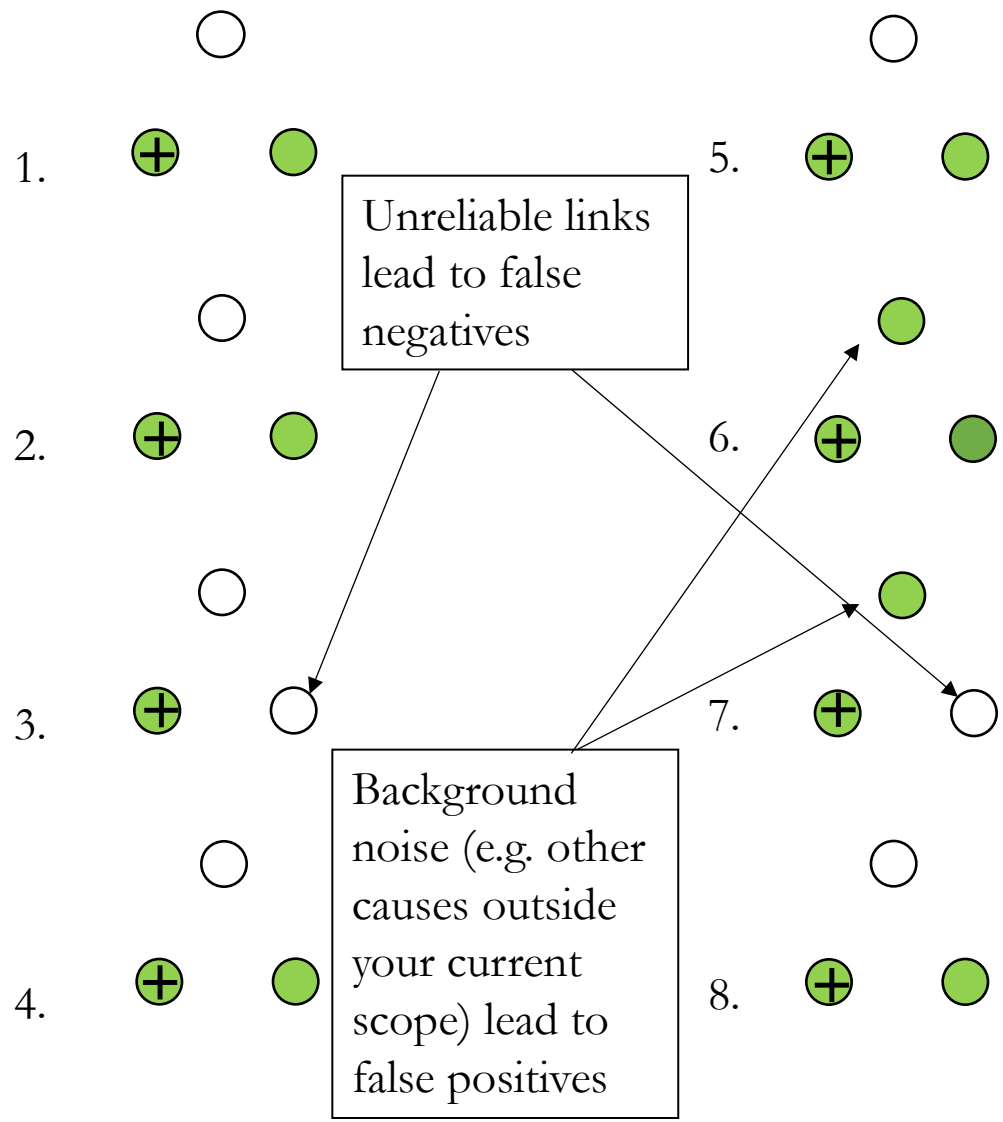


Imperfect causal mechanisms

Repeating the same test multiple times might sometimes gives different answers.



But on average, we can become more sure about the structure:



Questions?

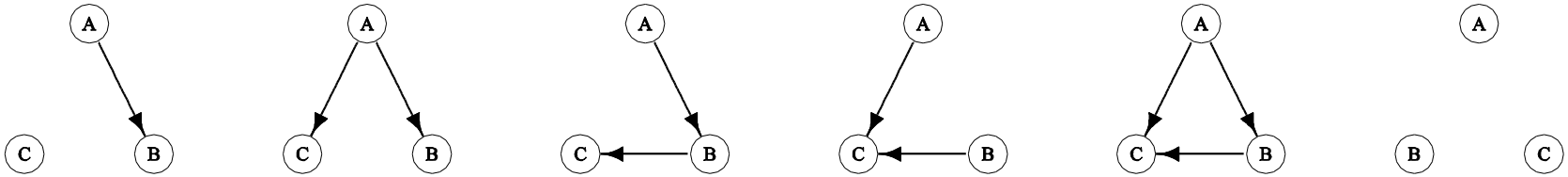
Experiment

Key questions

- How well can people learn imperfect/probabilistic causal relationships:
  - Is learning affected by the strength of the causal connections?
  - Is learning affected by the amount of background noise ?
- Is learning affected by working memory?
  - Does displaying current causal judgment during learning support/hinder performance?
  - Is the effect of displaying the current causal judgment different depending on the reliability of the connections or background noise?

Methods

- Participants and Apparatus
  - 111 UCL undergraduates
  - 89 female, 22 male
  - Mean age=18.73, SD = 0.92
  - Run on individual computers in lab class
  - Program written using Flash Professional CS5.5 (scripting language AS3)
- Stimuli
  - 6 test models (e.g. the different "Devices")



One link | Common cause | Chain | Common effect | Fully connected | No connections

Design – Between subjects

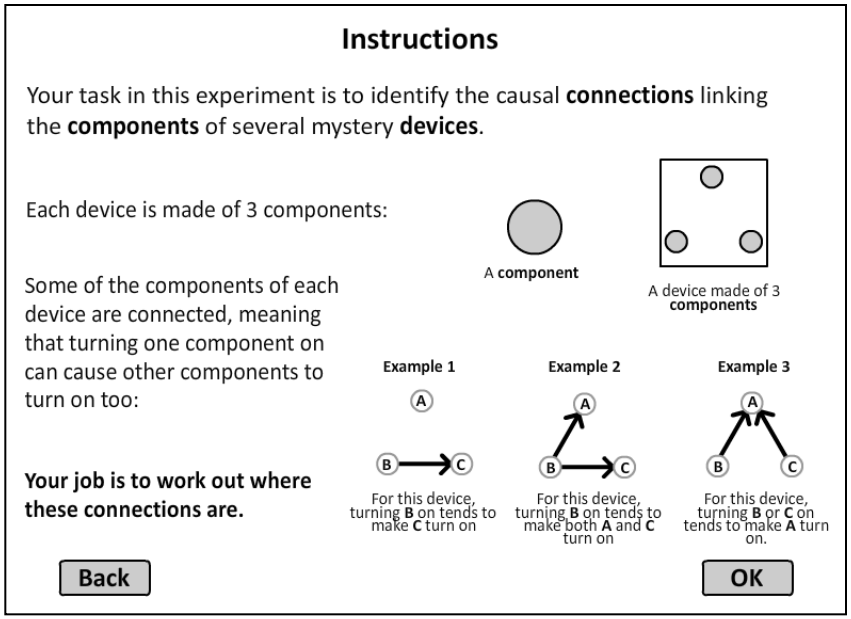
- Each subject randomly assigned to one of $2 \times 2 \times 2 = 8$ groups (determined by the code on post-it-note)
  - Roughly equal group sizes, range 12-15, Mean=13.88, SD=1.45

- 2 working memory conditions
  - Previous judgment (stays on screen / does not stay on screen) while participant chooses next test and makes next judgment.
- 2 background noise conditions
  - Components activate at random 10% vs. 25% of the time
- 2 causal connection strength conditions
  - Components successfully cause their effects 90% vs. 75% of the time

Procedure

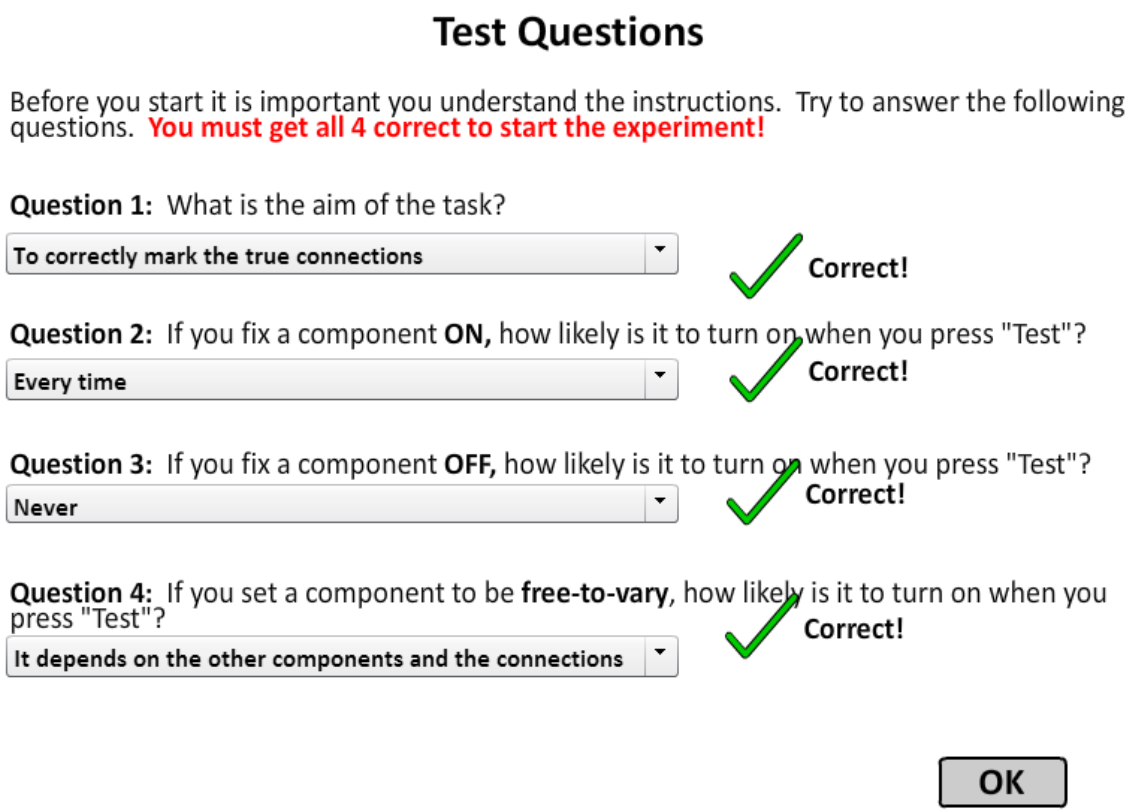
1. Participants first completed extensive interactive instructions, familiarising them with:

   - Task objectives
   - How to set up tests/interventions
   - How to make causal judgments
   - The meaning of active and inactive components
   - The unreliable nature of the connections and the background noise



Procedure

2. Then participants had to correctly answer 4 check questions to ensure they understood the key tenants of instructions, otherwise sent back to beginning of instructions



Procedure

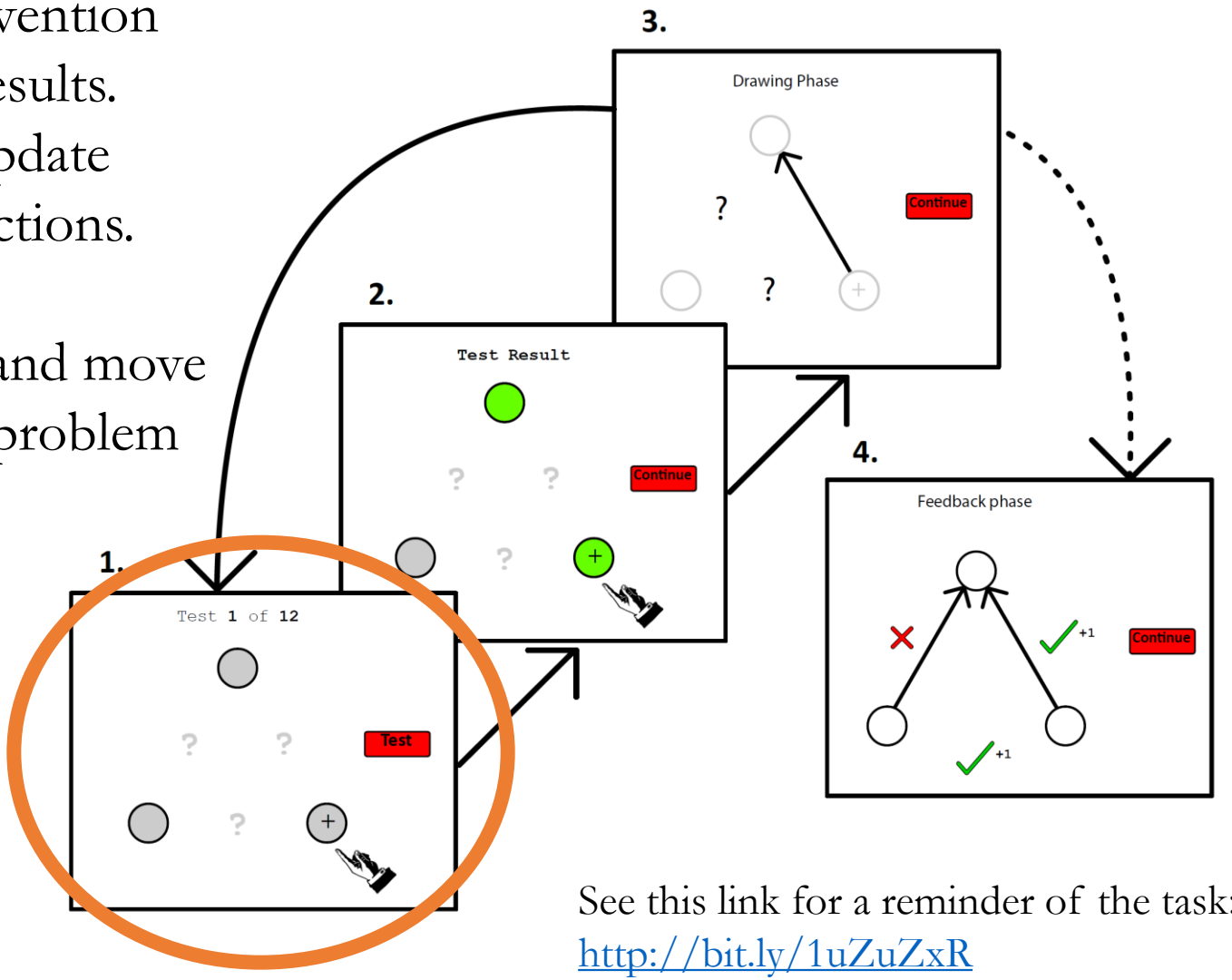
3. Then, participants completed 1 practice problem drawn randomly from the 6 possible test problems
4. Participants completed all 6 test problems in random order:
   - Location of components 'A', 'B' and 'C' also randomised for each problem
   - Outcome of each test generated probabilistically from the true causal structure and the true background noise and reliability levels
5. After each choice of test/intervention participants reported:
   - Expectation about test outcome
   - How much more/less certain they were about the structure
   - After each outcome participants registered their best guess about the structure

Procedure

1. Select an intervention
2. Observe the results.
3. (Optionally) update marked connections.

After 6 trials:

4. Get feedback and move onto the next problem



See this link for a reminder of the task: http://bit.ly/1uZuZxR

Procedure

6. Then participants completed the ‘chain’ problem a second time. This time reporting how they had chosen each intervention with free text responses.
7. Finally participants estimated the amount of noise and the causal strength of the connections in the problems they had just faced.

Dependent measures

- Accuracy of final judgments – What proportion of participants' final connection judgments were correct
  - From 0/21 connections correct = 0
  - To 21/21 connections correct = 1
- Interventions chosen – How useful were the interventions/tests chosen by the participant for identifying the true connections?
- Final judgments about level of background noise and reliability/strength of true causal connections

Independent measures

Between subjects:

- Previous judgments: remain vs. disappear
- Background noise: 10% vs. 25%
- Connection reliability: 90% vs 75%

Within subjects:

- The structure to be learned (1-6)
- The which test for that structure (1-6)

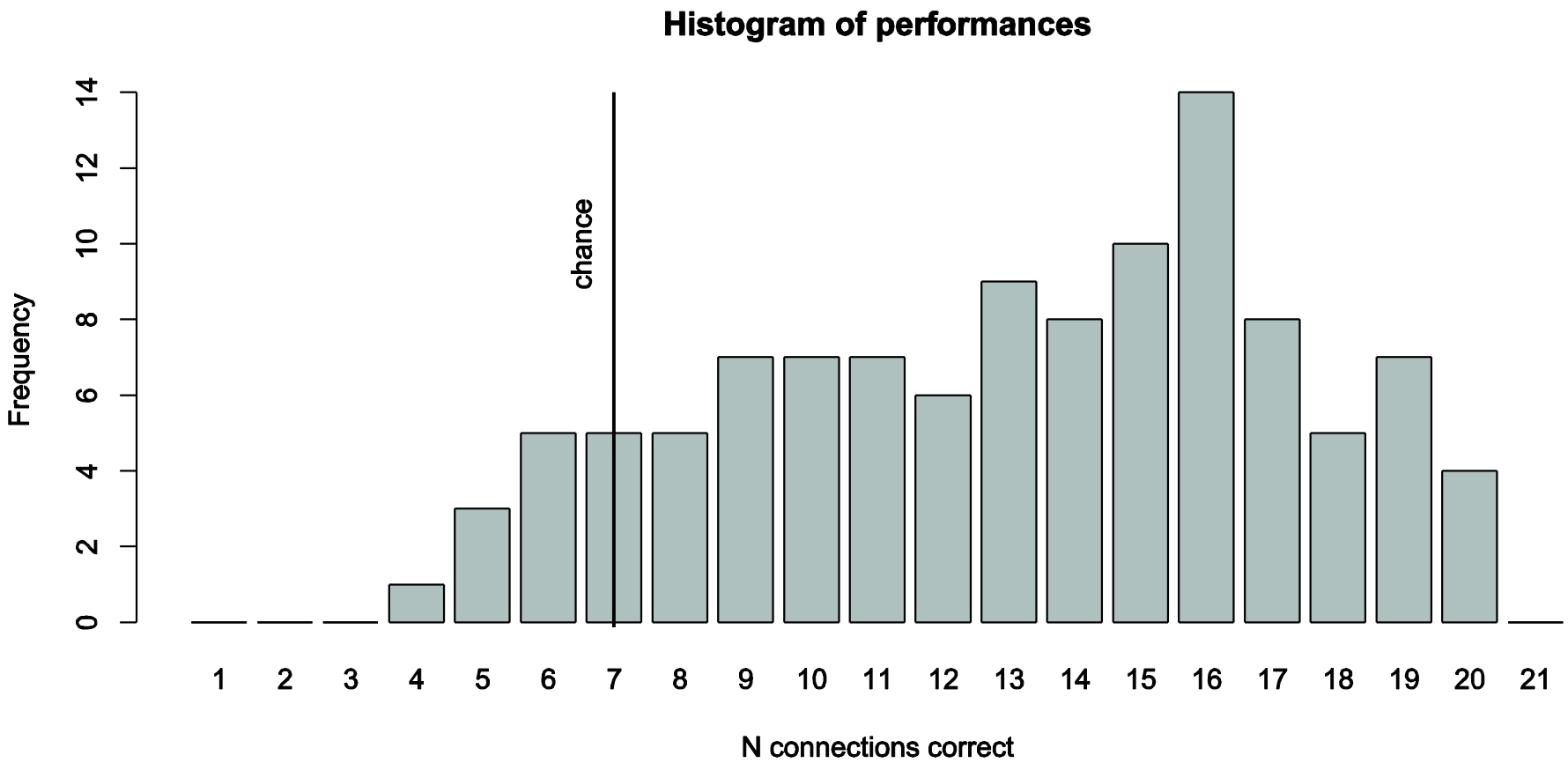
Other measures

- Expected outcome slider judgments
- Helpfulness of outcome slider judgments
- Connection-wise probability slider judgments
- "Type-out-loud" explicit descriptions of own strategy
- Reaction times

Questions?

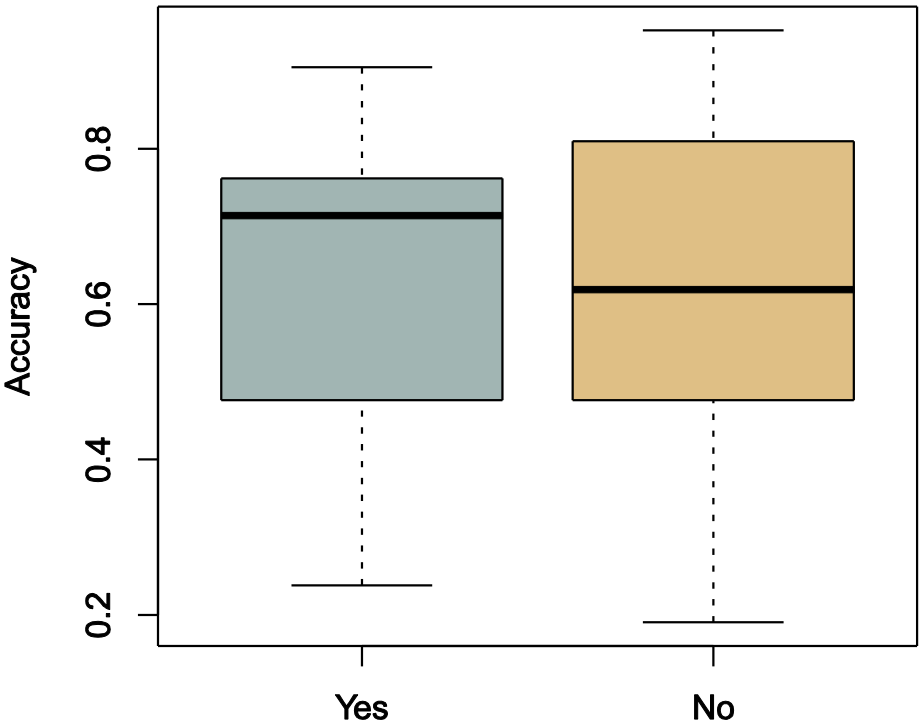
Results

Overall accuracy



Mean=13.05, Median=14, Mode=16

Did removing record of previous causal judgment affect performance?



Previous causal judgment remains?

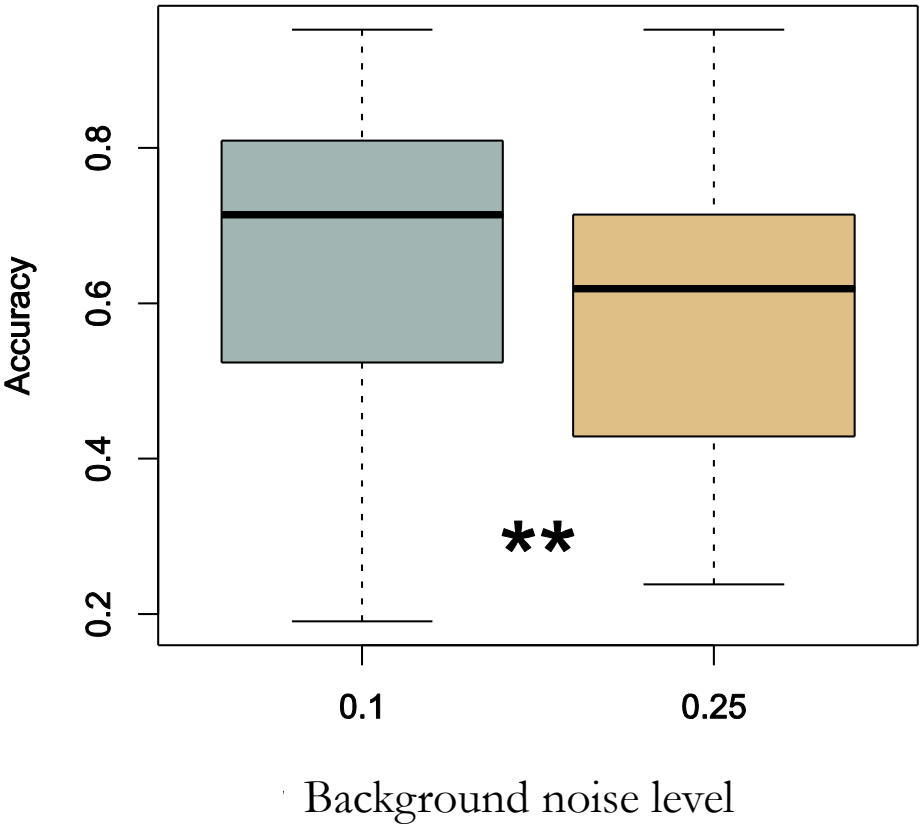
Independent samples t-test:

Means = .63, .61
t(109) = -.669
p=.51

Not significant

No

Did level of background noise affect performance?



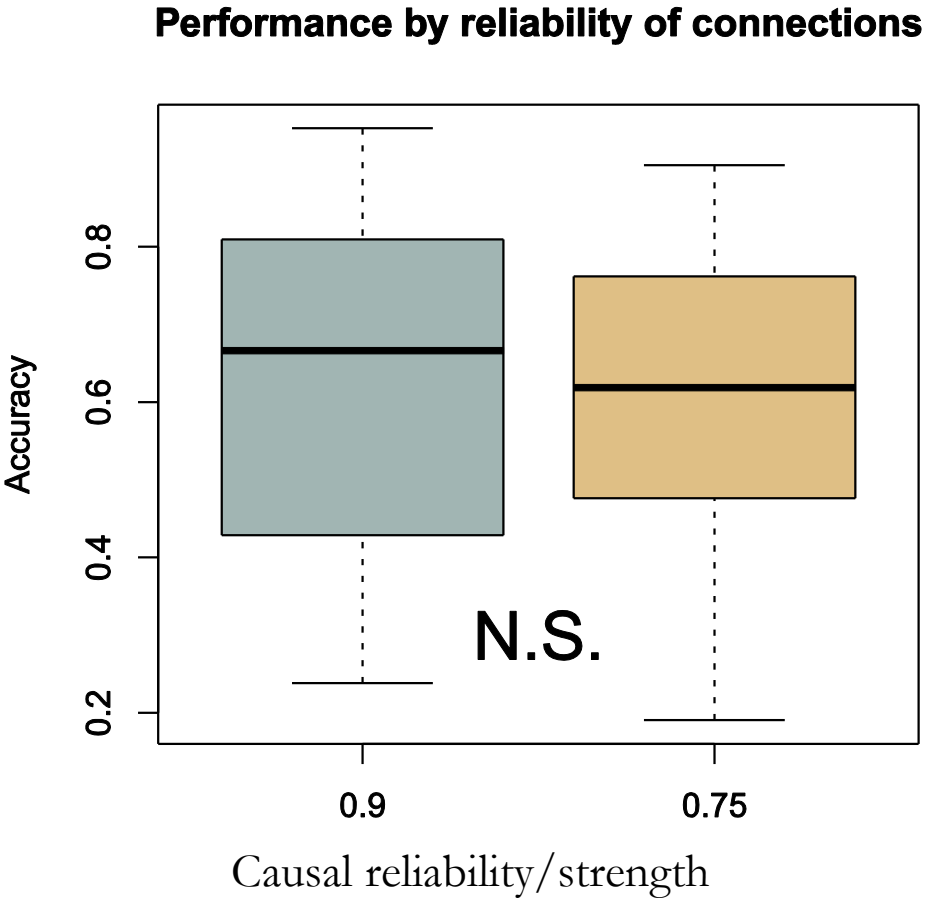
Independent samples t-test:

Means = .67, .57
t(108) = -2.70
p=.007**

Significant

Yes

Did reliability/strength of the causal connections affect performance?



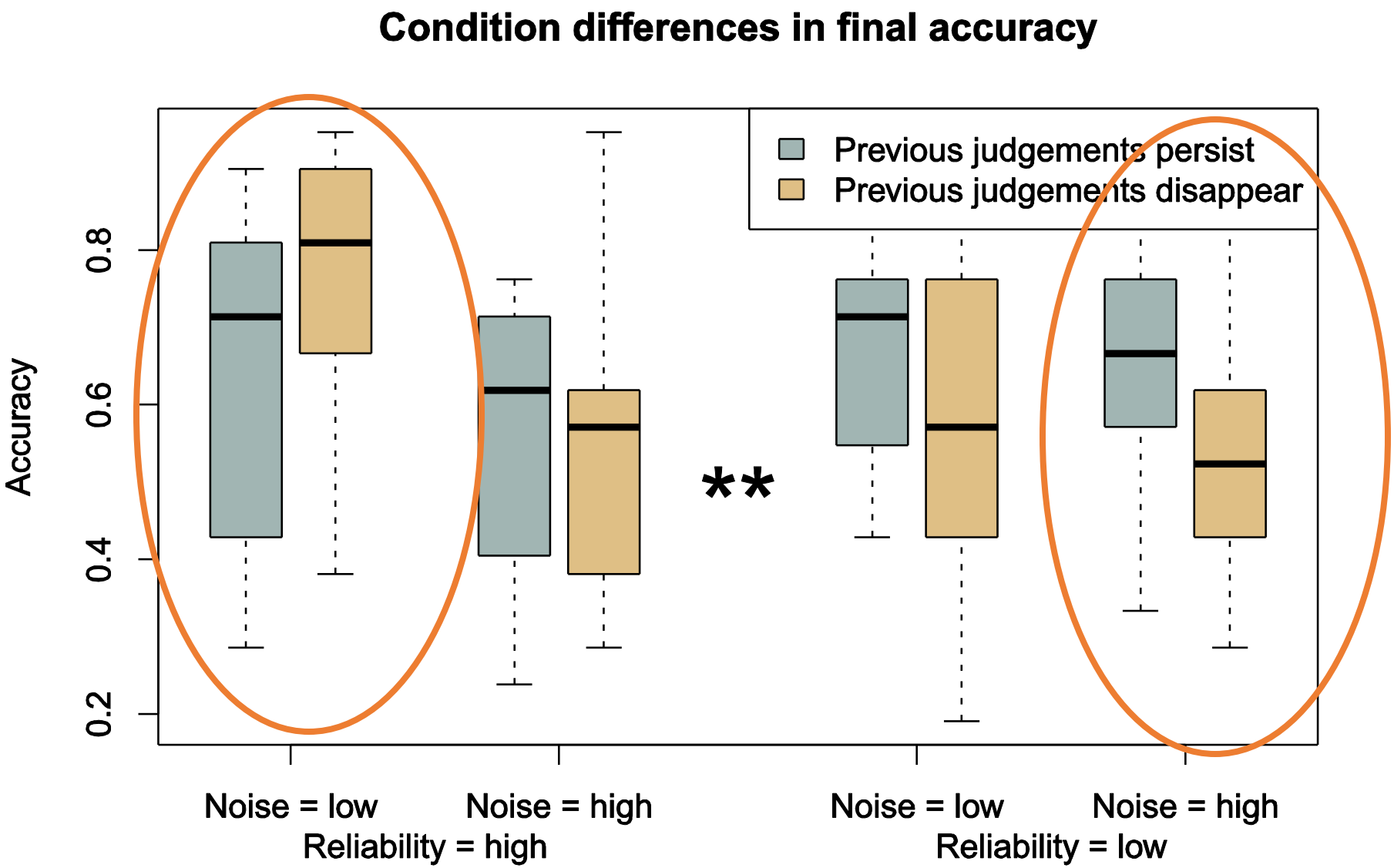
Independent samples t-test:

Means = .64, .60
t(107) = -.88
p=.37

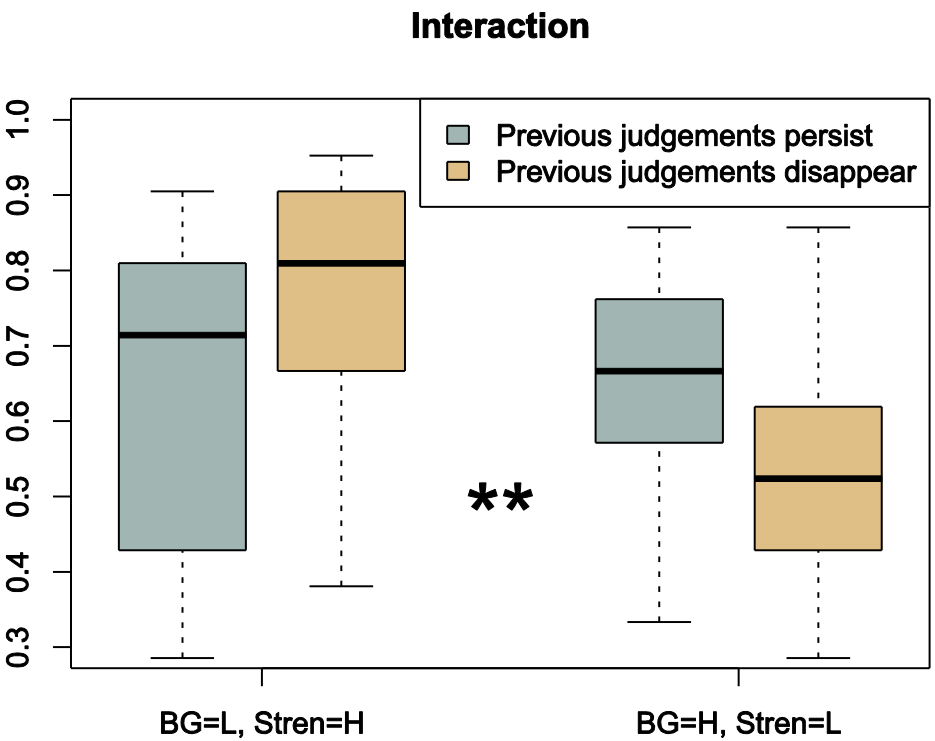
Not significant

No

Looking at all the groups together



Interaction between removing connections and noise level

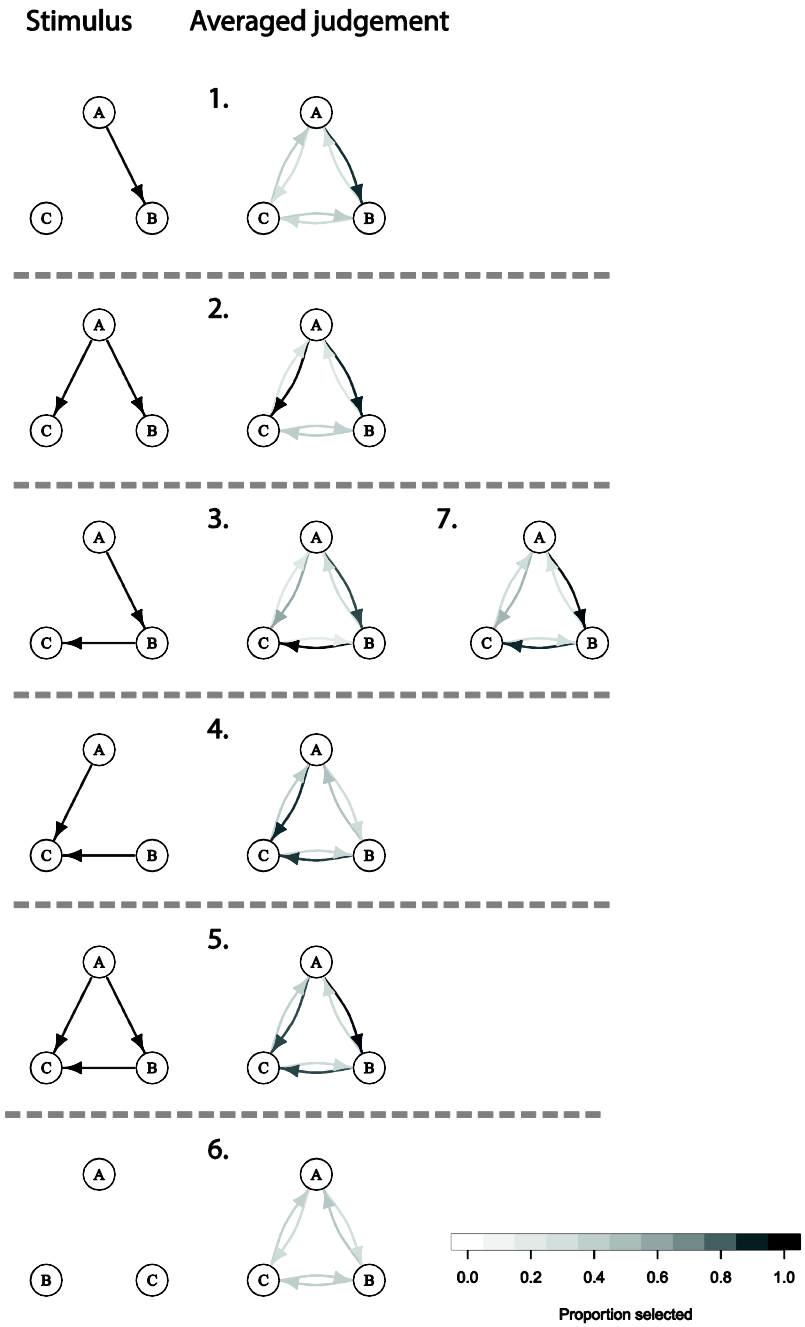


- Performance is unaffected by noise level when causal judgments persist into new trial:

  Means = .65, .64
  t (22)= -0.03,
  p = 0.97
  Not significant

- But performance is highly affected by noise when causal judgments disappear after each trial:

  Means = .77, .52
  t(26) = -3.9,
  p = 0.0007***
  Highly significant


Stimulus
Averaged judgement
1.
2.
3.
7.
4.
5.
6.
A
B
C
0.0
0.2
0.4
0.6
0.8
1.0
Proportion selected

Summary

- Participants were better able to learn causal connections when there was not much background noise (components did not activate by chance very often)
- But their performance was *not* affected by the strength/reliability of the connections.
- On average accuracy was not affected by whether a record of previous judgment persisted from test to test.
  - However, there was a significant interaction, meaning when there was a lot of noise, and unreliable causal connections then participants who could not see their previous judgment performed a lot worse, while those who could see their previous judgement were not affected.

Discussion

Why might false positives be harder to deal with than false negatives when it comes to identifying causal structure?

Several possible effects of the record of previous judgments:

- Easier to keep track of one's evolving causal belief
- But perhaps correspondingly less deep encoding required as information comes in / less of a "real" learning situation
- Might create primacy e.g. overreliance on earlier stated beliefs that become hard to overturn

Extensions

- We have analysed performance in terms of the between subjects factors but not really looked at psychological processes
- How did people come up with the tests/interventions they performed?
- How exactly were they moved by the evidence they saw?
- What psychological variables and representations went into the learning and the judgments?

Recommended reading

Lagnado, D. A., Waldmann, M. R., Hagmayer, Y., & Sloman, S.A. (2007). Beyond covariation: Cues to causal structure. Gopnik, A., & Schultz, L. (Eds.), *Causal learning: Psychology, philosophy, and computation*. OUP. (http://bit.ly/1zKky32)

Bramley, N. R., Lagnado, D. A., & Speekenbrink, M. (2014). Conservative forgetful scholars-How people learn causal structure through sequences of interventions. *Journal of Experimental Psychology: Learning, memory and Cognition*, online-first access (http://bit.ly/1tYFwX3).

Or find more of David Lagnado's papers here:

http://www.ucl.ac.uk/lagnado-lab/david_lagnado.html

Additional references

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological review*, *104*(2), 367.

Pearl, J. (2000). *Causality: models, reasoning and inference*. Cambridge: MIT press.

Shanks, D. R. (2004). Judging covariation and causation. *Blackwell handbook of judgment and decision making*, 220-239.

Sloman, S. (2009). *Causal models: How people think about the world and its alternatives*. Oxford University Press.

Sobel, D. M., & Kushnir, T. (2006). The importance of decision making in causal learning from interventions. *Memory & Cognition*, *34* (2), 411-419.

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E. J., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive science*, *27*(3), 453-489.

Questions?

- For further questions about the study for your reports, email:

David Lagnado - d.lagnado@ucl.ac.uk

Or me - neil.bramley.10@ucl.acuk

£20 Amazon vouchers

- Most accurate participant:

211v8xbqsw

- Most well calibrated participant (most accurate slider judgments):

211rQdG6M6

Email David Lagnado or myself to receive your prizes.