A review of recent developments in research and theories on human contingency learning

Jan De Houwer
University of Ghent, Ghent, Belgium
Tom Beckers
University of Leuven, Leuven, Belgium

Percy Mistry
11 February 2016

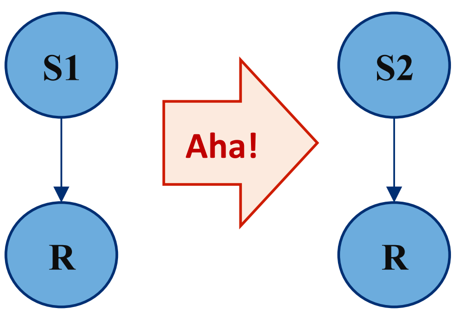
Human Contingency Learning

learning about the relation between the presence of two stimuli

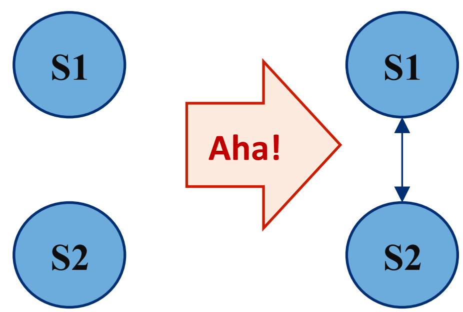
Difference between HCL & Pavlovian

(does not have to be causal)

Pavlovian
Mapping of innate response associated with a potent stimulus now also mapped to a previously neutral stimulus



HCL
Explicit mapping between two stimulus



Models and explanations

	Probabilistic			Associative		
HCL characteristics	**ΔP**	**PC**	**Power PC**	**RW**	**Revised RW**	**Revised SOP**
Forward blocking	✗	✓	✓	✓	✓	✓
Backward blocking	✗	✓	✓	✗	✓	✓
Retrospective evaluation only when A-T co-occur	✗	✗	✗	✗	✓	✓
Retrospective evaluation without A-T co-occurring	✗	✓	✓	✗	✗	✗
Recency effects	✗	✗	✗	✗	✓	✓
Primacy effects	✗	✗	✗	✗	✗	✗
Learning curves	✗	✗	✗	–	✓	✓
Ceiling effects	✗	✗	✓	✗	✗	✗
Active recoding	✗	✗	✗	✗	✗	✗
Variable blocking	✗	✗	✗	✗	✗	✗

Characteristics of HCL:
Blocking

Contingency judgments exhibit forward and backward blocking

FB: Response (T) { A+, AT+ } < Response (T) { AT+ }

BB: Response (T) { AT+, A+ } < Response (T) { AT+ }

T = Target cue
A = Alternate cue
+ = Presence of US (outcome O)

Note that FB changes the association between O-T on trials when T is present but BB changes the association between O-T on trials when T is not present

Characteristics of HCL: Asymmetry in forward vs retrospective cue competition

{A+, AT+} always influences Response (T)

{AT+, A+} only influences Response (T) when A and T repeatedly co-occur *

Implications: Different mechanisms for FB and BB

* But not always!
Retrospective evaluation effects have been shown even when the alternative cue is not directly associated with the target cue

Retrospective Revaluation Study Example

Forward cue competition

Bottom of page 294

{ [A1+, A2+, A3+], [A1T+, A2T+, A3T+] }

{ [A1+, A1+, A1+], [A1T+, A1T+, A1T+] }

A blocks ***T*** in both conditions

Backward cue competition *(pg. 294)*

{ [A1T+, A2T+, A3T+], [A1+, A2+, A3+] }

{ [A1T+, A1T+, A1T+], [A1+, A1+, A1+] }

A only blocks ***T*** in the second condition

Predicted by revised RW and SOP models but not by non-associative models

Characteristics of HCL:
Order effects (Primacy vs Recency)

Contradictory findings:

Recent information has a stronger impact on contingency judgements than does information that was presented earlier

Stronger influence of initial than of later trials under certain conditions (confirmation bias)

Characteristics of HCL: Learning curves

Contingency judgements become more accurate as the number of trials concerning a particular cue–outcome relation increases *

- when a cue is repeatedly paired with an outcome, associative strength will gradually increase over trials until it reaches an asymptote that reflects the actual contingency between the cue and outcome.

- increases in associative strength will be bigger on initial trials

* But not always!

Characteristics of HCL: Ceiling Effects

Impact of **$p(O|A,\sim T) = 1$**

Normatively, inferences about the relation between T and O should be based on the comparison **$p(O \mid Situation, T) - p(O \mid Situation, \sim T)$**

Ceiling effect: This inference breaks down when **$p(O \mid Situation, \sim T) = 1$**

Characteristics of HCL:
Variable blocking

Cue competition (and resulting blocking) can be modulated by task instructions and suggestions of causal direction

Characteristics of HCL: Active recoding

People are able to recode prior events on the basis of new information about those events and to adjust their judgements accordingly

ΔP Model

Models contingency judgements as a fairly accurate reflection of the actual statistical contingencies between events.

Response (T) depends on $\mathbf{p(O|T) - p(O|{\sim}T)}$

Cannot explain forward or backward blocking

Rescorla-Wagner Model (associative)

$$\Delta V_n = \alpha\, \beta\, (\, \lambda - \Sigma\, V_{n-1}\,)$$

ΔV_n = Change in associative strength on a given trial

ΣV_{n-1} = Associative strengths of other cues present on that trial

λ = Fixed +ve maximum associative strength when US present;
= 0 when US absent

α = Positive is cue present; 0 if cue absent

<u>Explains forward, but not backward blocking</u>

Probabilistic Contrast Model

Response (T) depends on **p(O|A,T) – p(O|A,~T)**

Ignores trials involving p(O|~A,T) or p(O|~A,~T)

<u>Explains forward and backward blocking, but not other HCL characteristics</u>

Revised Rescorla-Wagner Model

Original model: Associative strength of a cue could change only on trials where that cue was actually present

Revised: Associative strength of a cue could also change on trials where the cue was absent but expected.

$$\Delta V_n = \alpha \beta (\lambda - \Sigma V_{n-1})$$

α = Positive is cue present; **Negative if cue absent**

e.g. on trial sequence { AT+, A+ } T is expected but absent on the second trial. This results in a change in V_n in the direction opposite to the direction of change when cue T is present.

Explains recency effects, retrospective evaluation only when A-T co-occur, and learning curves

Revised SOP (Alternate Associative Model)

Standard SOP:

- Stimuli represented as nodes that consist of several elements **[cues?]**
- Each element can be in one of three states: an inactive state (I), a primary active state (A1), or a secondary active state (A2).
- Transition from I to A1 only when the particular stimulus presented.
- Transition from I to A2 even when an associated stimulus is presented.

What happens if one stimulus node has more features than another, are non-existant features fixed to an inactive state? **Yes, unless this.**

Revised SOP:

- Excitatory association between two representations will increase in strength as a function of the number of elements of both stimulus representations that are in the same active state (A1–A1 or A2–A2).
- Inhibitory associations will increase in strength as a function of the number of elements in different active states (A1–A2 or A2–A1).
- Contingency judgements: Function of the difference between the strength of the excitatory and inhibitory association that link the representation of cue and outcome.

Explains recency effects, retrospective evaluation only when A-T co-occur, and learning curves

- Note: SOP is actually "Sometimes opponent process"

Power PC (Causal learning)

Response (T) depends on

$$\frac{p(O|A,T) - p(O|A,\sim T)}{1 - p(O|A,\sim T)}$$

<u>Explains retrospective evaluation regardless of whether A-T co-occur, and ceiling effects</u>

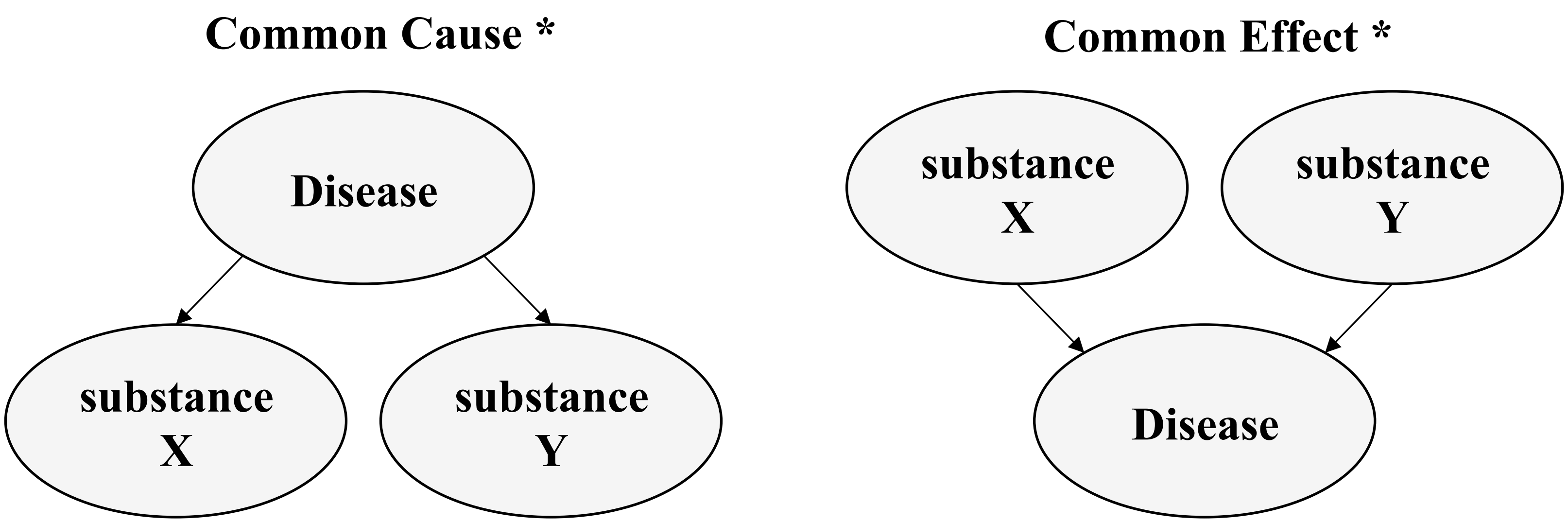
Revisiting models and explanations

	Probabilistic (Non-associative)			Associative		
HCL characteristics	**ΔP**	**PC**	**Power PC**	**RW**	**Revised RW**	**Revised SOP**
Forward blocking	✗	✓	✓	✓	✓	✓
Backward blocking	✗	✓	✓	✗	✓	✓
Retrospective evaluation only when A-T co-occur	✗	✗	✗	✗	✓	✓
Retrospective evaluation without A-T co-occurring	✗	✓	✓	✗	✗	✗
Recency effects	✗	✗	✗	✗	✓	✓
Primacy effects	✗	✗	✗	✗	✗	✗
Learning curves	✗	✗	✗	–	✓	✓
Ceiling effects	✗	✗	✓	✗	✗	✗
Active recoding	✗	✗	✗	✗	✗	✗
Variable blocking	✗	✗	✗	✗	✗	✗

Causal Model:
Can account for variable blocking & active recoding

Blocking depends upon underlying causal model assumed:

A figure would've been nice here: Here it is!



Common Cause: Disease causes X & Y in the blood. Observing X should not block Y.

Common Effect: Disease can be caused by X or Y. Observing X would block Y.

* Note: Paper flips the labels for Common Cause & Common Effect

Configural learning

Going beyond elemental approaches

Considering compound cues (e.g. XOR learning)

Cannot account for retrospective cue competition effects

What factors underlie whether a participant chooses a configural or elemental encoding strategy when processing cues?

[My opinion]:

- *Matter of representation;*
- *Varies on a continuous rather than discrete scale.*

What sort of research has been done on HCL in terms of context-dependent learning? I would be interested in hearing about occasion setting with contingency learning.

I'm pretty curious about how abstract Rule Learning could be integrated into models of HCL.

Model based learning?

What if people are using observations to build up conditional dependencies between cue events, outcome events, and potentially unobservable events (i.e. building Bayes nets).?

Yes, that is what the causal model theory is proposing (in part).

Conclusions

Large number of diverse observations

No single model can accommodate all

Multiple systems of learning?

Need to account for task level modulators of performance?
e.g. number of cues

Thank You!

Example:
Learning Curves in probabilistic models of contingency learning