With thanks to:
Tom Barrasin, Catherine Bryant, Sue Chapman, Steven Graham, Tim Hodgson, Stephen Lea, Lisa Leaver, Chris Longmore, Ian McLaren, Lou Millar, Fraser Milton, Britta Osthaus, Matt Rolland, Mark Suret

Supported by:
British Council, EC and ESRC.

Recurrence and higher-order thought

Andy Wills
University of Exeter

Definitions and connections

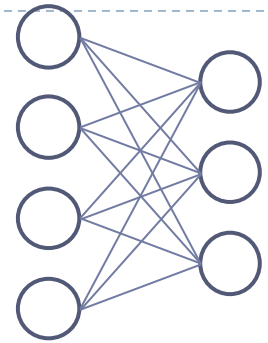
- **Higher-order thought** as those processes most likely to take place when the task has our full attention, and when we have sufficient time.
- **Recurrence** as the process by which the outputs of a system feed back into its inputs
  - Addition e.g. 5 + 2, *when calculated thus*: 5 (+1) 6 (+1)
- Recurrence as a candidate characteristic of higher-order thought.
  - Higher-order thought as dependent on our ability to maintain intermediate states (working memory)
- Higher-order thought in pigeons and squirrels?

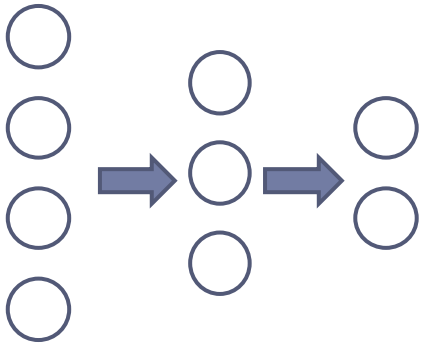
Components of this talk

- Computational power argument
- Recurrence and higher-order thought in free classification (including some pigeon and squirrel work).
- Recurrence and higher-order thought in contingency learning

Recurrence and higher-order thought: A rough theory

- Simple networks
  - Computationally limited (e.g. XOR).
- Hidden-layer networks
  - Universal approximator (Hornik, Stinchcombe & White, 1989), but still a bounded computation, and hence computationally limited (Sima & Orponen, 2003).
- Recurrent networks
  - Certain types of recurrent network have the *computational power of a Turing machine* (Sima & Orponen,2003)





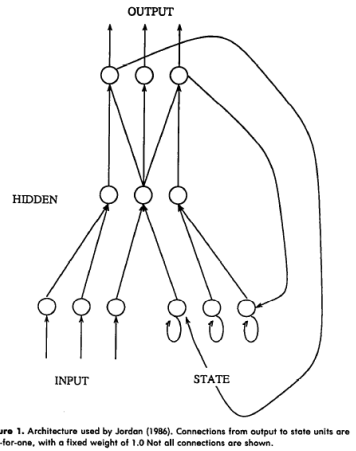
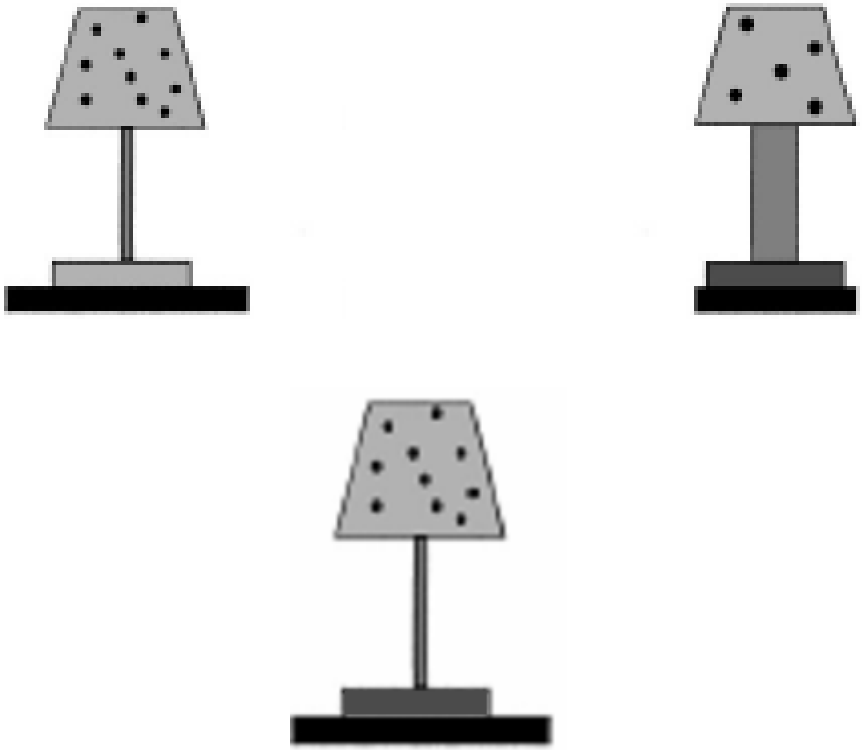


Figure 1. Architecture used by Jordan (1986). Connections from output to state units are one-for-one, with a fixed weight of 1.0 Not all connections are shown.

Components of this talk

- Computational power argument
- Recurrence and higher-order thought in free classification (including some pigeon and squirrel work).
- Recurrence and higher-order thought in contingency learning

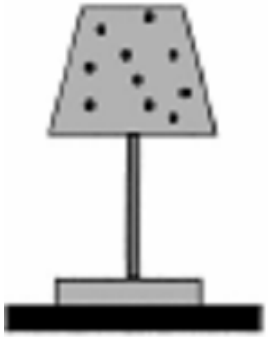
Free classification



- Match-to-standards procedure
  - Regehr & Brooks (1995); Milton & Wills (2004).

DV: UD, OS, Other

1 1 1 1



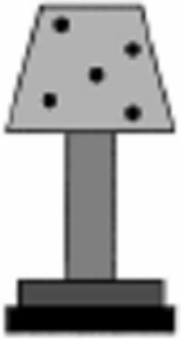
Overall similarity

1 1 1 1
0 1 1 1
1 0 1 1
1 1 0 1
1 1 1 0

Unidimensional

1 1 1 1
1 0 0 0
1 0 1 1
1 1 0 1
1 1 1 0

0 0 0 0



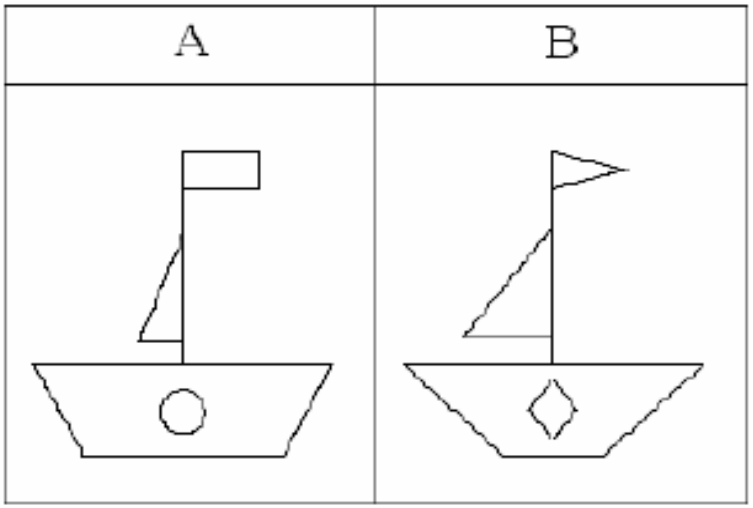
0 0 0 0
1 0 0 0
0 1 0 0
0 0 1 0
0 0 0 1

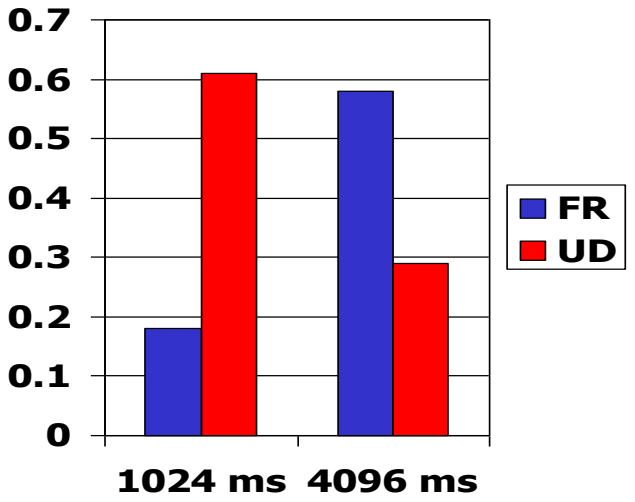
0 0 0 0
0 1 1 1
0 1 0 0
0 0 1 0
0 0 0 1

Dimensional summation hypothesis (Milton & Wills, 2004)

- Overall similarity sorting in this task occurs through a process of serial summation of evidence from each of the constituent dimensions in turn.
- It's a theory based on recurrence, and the assumption is the overall similarity sorting is more characteristic of higher-order thought than single-dimension classification.
- Contrary to some definitions of a rule-based process (e.g. Pothos, 2005)
- The hypothesis also seemed unlikely to many, given the widely-held belief that overall similarity classification is non-deliberative
  - Kemler Nelson (1983) – Intentional vs. incidental training
  - Smith & Kemler Nelson (1984) – Time pressure; concurrent load
  - Smith & Shapiro (1989) – Concurrent load.
- There are methodological problems with the above studies.

Stimulus presentation time

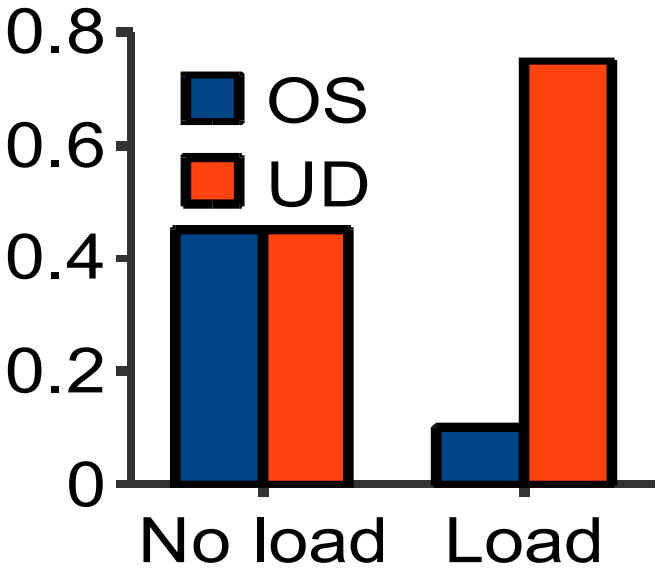


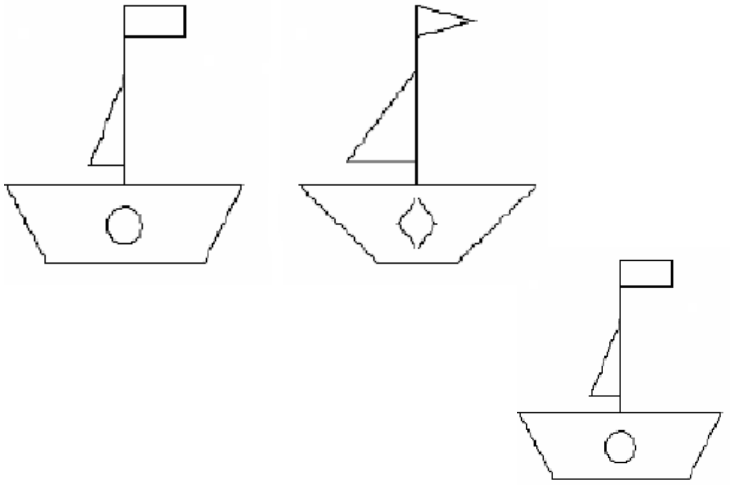


- Milton, Longmore & Wills (2008)

Concurrent load

"...11...48 ... 9 ... 87 ... 45 ... 78 ... 23 ... 91 .. 43 ... 82 ..."



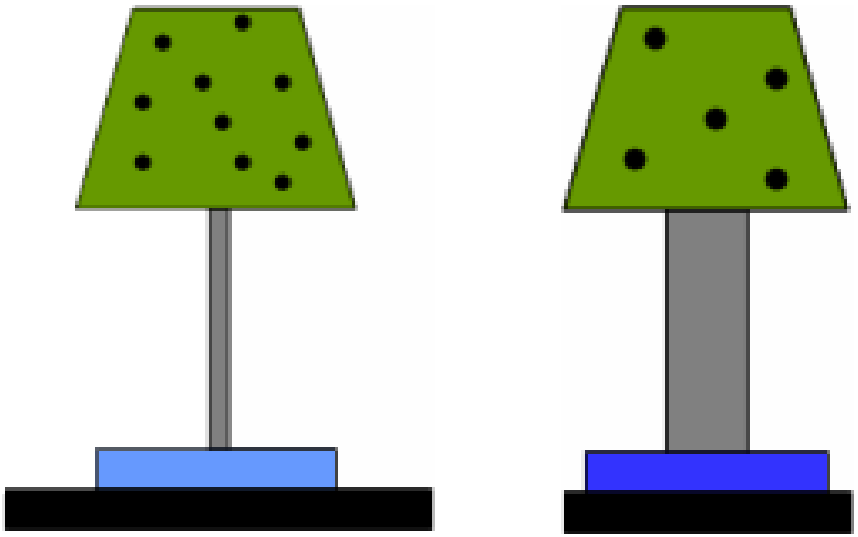


A or B?

- Wills, Longmore & Milton (under review)

Working memory capacity

- Operation span (OSPAN)
- OS sorters' mean span: 3.7
- UD sorters' mean span: 2.3



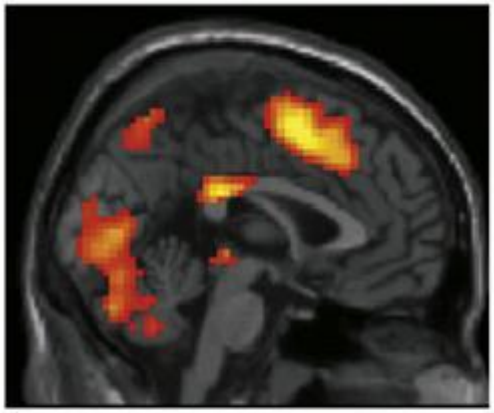
- Wills, Longmore & Milton (under review)

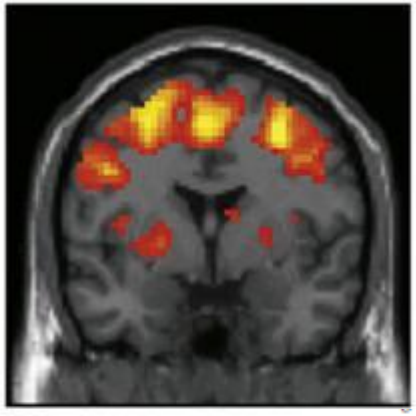
Neuroscience

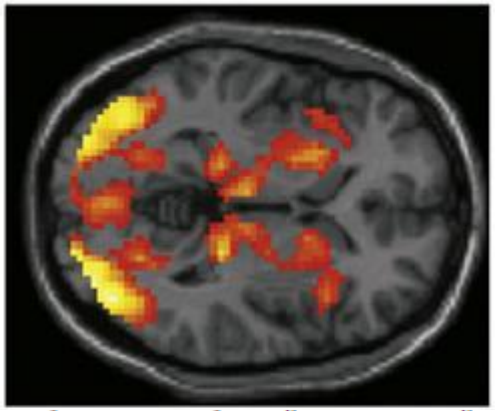
"Generally, it seemed that the authors did not meet the burden of defending the idea that FR responding is an effortful, explicit, analytic thing… They are arguing counter to the excellent literature on the cognitive neuroscience of category learning. There, it seems clear that prefrontal areas do unidimensional, rule-based things. It seems clear that the basal ganglia or the tail of the caudate do multidimensional, information-integration things".

- Anonymous JEP:LMC reviewer

Conjunction analysis





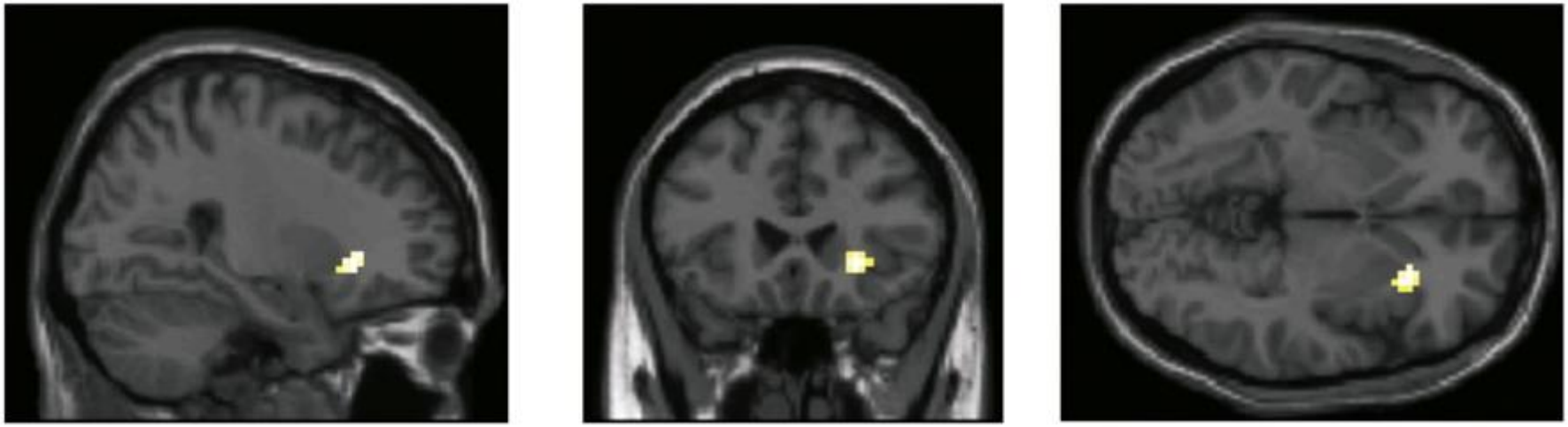


Region	Cluster size	BA	Talairach coordinates			z-score
			x	y	z	
Right posterior cerebellum	9629	–	33	−56	−12	>8.00
Right middle occipital gyrus		19	42	−78	7	7.83
Left middle occipital gyrus		19	−33	−84	18	7.76
Anterior cingulate	115	23	0	−28	24	6.23
Right sub-lobar thalamus	1145	–	21	−29	4	6.07
Left sub-lobar thalamus		–	−12	−20	9	5.88
Left brainstem		–	−6	−20	−1	5.47
Left dorsolateral frontal cortex	66	46	−42	30	23	5.17
Left dorsolateral frontal cortex		46	−48	25	26	3.24
Left sub-lobar insula	25	13	−36	−2	11	4.82
Right anterior cerebellum	29	–	6	−39	−26	4.21
Right anterior cerebellum	12	–	0	−59	−27	3.94
Left brainstem	14	–	6	−39	−26	3.76
Right middle prefrontal cortex	7	10	42	41	12	3.37

Note. BA = Brodmann's area. Indented rows indicate voxels in the same cluster as the non-indented row above them.

Milton, Wills & Hodgson (2009)

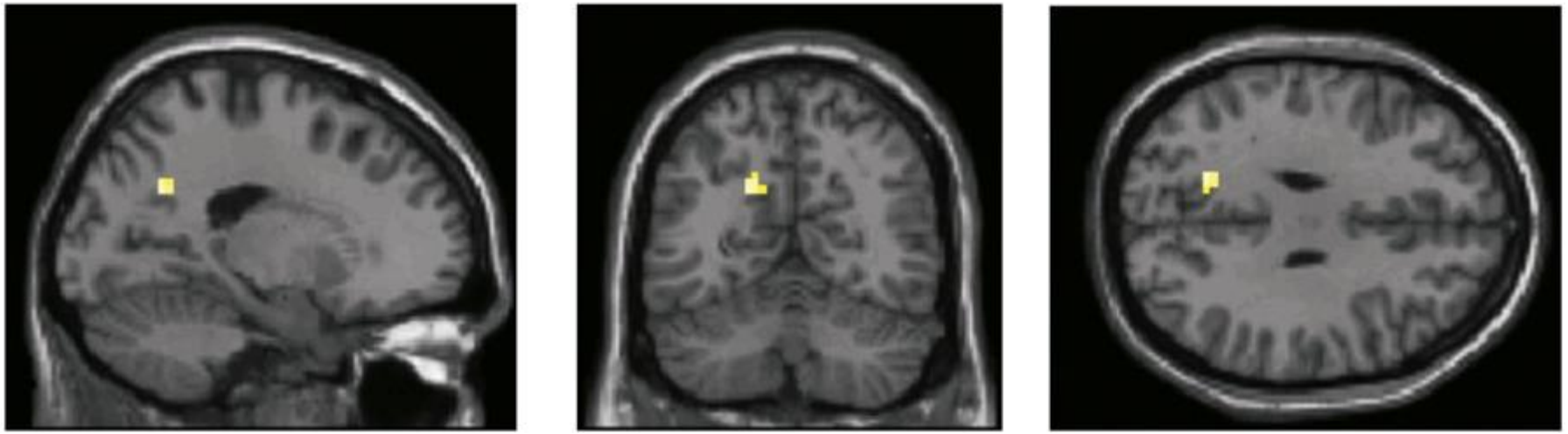
OS - UD



Includes: ventrolateral frontal cortex

Milton, Wills & Hodgson (2009)

UD - OS



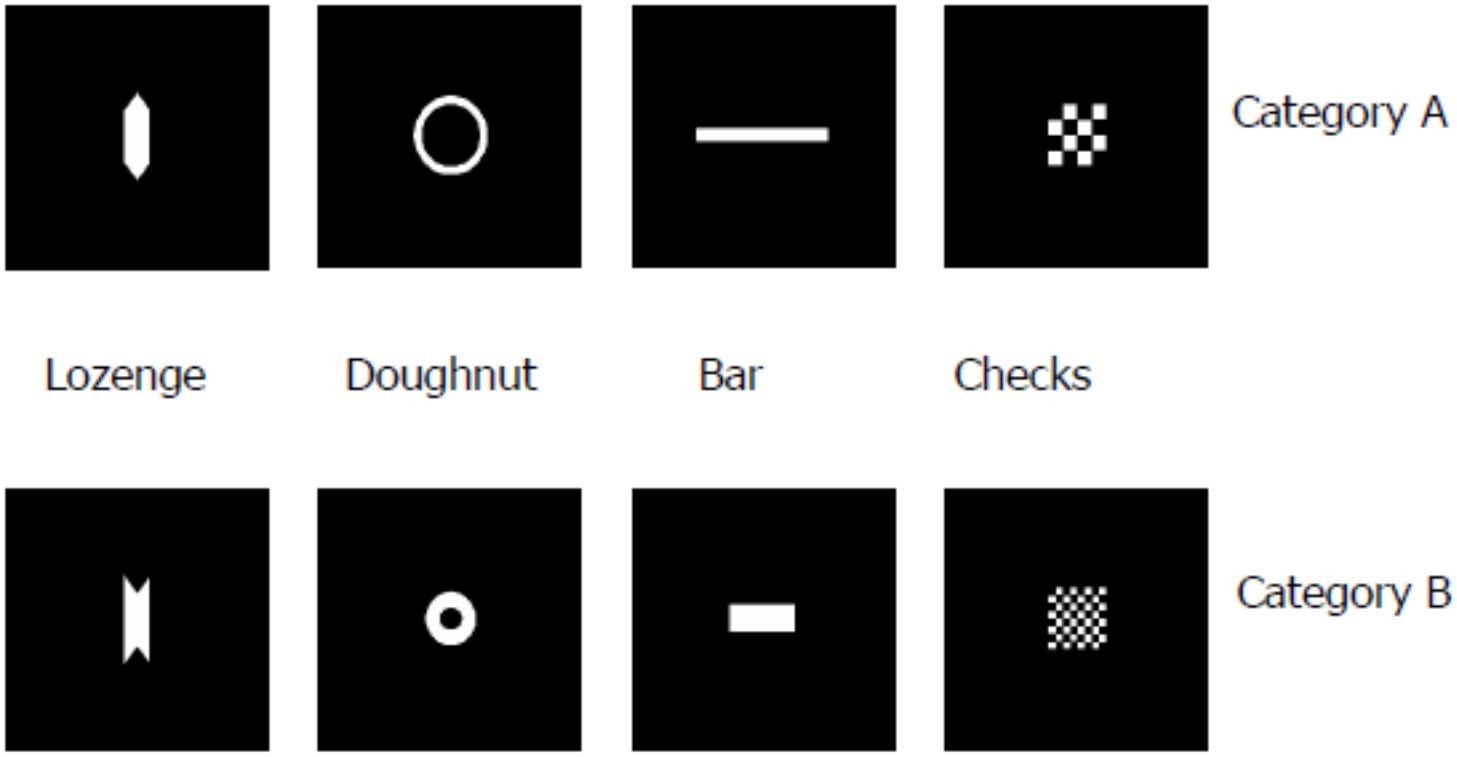
Includes: superior parietal lobe

Milton, Wills & Hodgson (2009)

Components of this talk

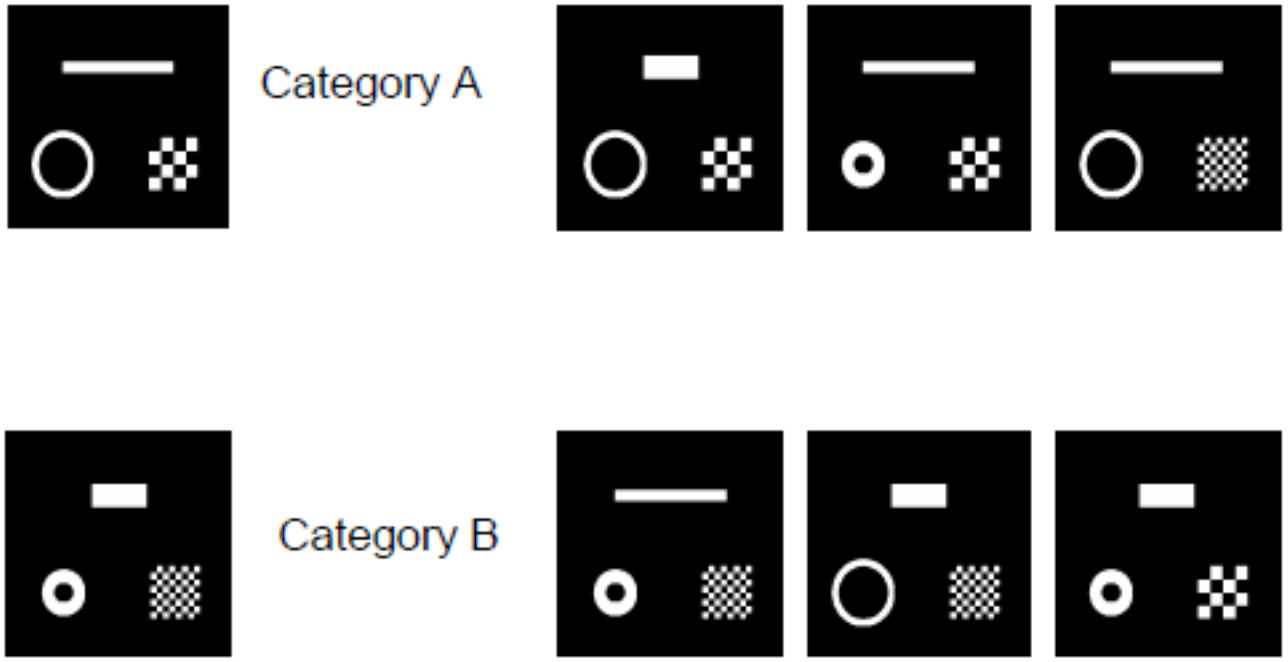
- Computational power argument
- Recurrence and higher-order thought in free classification (*including some pigeon and squirrel work*).
- Recurrence and higher-order thought in contingency learning

Free sorting in pigeons



- Wills, Lea, Leaver, Osthaus, Ryan, Suret, Bryant, Chapman & Millar (2009)

Free sorting in pigeons



- Wills, Lea, Leaver, Osthaus, Ryan, Suret, Bryant, Chapman & Millar (2009)

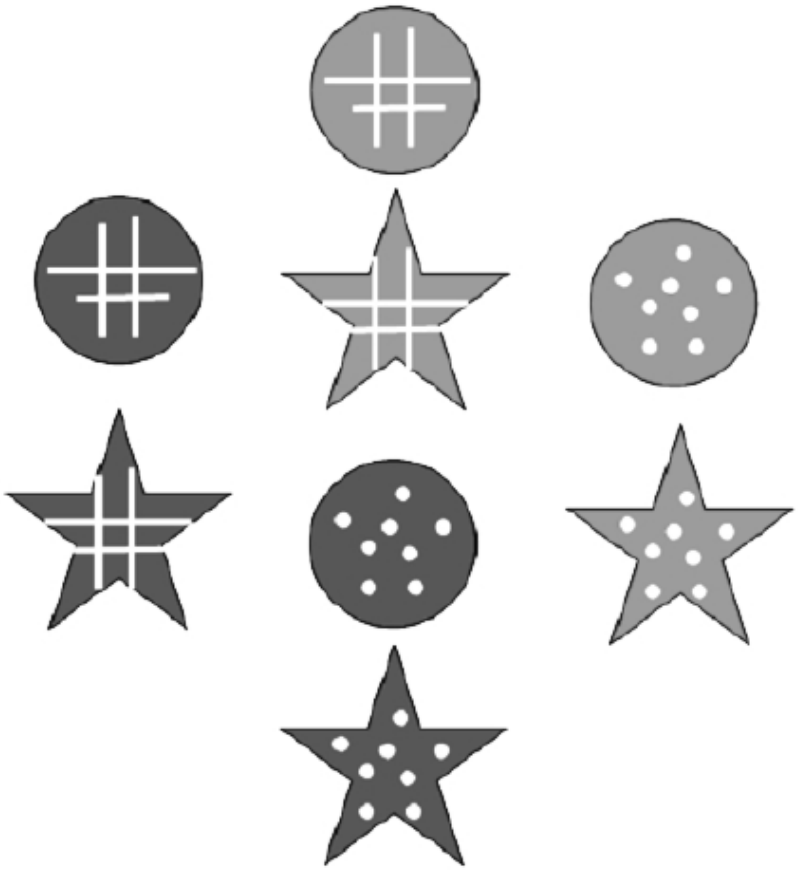
Free sorting in pigeons

Pigeons – 50% UD, 50% OS

Humans – 30% UD, 70% OS

- Wills, Lea, Leaver, Osthaus, Ryan, Suret, Bryant, Chapman & Millar (2009)

Free sorting in pigeons



Pigeons – 83% UD, 17% OS

Humans – 78% UD, 22% OS

Squirrels – 100% UD, 0% OS

- Wills, Lea, Leaver, Osthaus, Ryan, Suret, Bryant, Chapman & Millar (2009)

Components of this talk

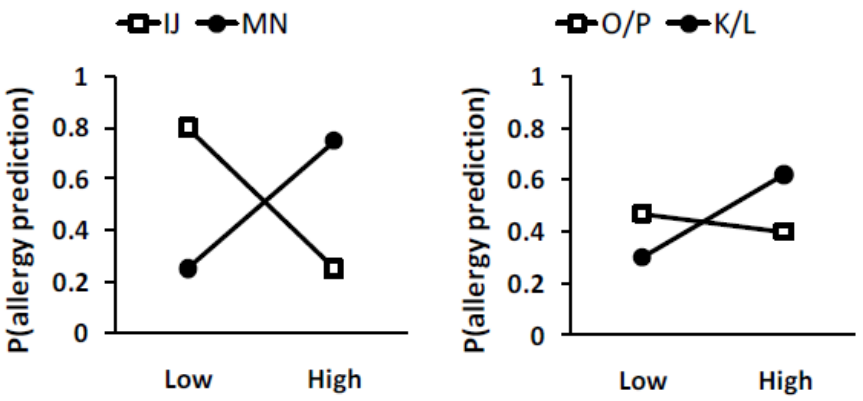
- Computational power argument
- Recurrence and higher-order thought in free classification (including some pigeon and squirrel work).
- Recurrence and higher-order thought in contingency learning

Shanks & Darby (1998)

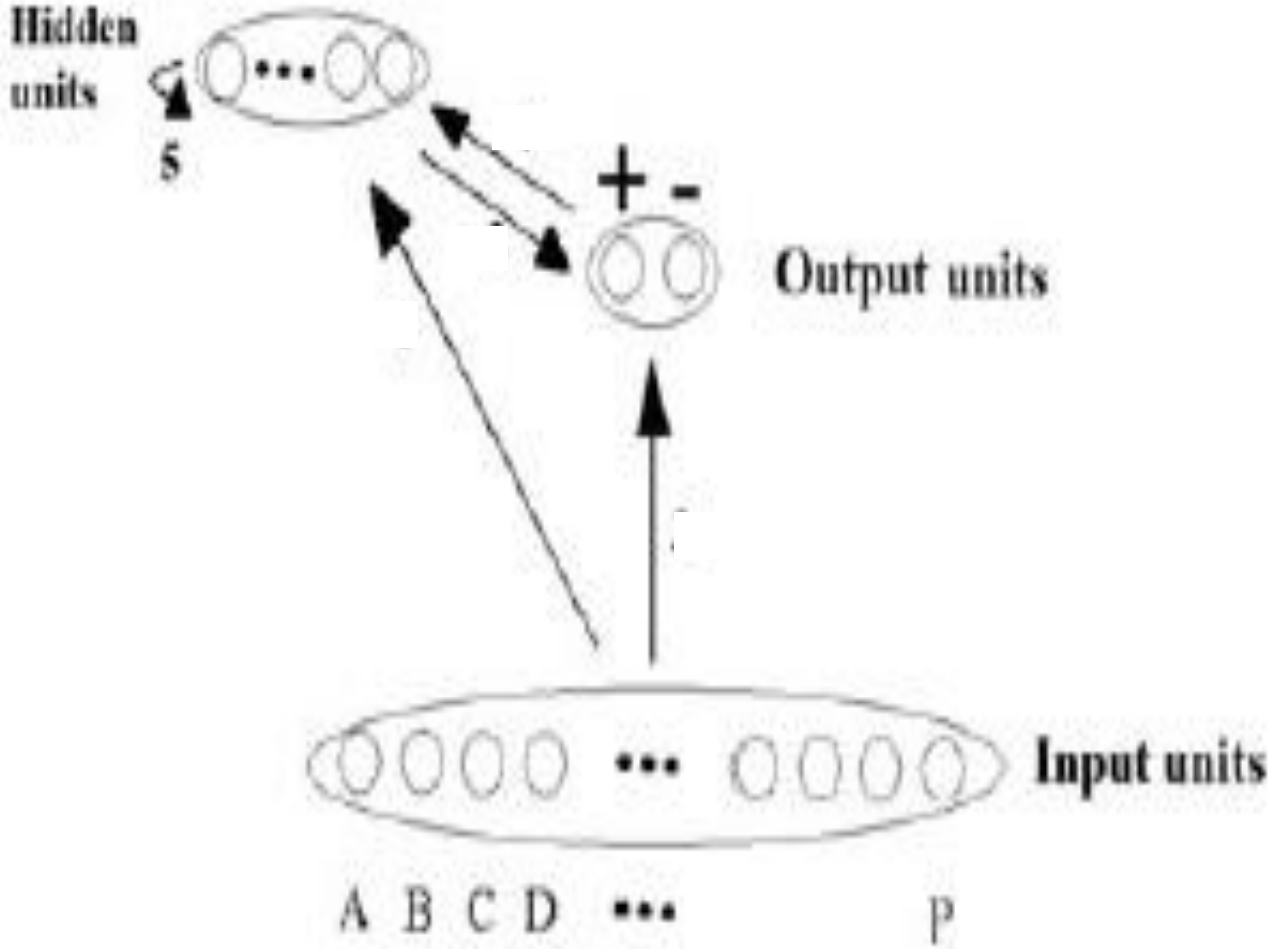
A.

Training			Test		
A+	B+	AB-	A?	B?	AB?
C-	D-	CD+	C?	D?	CD?
E+	F+	EF-	E?	F?	EF?
G-	H-	GH+	G?	H?	GH?
I+	J+		I?	J?	**IJ?**
		KL-	**K?**	**L?**	KL?
M-	N-		M?	N?	**MN?**
		OP+	**O?**	**P?**	OP?

B.

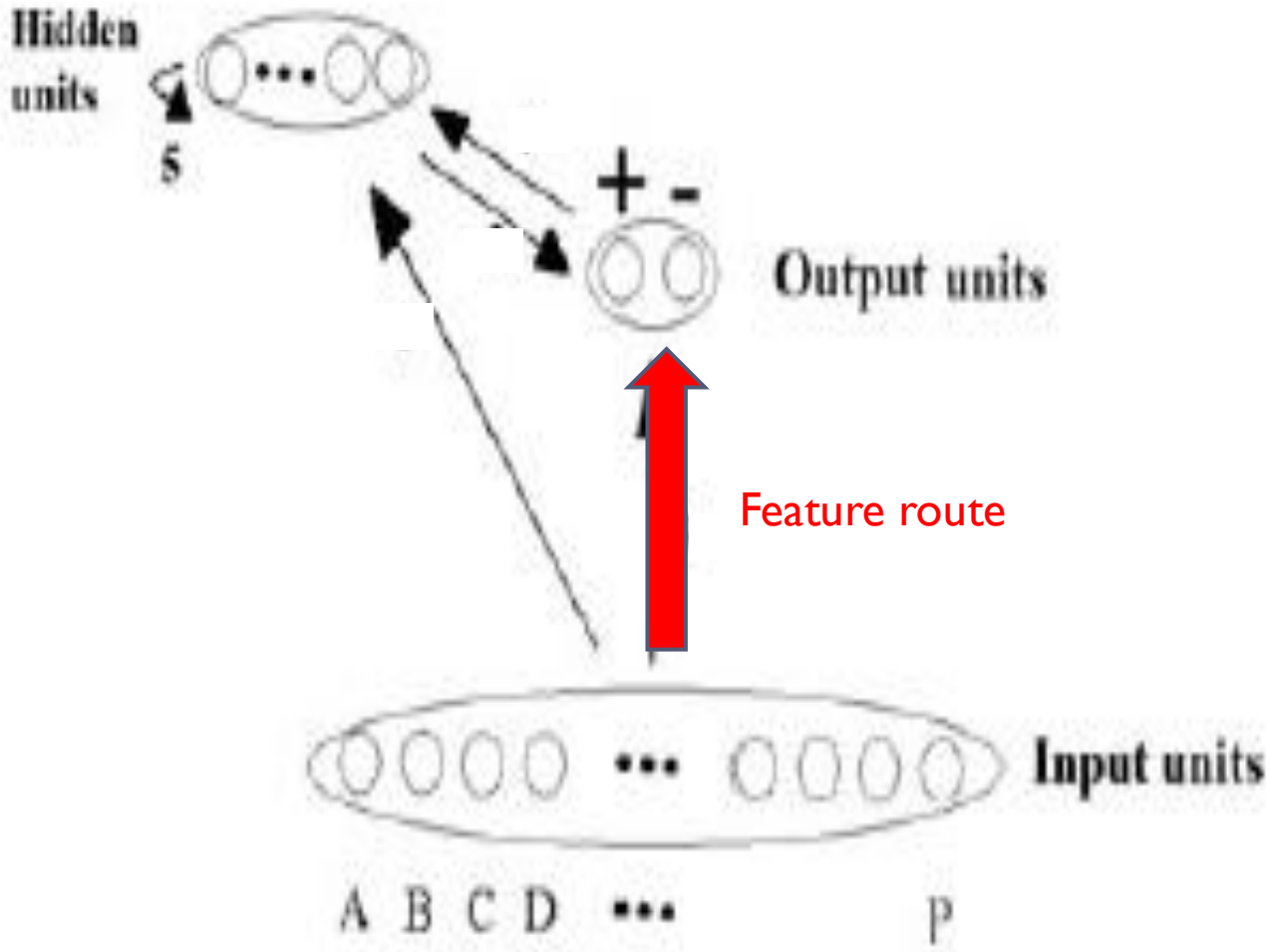


Recurrent network model of Shanks-Darby



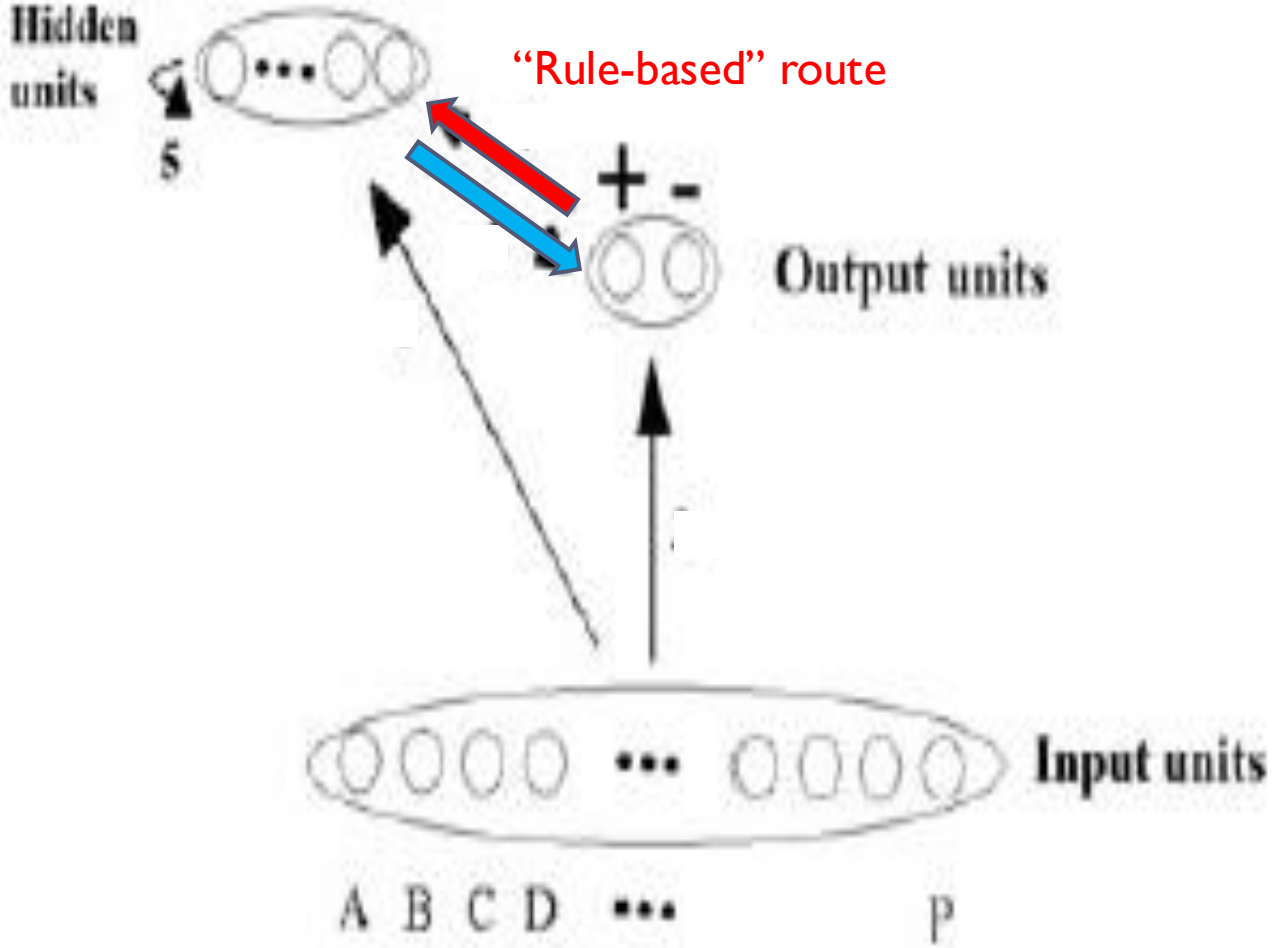
- Verguts & Fias (2009)

Recurrent network model of Shanks-Darby



- Verguts & Fias (2009)

Recurrent network model of Shanks-Darby

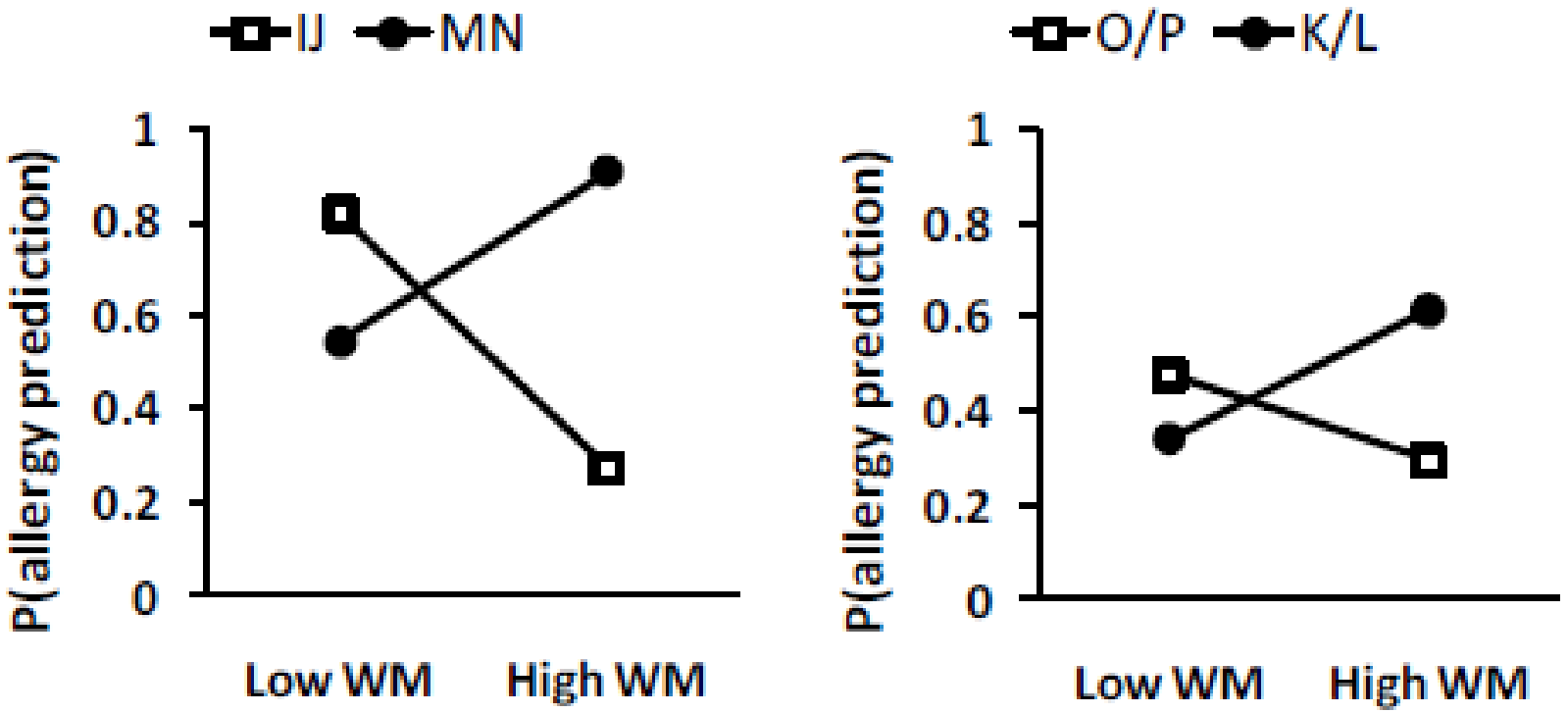


- Verguts & Fias (2009)

Predictions

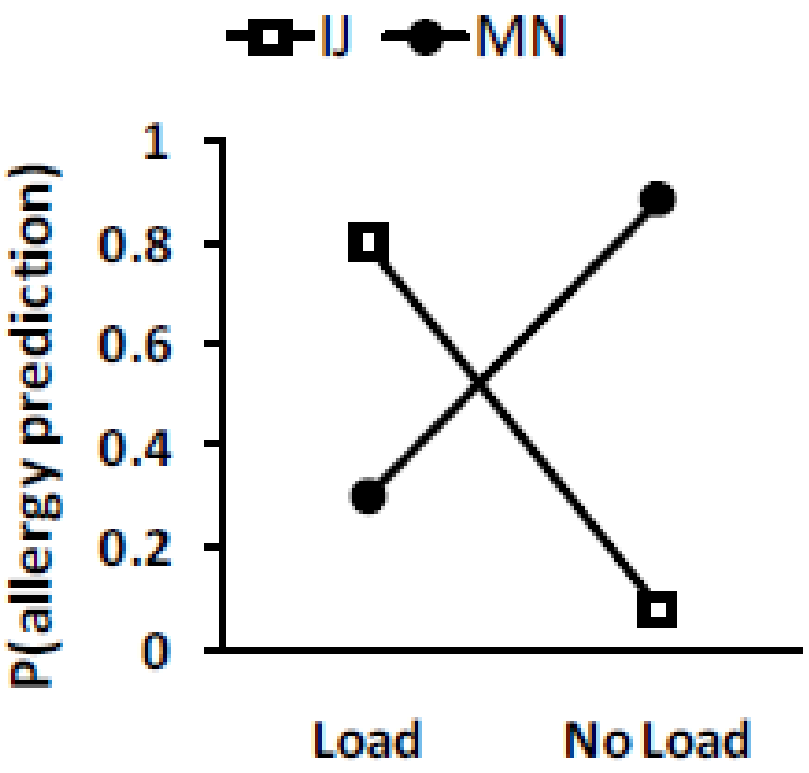
- High WMC more likely to show "rule-based" generalization than low WMC
- Rule-based generalization more likely under full attention than under concurrent load.

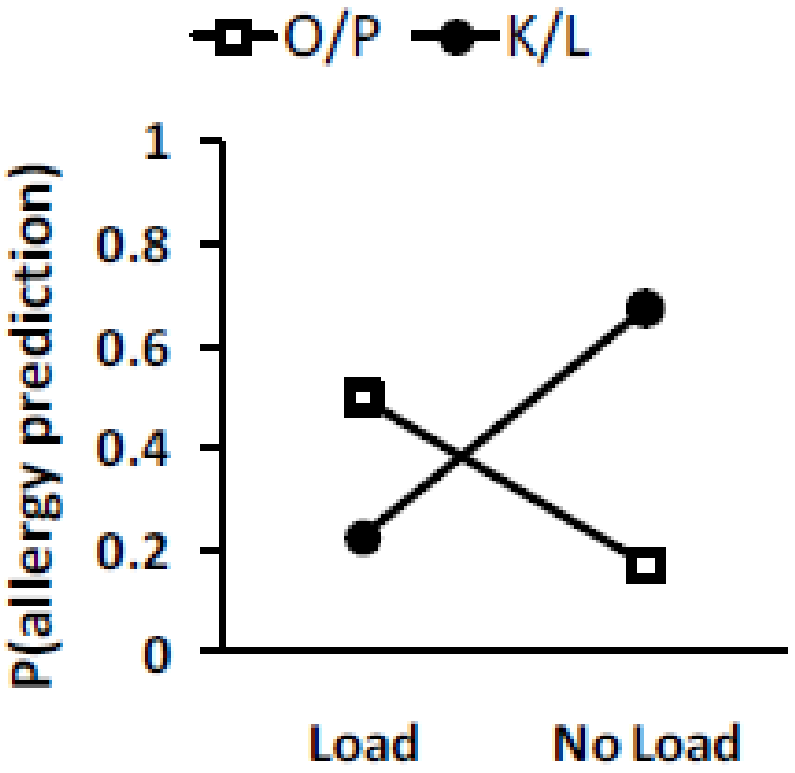
Shanks-Darby and WMC



- Wills & Barrasin (in preparation)

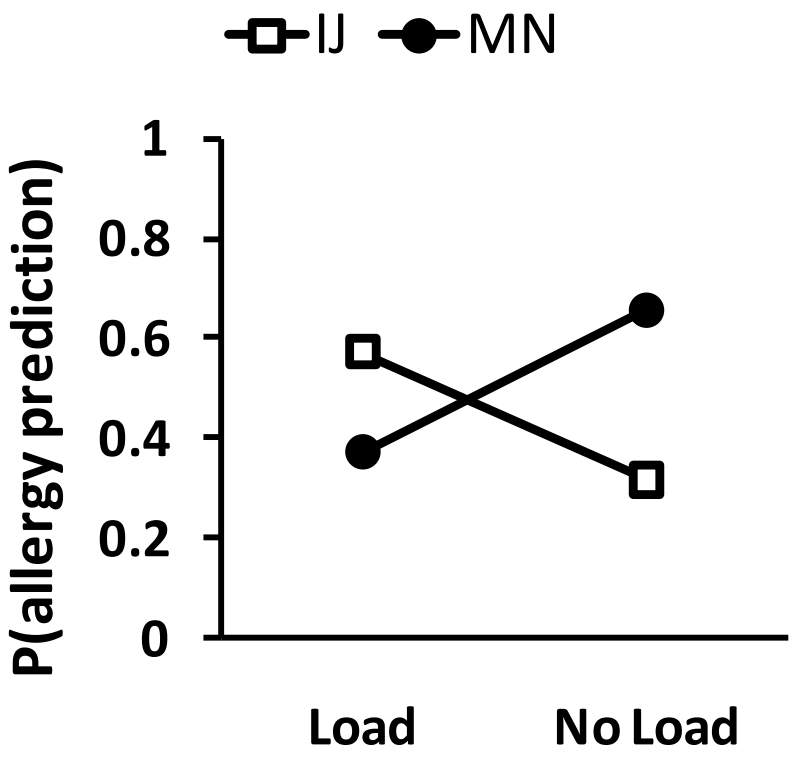
Shanks-Darby and Concurrent Load

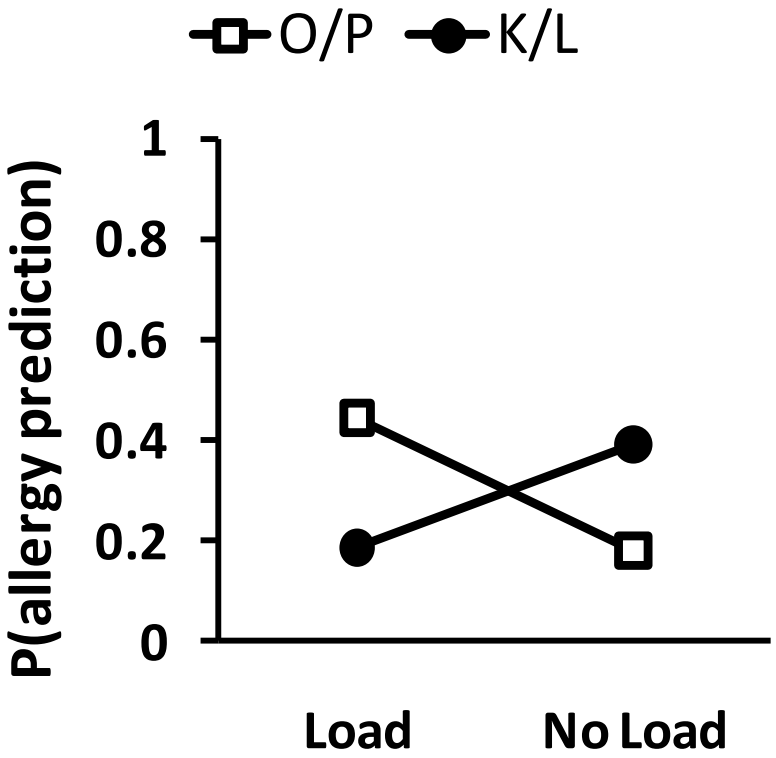




- Wills, Graham, Koh, McLaren & Rolland, (under review)

Shanks-Darby and Concurrent Load during training





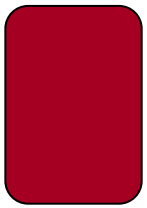
Closing remarks

- Recurrence as a characteristic process of higher-order thought
- Recurrent processes as
  - Requiring time (MTS)
  - Disrupted by concurrent load (MTS, SD)
  - Facilitated by a larger working memory capacity (MTS, SD)
  - VLFC mediated in part? (MTS)
- Unfinished business
  - Lack of a reliable species difference (MTS)
  - Shanks-Darby in pigeons?

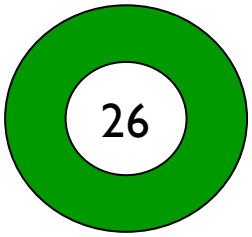
Additional slides

Turing machine

Magnetic tape (memory)



Read / write head

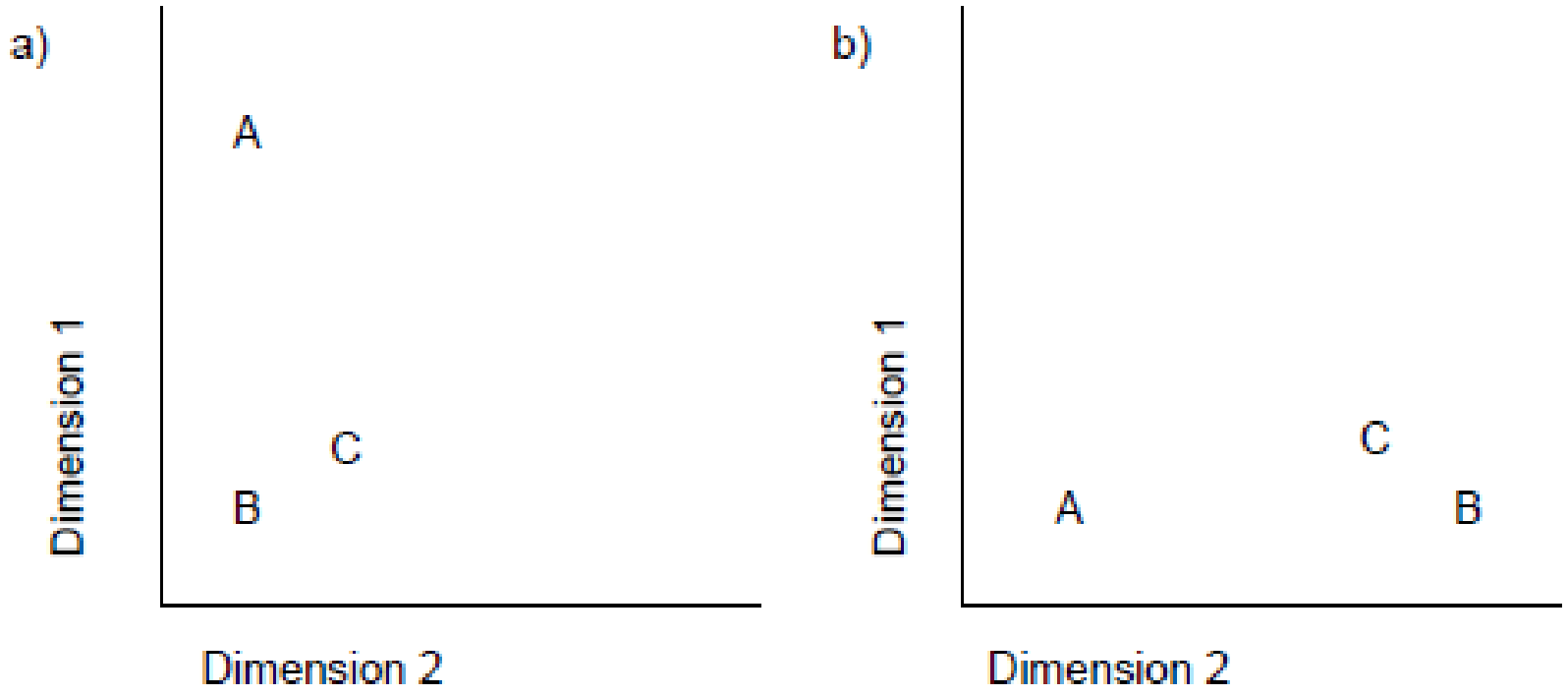


State

State	Read	Action	New state
26	0	>>	21
18	1	<<	12
09	0	Write 1	08
…	…	…	…

Program

Triad procedure



Criterial-attribute procedure

Category A			
0	0	0	0
0	1	0	0
0	0	1	0
0	0	0	1

Category B			
1	1	1	1
1	0	1	1
1	1	0	1
1	1	1	0

Test			
0	1	1	1
1	0	0	0

Information-integration procedure