Cognitive Constraints on Rule Induction

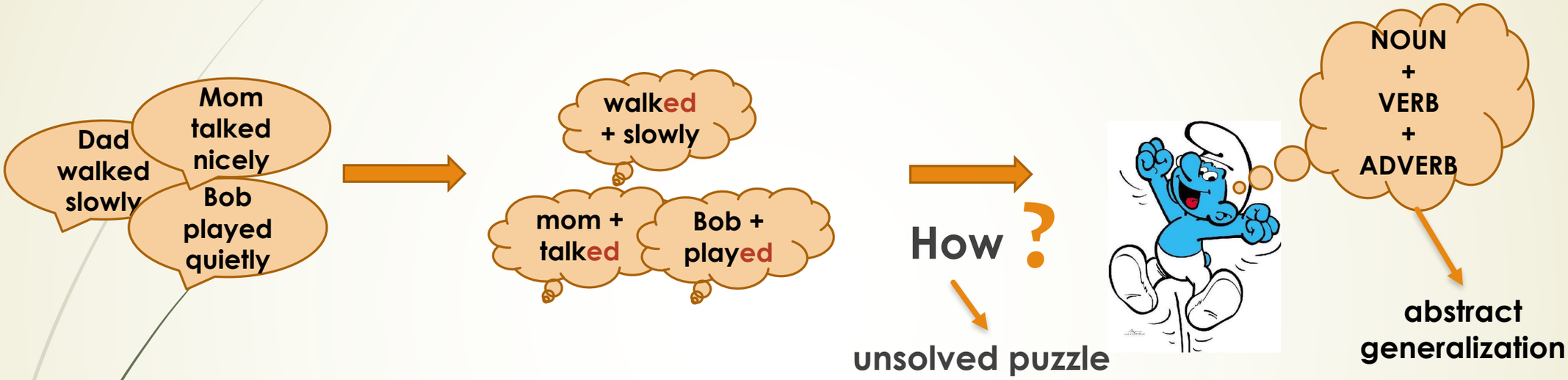
An Entropy Model

Silvia Rădulescu & Efi Giannopoulou
Sergey Avrutin, Frank Wijnen
Utrecht University

Netherlands Organisation for Scientific Research

UiL OTS
UTRECHT INSTITUTE OF LINGUISTICS

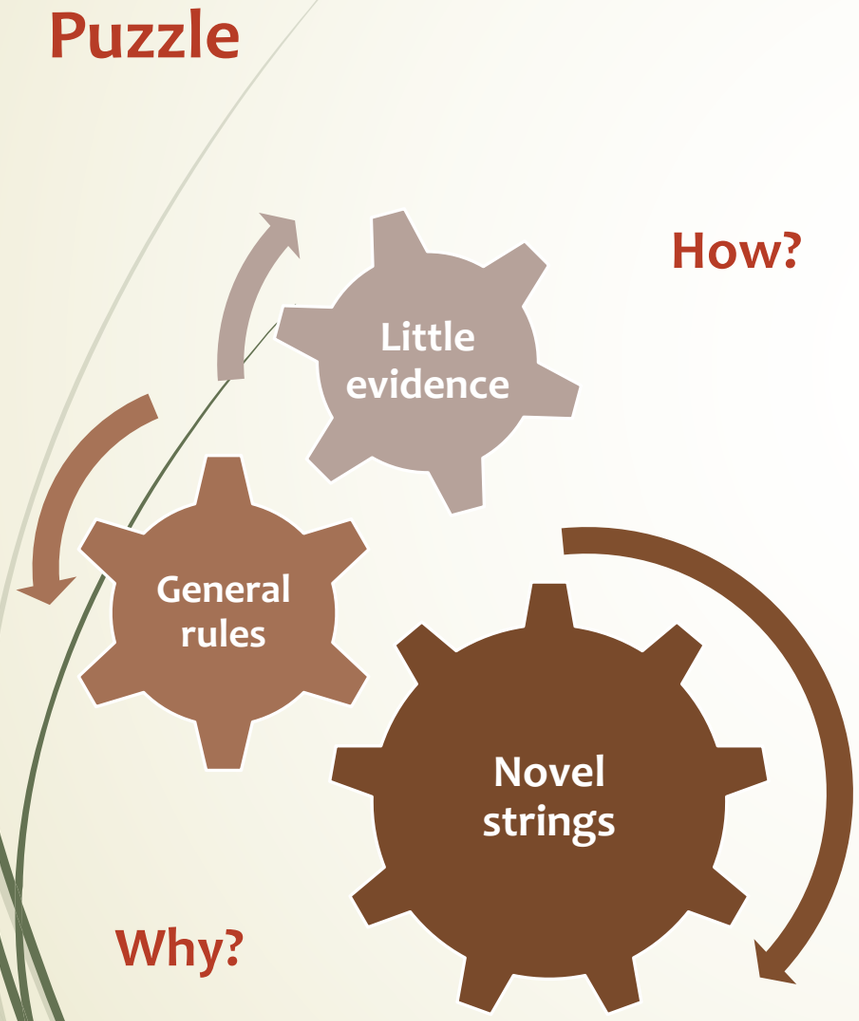
How do we make generalizations from little evidence?



? How do learners converge on abstract generalizations?

What triggers the inductive leap from memorizing to rule induction? ?

Types of Rule Induction (generalization)



Item-bound generalizations

→ relations between specific items

- e.g. *verb* + ***"-ed"***

Category-based generalizations

→ operations beyond specific items

→ over abstract categories

- e.g. **NOUN + VERB + ADVERB**

Independent underlying mechanisms ?

1. • **Statistical learning -> Item-bound generalizations**
   - *ba* follows *ba*, end in *di*
   - Saffran et al. (1996); Aslin et al. (1998)

2. • **Abstract rule learning -> Category-based generalizations**
   - *varX* follows *varX*, end in *varY*
   - Marcus et al. (1999)

OR • **Statistical learning -> BOTH item-bound & category-based generalizations**
- Aslin & Newport (2012)

OR
Phased mechanism?

- *dedeje*
- *kokoba*
- leledi

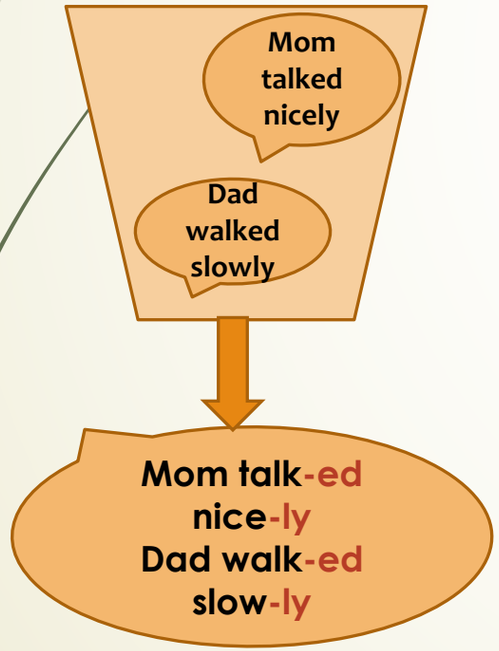
- *di* after 2x *le*
- end in je, *ba* or *di*

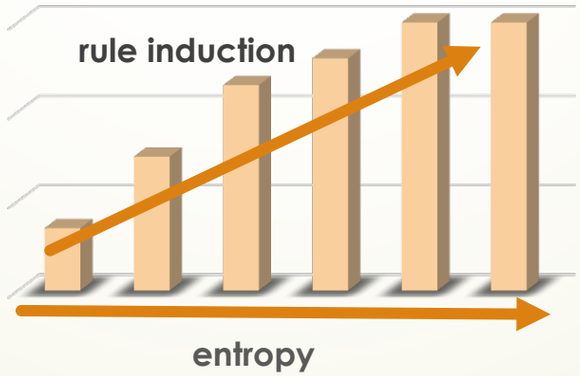
XXY

Entropy Model

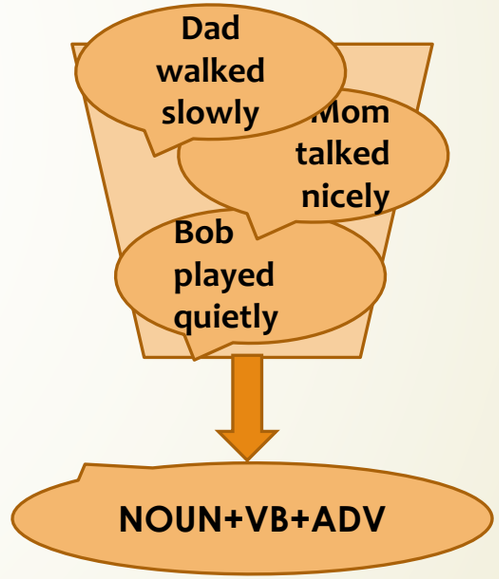
Rule Induction → the interaction between ***input complexity*** **(entropy)** and the limited encoding power of the human brain (***channel capacity***)

Low complexity (entropy) → item-bound generalizations





High complexity (entropy) → category-based generalizations



? What is the effect of **Input Complexity** on Rule Induction ? ?

Previous research

- **input complexity (variability) plays a crucial role in rule induction** √
  - Gomez (2002); Gerken (2006); Reeder, Aslin & Newport (2013)
- **NOT mere variability, BUT a specific pattern of variability**

Question

- **How can we capture this specific pattern of variability by incorporating all variables ?** ?

? What is the effect of **Input Complexity** on Rule Induction ? ?

→ vary **Input Complexity** & keep **Channel Capacity** constant

Artificial Grammar Learning - Experiment 1 + 2

- **71 adults, ~22y, ~4min, between-subjects**
- **3-syllable XXY: *goo_goo_sjie***
- **manipulated *ENTROPY* (number & frequency)**
  - **LowEN**
    - 2.8 bits (4 × 7Xs / 4 × 7Ys)
    - 3.5 bits (4 × 6Xs / 4 × 6Ys)
  - **MedEN**
    - 4 bits (2 × 12Xs / 2 × 12Ys)
    - 4.25 bits (2 × 14Xs / 2 × 14Ys)
  - **HiEN**
    - 4.58 bits (1 × 24Xs / 1 × 24Ys)
    - 4.8 bits (1 × 28Xs / 1 × 28Ys)

Test

Could this string be possible in the language that you heard?

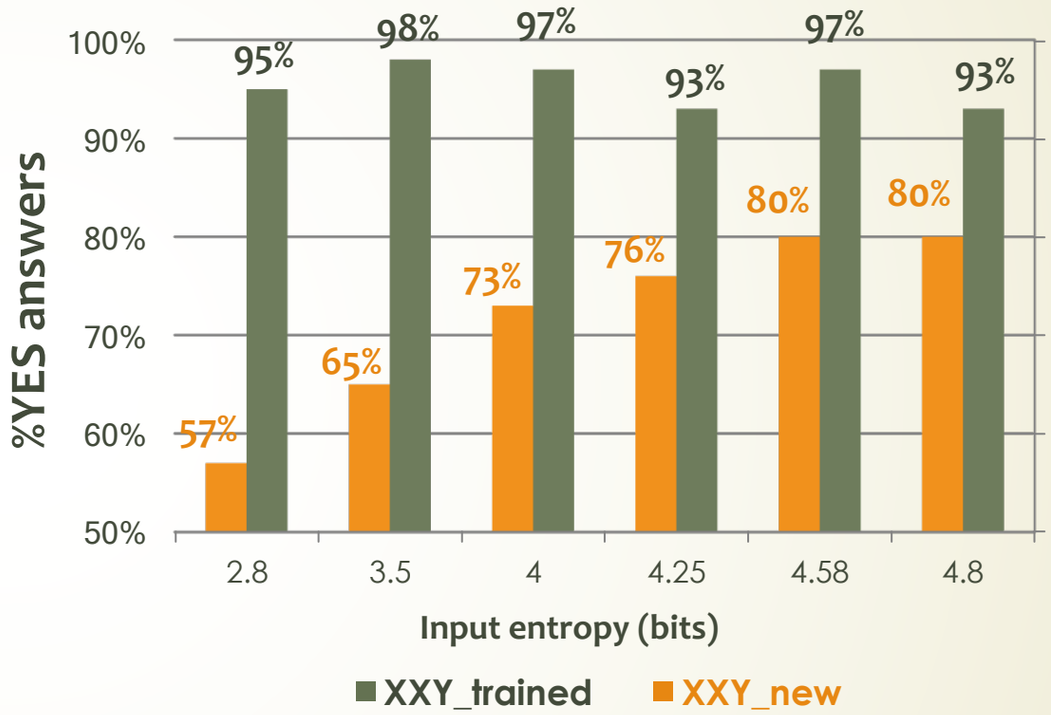
- *5 x 4 types = 20 strings*

- **XXY_new_syll:** ***too_too_suu*** √
- **XXY_trained_syll:** ***goo_goo_sjie*** √
- **XYZ_trained_syll:** ***teu_duu_saa****
- **XYZ_new_syll:** ***reu_loo_gee*** *

Results

- the higher the entropy, the higher the tendency to accept **new *XXY*** strings

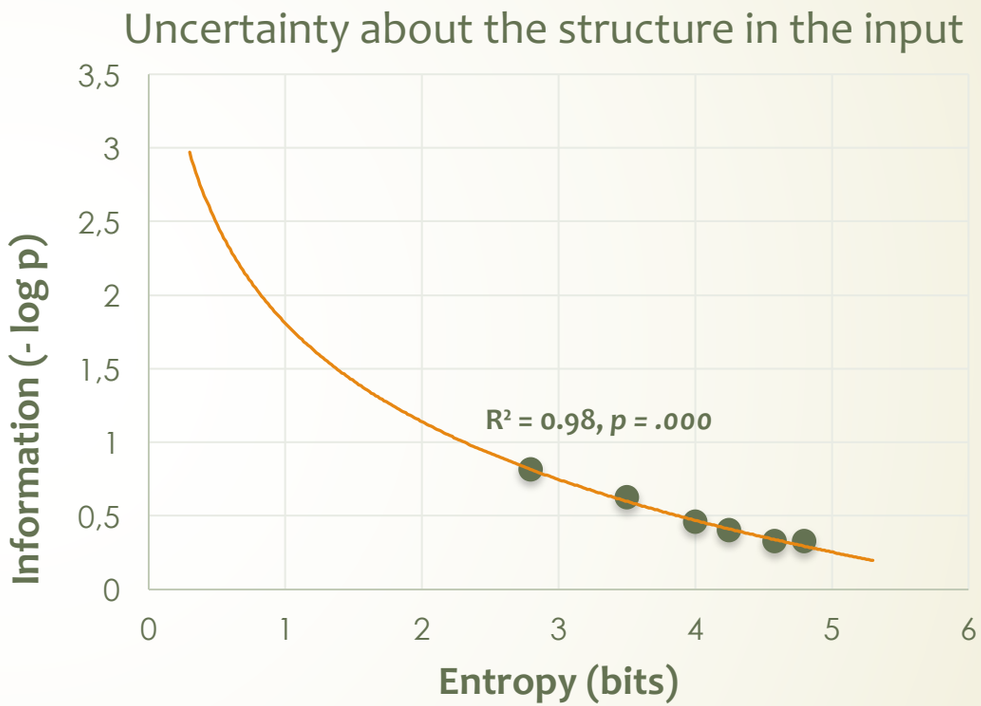
- a very similar high acceptance of ***XXY* trained** strings



Uncertainty about the structure in the input

What is information here?

→ a quantitative measure of how uncertain the mind is about the structure when exposed to a certain input entropy

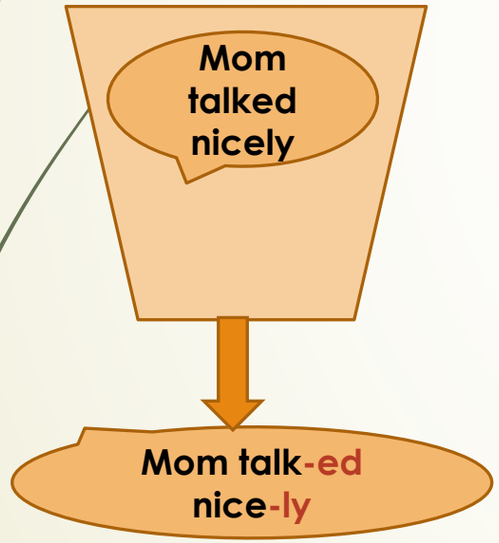


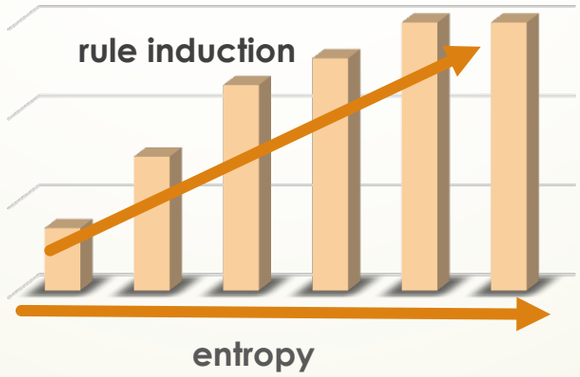
The uncertainty about structure decreases as the input entropy increases.

Entropy Model

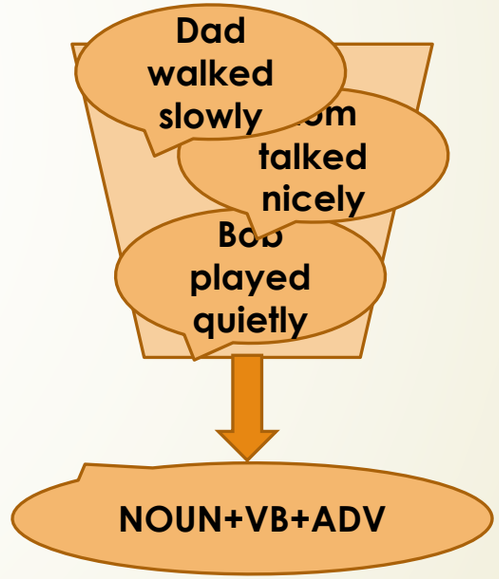
Rule Induction → interaction of ***input complexity*** **(entropy)** and ***channel capacity***

Low complexity (entropy) → item-bound generalizations





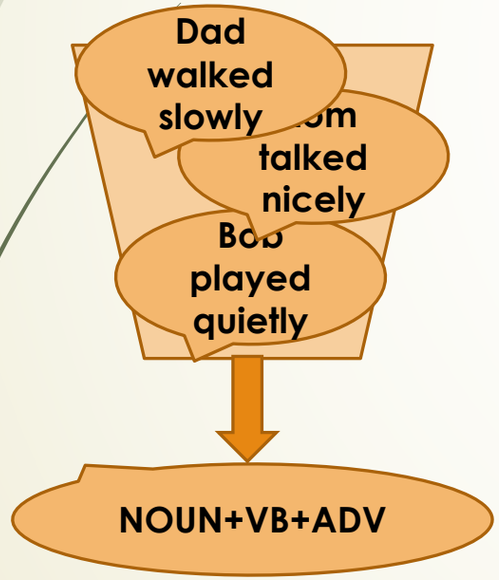
High complexity (entropy) → category-based generalizations

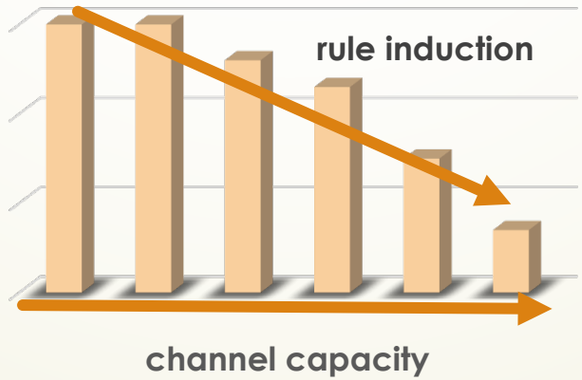


Entropy Model - hypotheses

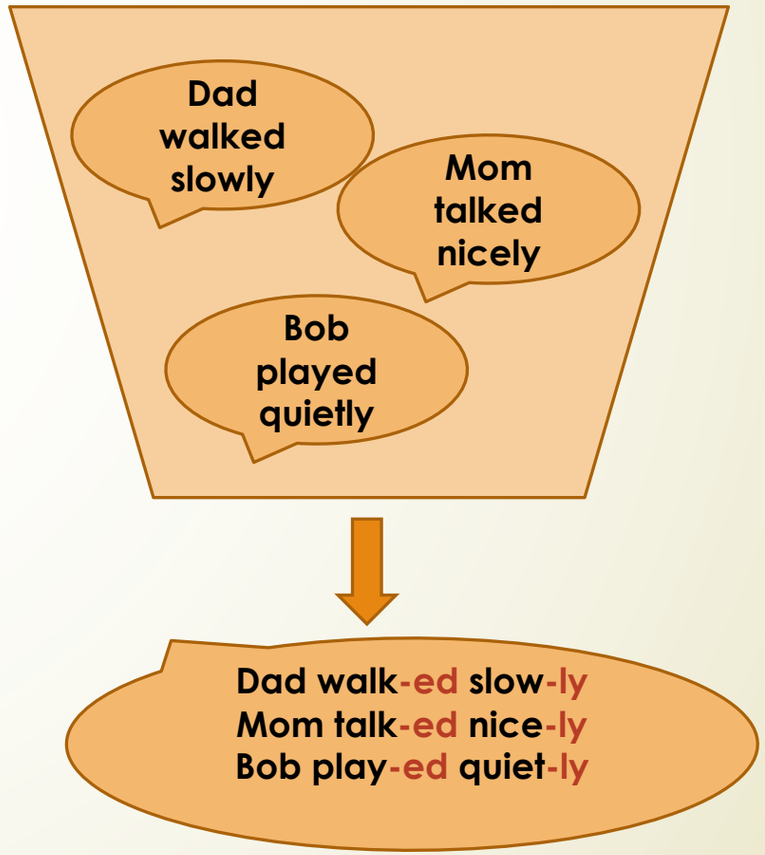
Rule Induction → interaction of ***input complexity*** **(entropy)** and ***channel capacity***

Low channel capacity →
category-based generalizations





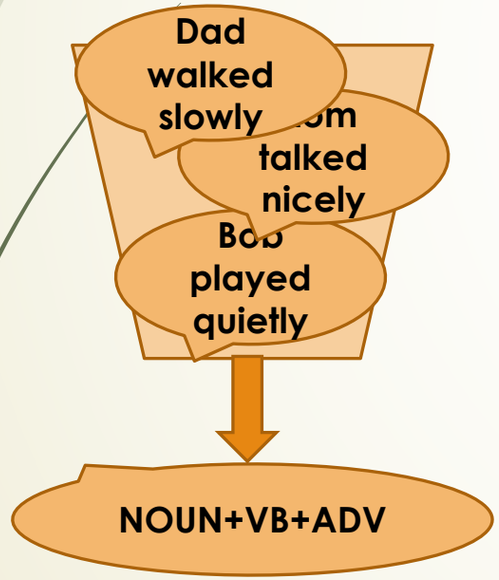
High channel capacity →
item-bound generalizations

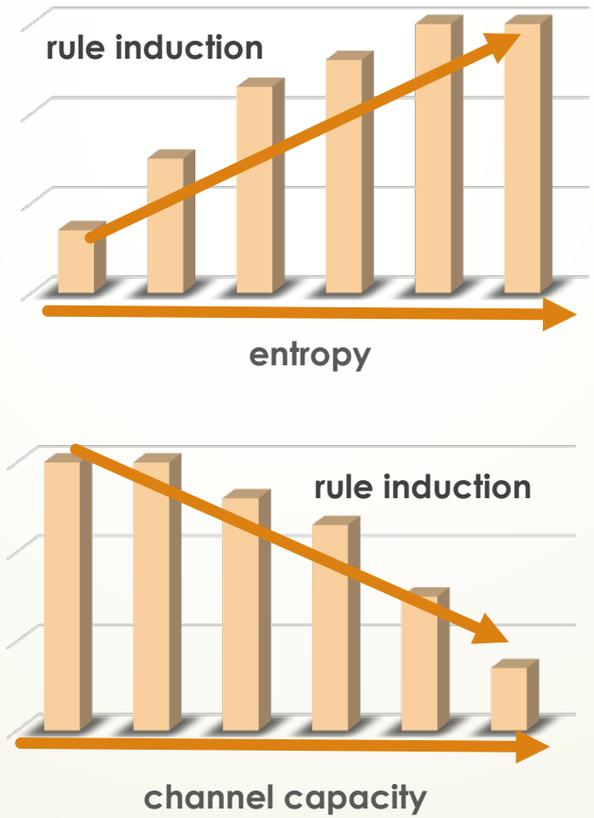


Entropy Model - hypotheses

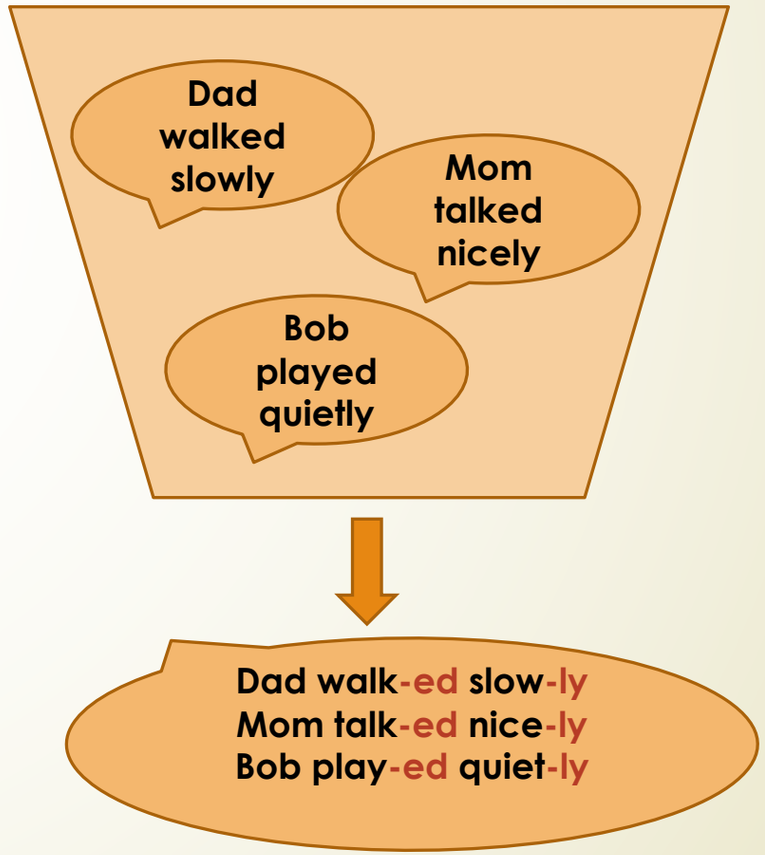
Rule Induction → interaction of ***input complexity*** **(entropy)** and ***channel capacity***

Low channel capacity →
category-based generalizations





High channel capacity →
item-bound generalizations



Effect of Channel Capacity on Rule Induction

Channel Capacity

- information-theoretic concept (entropy/time)
- model the limited encoding power of the brain

? What are the cognitive processes that modulate channel capacity? ?

- memory capacity
- pattern-recognition capacity
- attention ?
- …

? What makes us see a different picture even if the picture is the same ?

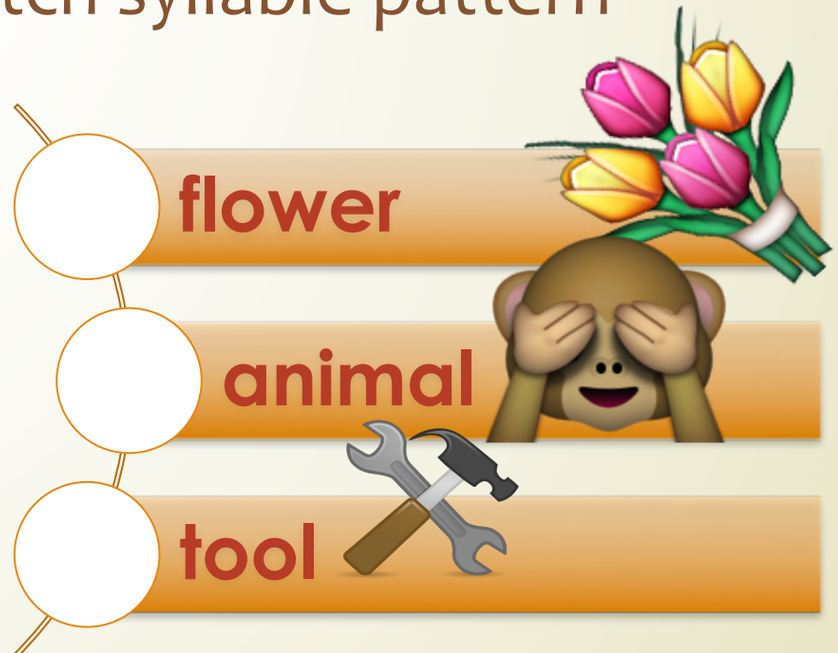
→ keep **Input complexity** constant & measure **Channel Capacity**

- **Experiment 3: 51 adults (age 19 -44; different background)**
  - **same AGL - 3-syllable XXY: *goo_goo_sjie***
- **entropy** *constant* **-> MedEn:** 2*14 X/2*14 Y (4.2 bits)
- **measure memory capacity**
  - **forward digit span**
  - **incidental memorization**
- **measure visual pattern-recognition capacity**
  - **Raven's Standard Progressive matrices**

Incidental memory task

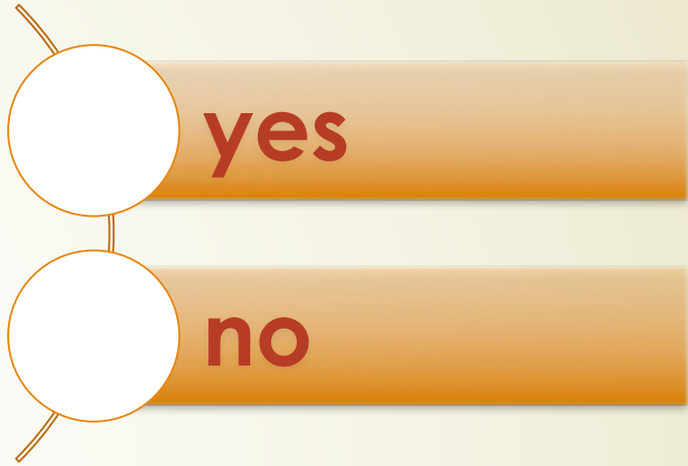
- ***Goal***
  - measure participants' capacity to memorize ***incidentally***, without having an explicit instruction to do so

- ***Training phase***
  - listen to 30 pseudo-words following a Dutch syllable pattern
    - e.g. ***go_pem, wa_dim***
  - from a forgotten language
  - What does this word sound like?

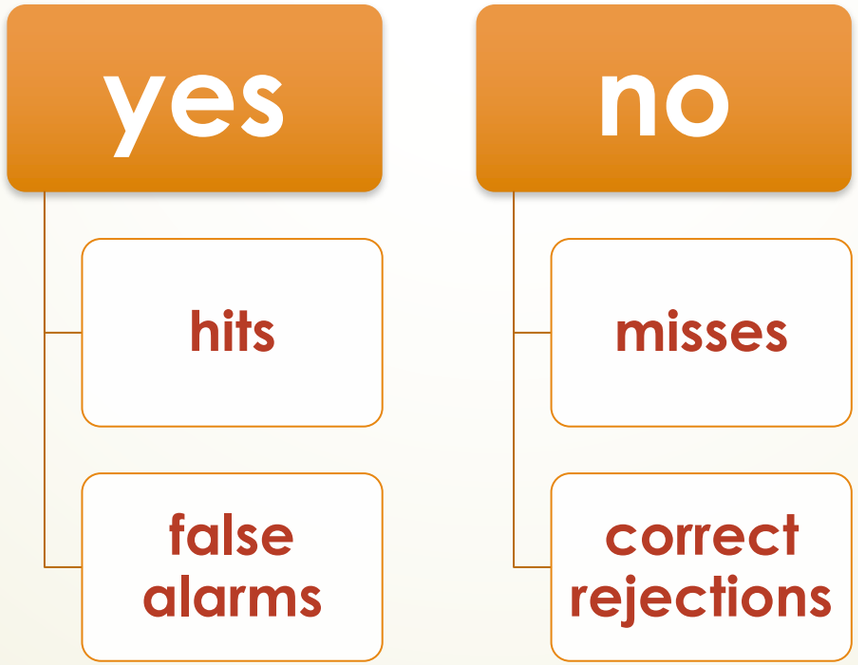


Incidental memory task

- ***Surprise Memory Test***
  - *Have you heard this word in the previous phase?*
  - *13 targets + 13 foils*

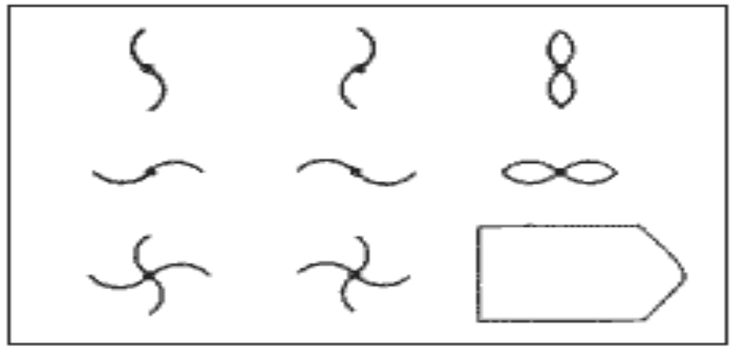


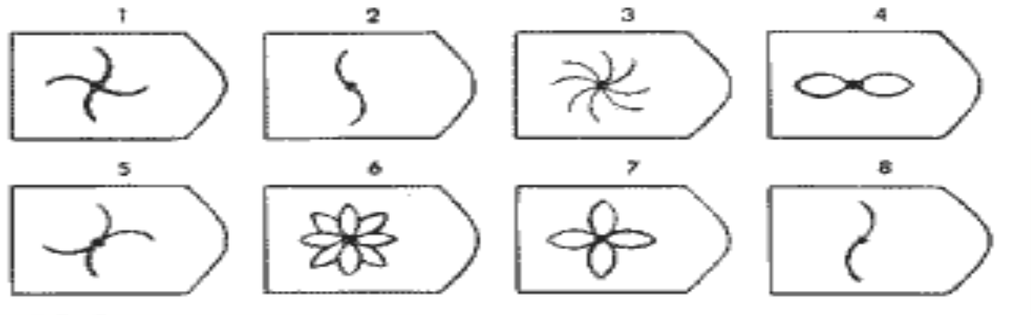
- answers

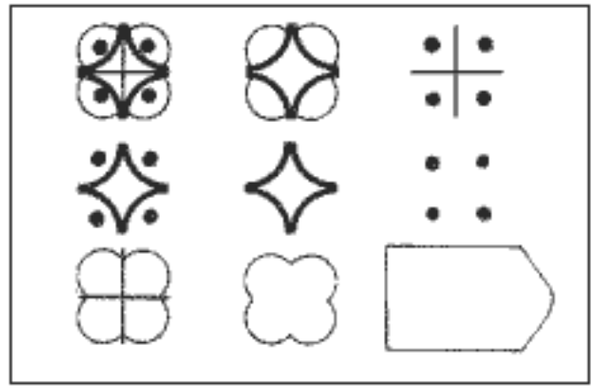


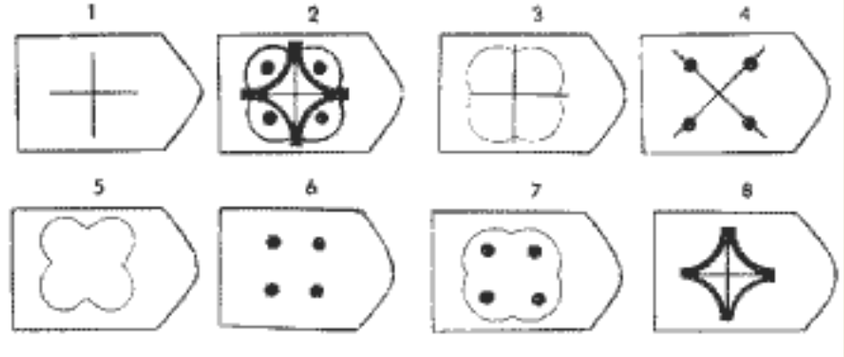
Raven's Standard Progressive Matrices

- *Visual pattern-recognition test*
  - *60 questions – logically complete patterns of shapes*
  - *Increasing difficulty*



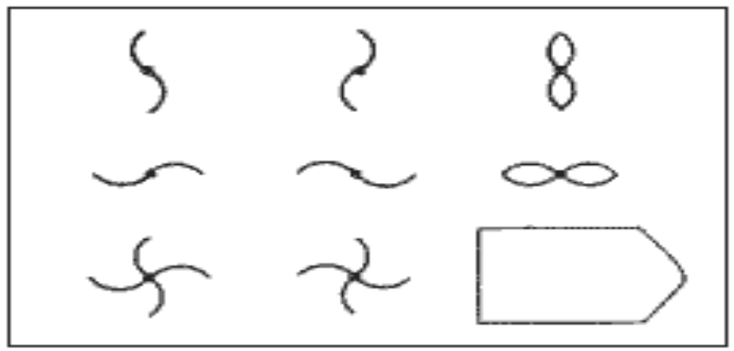


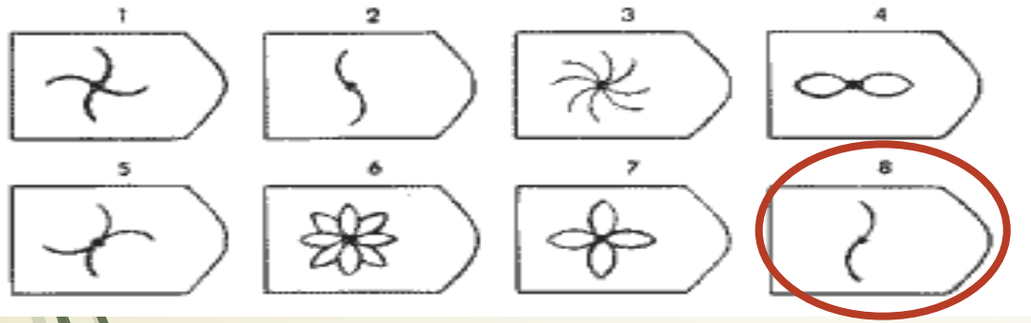


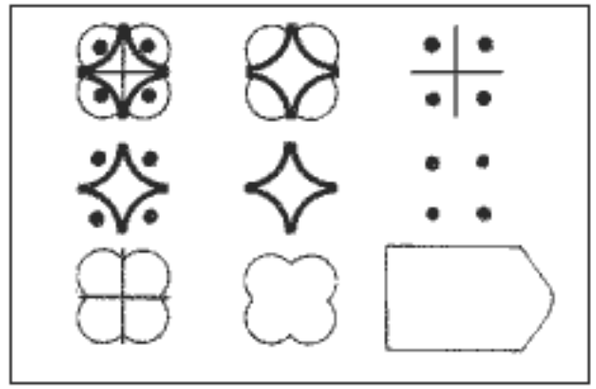


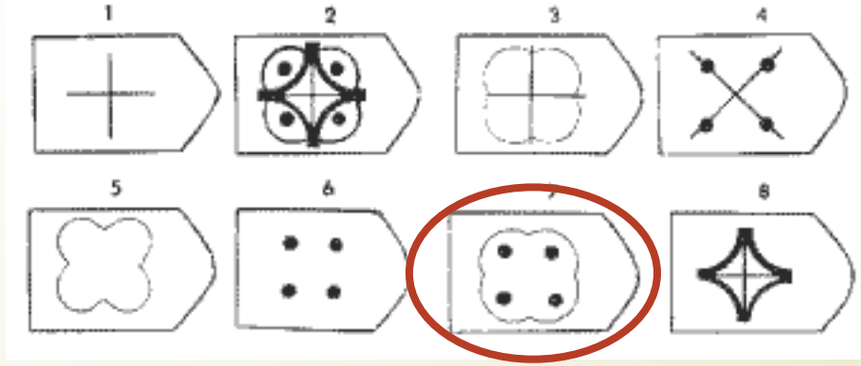
Raven's Standard Progressive Matrices

- *Visual pattern-recognition test*
  - *60 questions – logically complete patterns of shapes*
  - *Increasing difficulty*



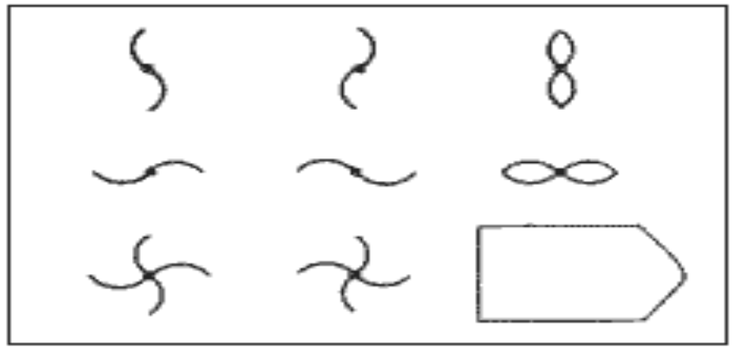


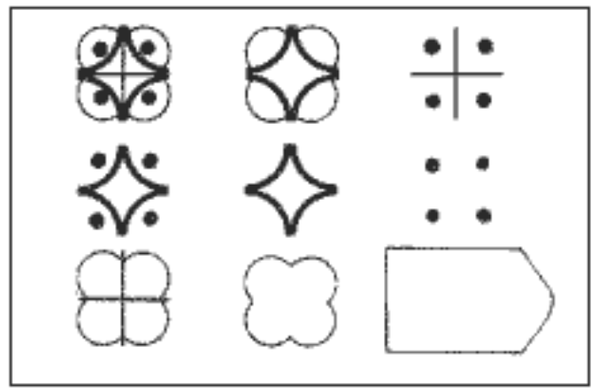


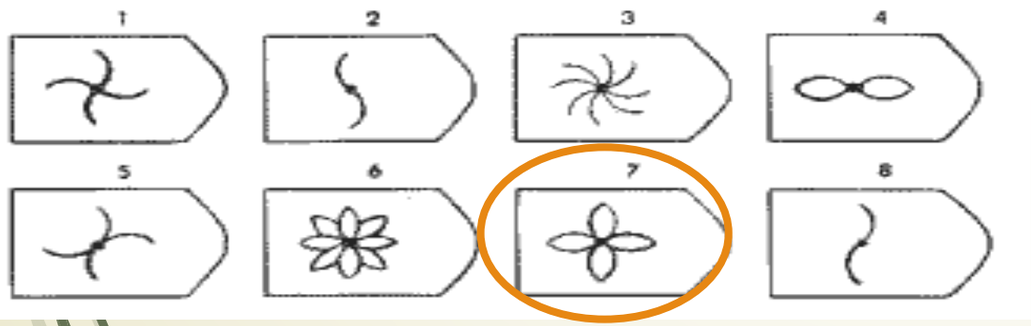


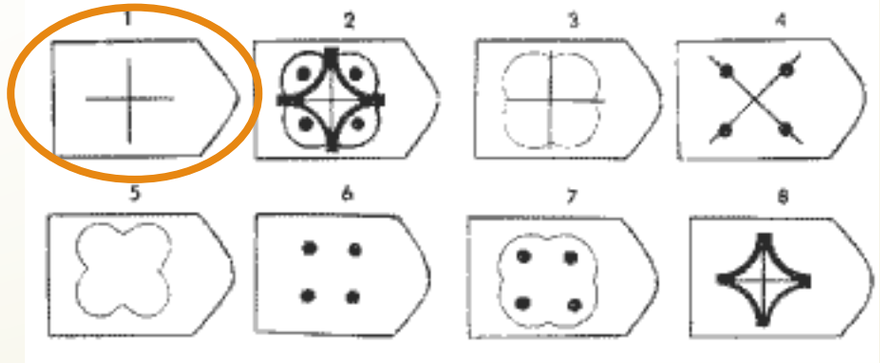
Raven's Standard Progressive Matrices

- *Visual pattern-recognition test*
  - *60 questions – logically complete patterns of shapes*
  - *Increasing difficulty*



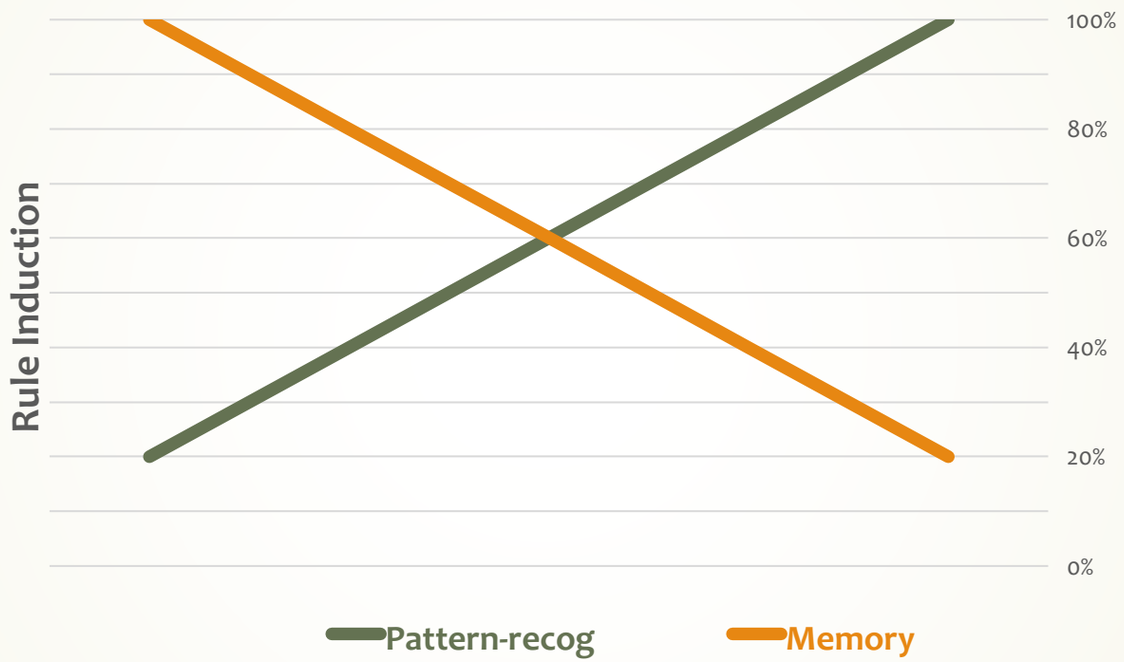






Effect of Channel Capacity on Rule Induction

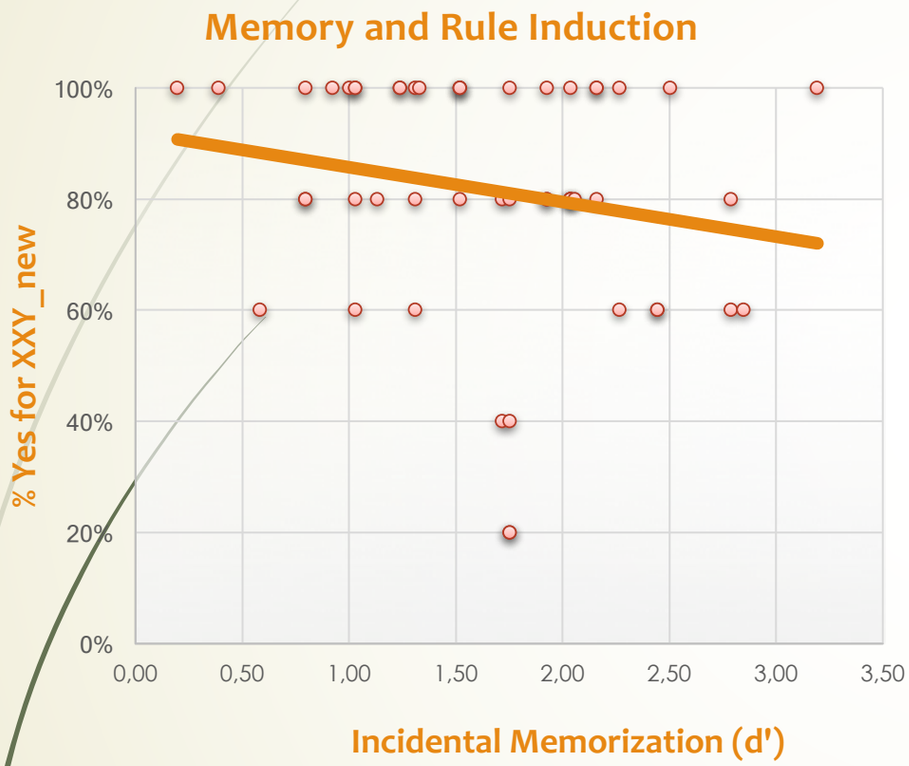
Predictions

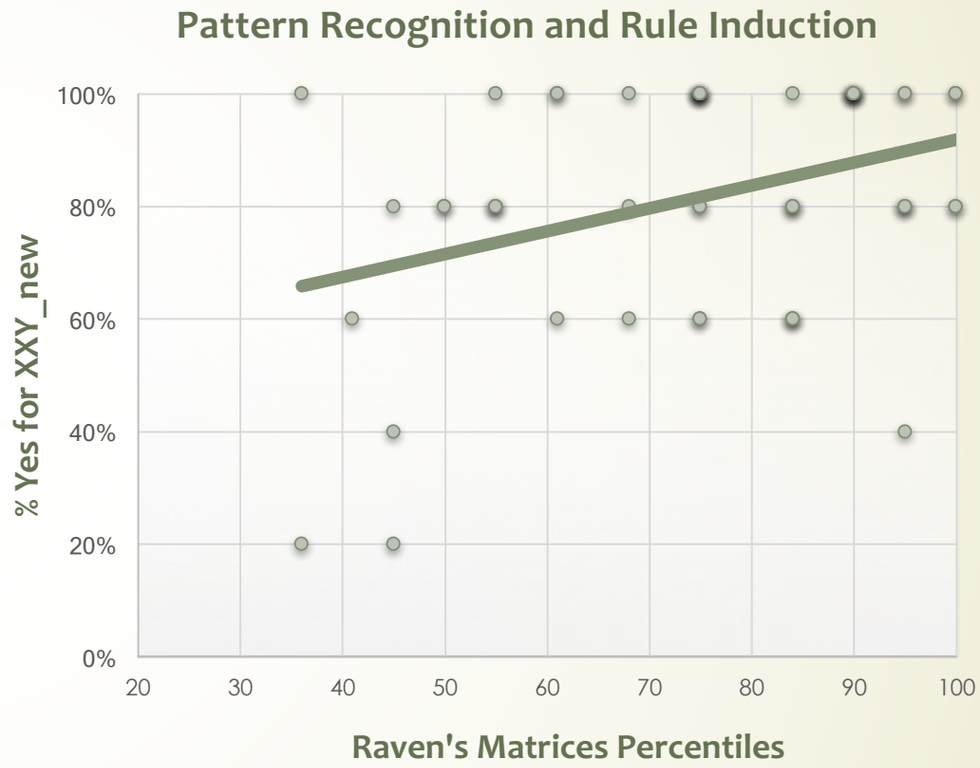


? Lower memory capacity and higher pattern-recognition capacity increase tendency to generalize ?

Effect of Channel Capacity on Rule Induction

Results: XXY new



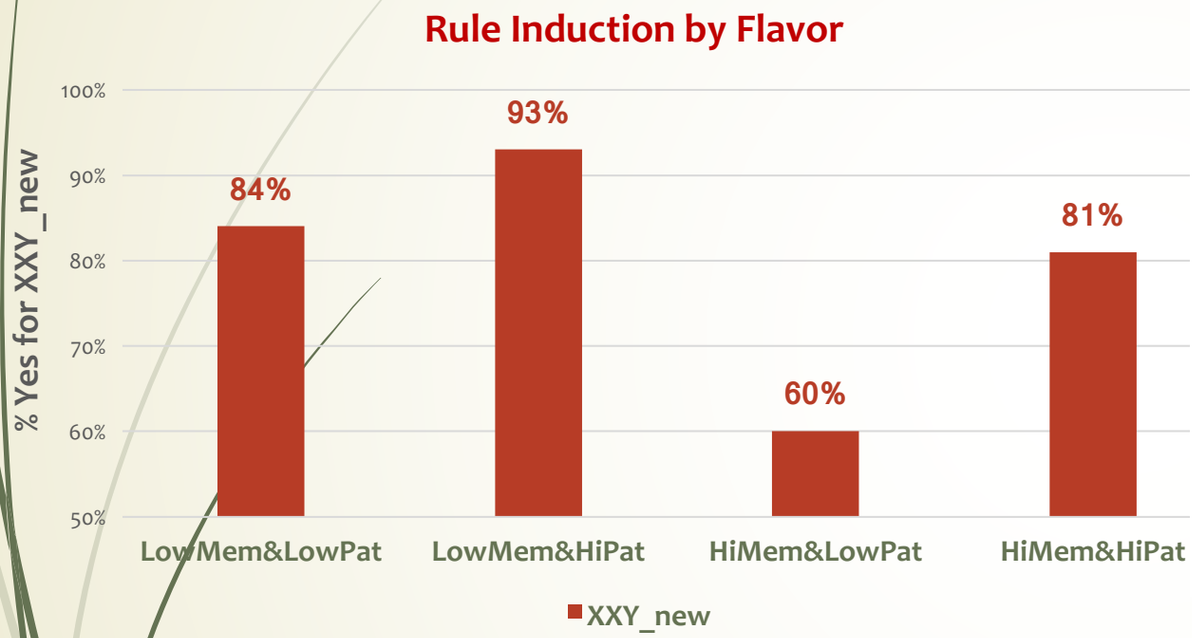


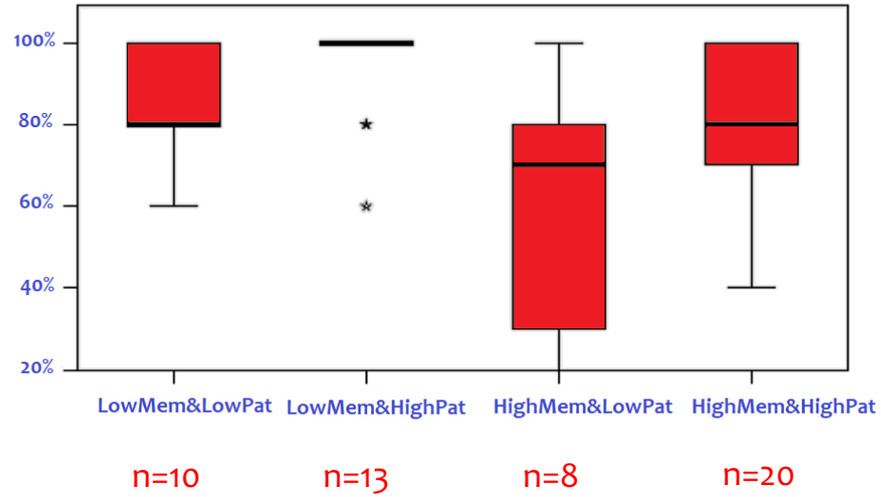
Lower incidental memory
Higher visual pattern-recognition

} predict higher tendency to generalize

Effect of Channel Capacity on Rule Induction

Results: new XXY





People with lower memory capacity AND higher visual pattern-recognition capacity have the highest tendency to generalize

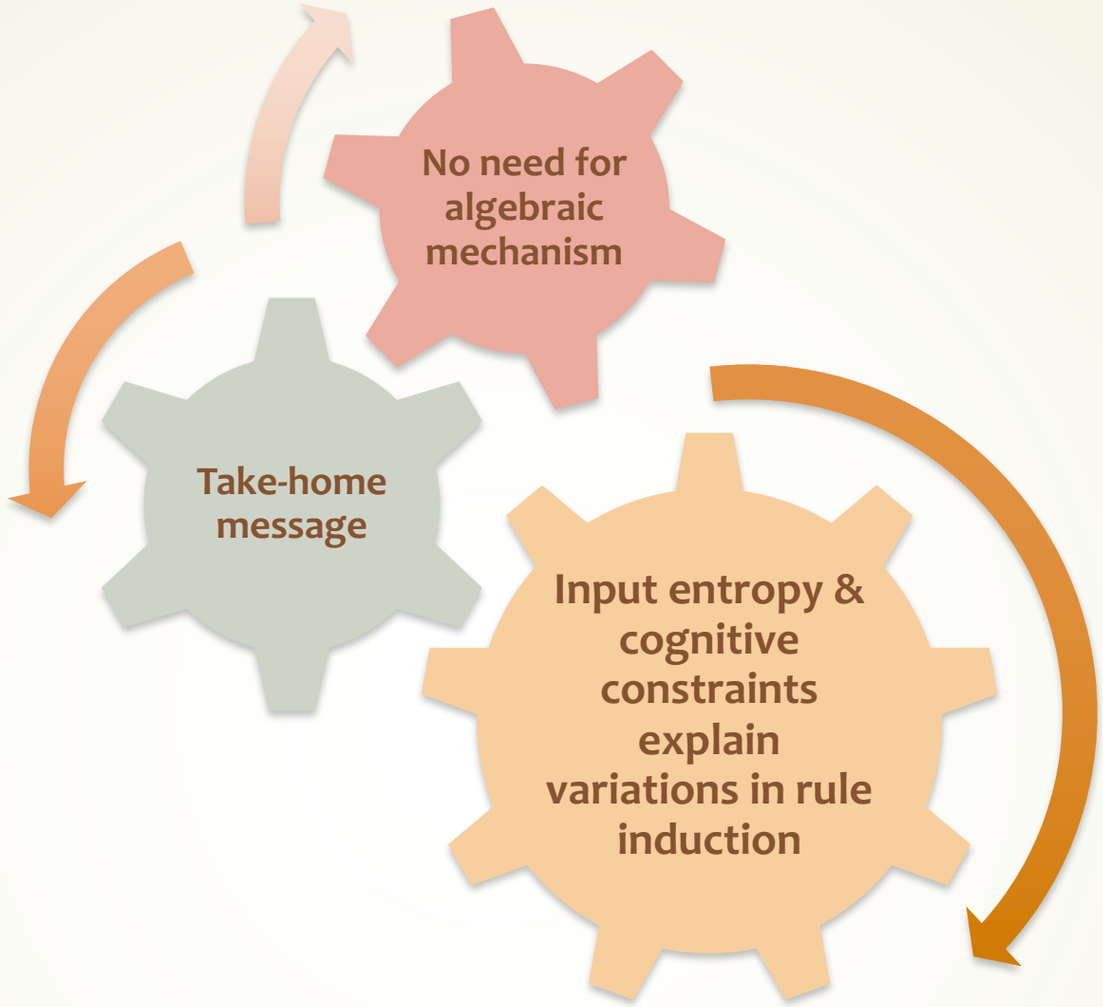
Conclusions

→ if input entropy increases, the tendency to generalize increases gradually

→ lower incidental memory predicts a higher tendency to generalize

→ higher visual pattern recognition predicts a higher tendency to generalize

→ the two types of rule induction are **outcomes of the same information encoding mechanism** → *gradually* move from lower-level *item-bound* encoding to higher-level *category-based* encoding in response to the interaction between **input entropy** and the encoding power (**channel capacity**)



Netherlands Organisation for Scientific Research

UiL OTS
UTRECHT INSTITUTE OF LINGUISTICS

"To think is to forget a difference, to generalize, to abstract.
In the overly replete world of Funes there were nothing but details,
almost contiguous details."

Funes, The Memorious

Jorge Luis Borges