M. A. Kelly, Penn State
Nipun Arora & Robert L. West, Carleton University
David Reitter, Google Research

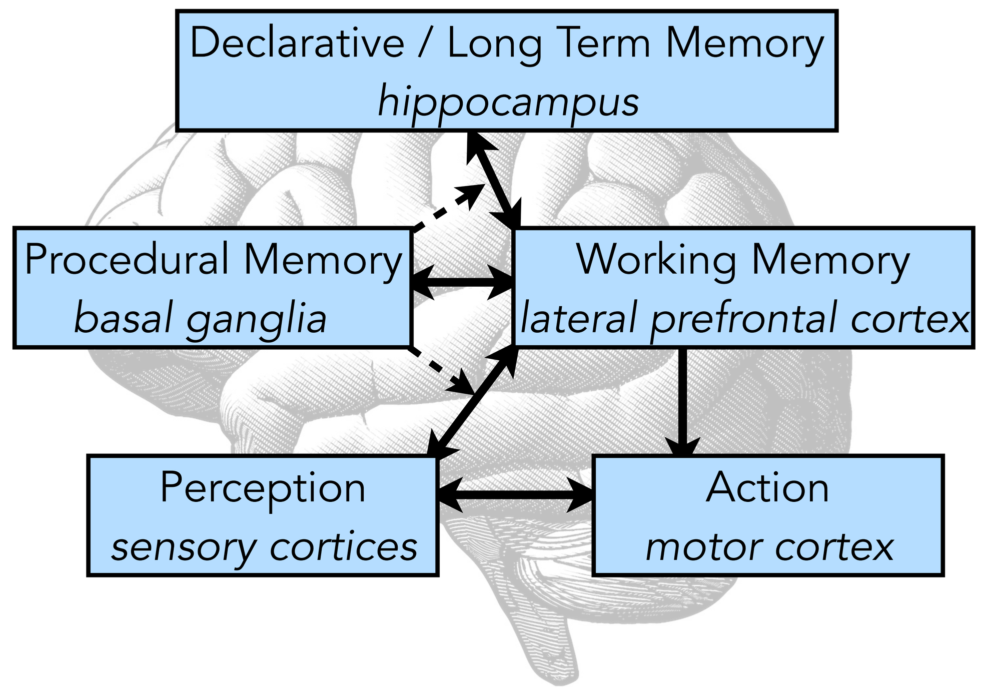
High-Dimensional Vector Spaces as the Architecture of Cognition

PennState
College of Information Sciences and Technology

The key parts of **cognitive architectures** - **declarative** and **procedural** memory - and their key abilities - **learning**, **memory retrieval**, **judgement**, and **decision-making** - can be realized as algebraic operations on vectors in a high-dimensional space.

Deep learning has an impressive ability to process data to find patterns, but does not model high-level cognition.
Symbolic architectures can capture the complexities of high-level cognition, but have limited ability to detect patterns or learn.
Vector-symbolic architectures, where symbols are represented as vectors, bridge the gap between approaches.

The Common Model of Cognition (Laird et al., 2017) describes a blueprint for realizing a cognitive architecture and associated brain areas (Steine-Hanson, Koh, & Stocco, 2018; Stocco et al., 2018).



Our model, **Holographic Declarative Memory (HDM)**, based on the BEAGLE (Jones & Mewhort, 2007) and DSHM models (Rutledge-Taylor et al., 2014) is a candidate for realizing **declarative memory** and aspects of **procedural memory**, in a manner that can be integrated with other vector-based approaches, such as deep-learning models of perception and action, or with symbolic approaches, e.g., **ACT-R** (Anderson & Lebiere, 1998). HDM uses ***holographic reduced representations*** (Plate, 1995), a means of instantiating arbitrarily complex concepts in high-dimensional vectors.

HDM accounts for **primacy** and **recency** effects in **free recall**, the **fan effect** in recognition, human **probability judgements**, and human performance on an **iterated decision** task. HDM provides a flexible, scalable alternative to symbolic architectures at a level of description that bridges ***symbolic***, ***quantum***, and ***neural*** models.

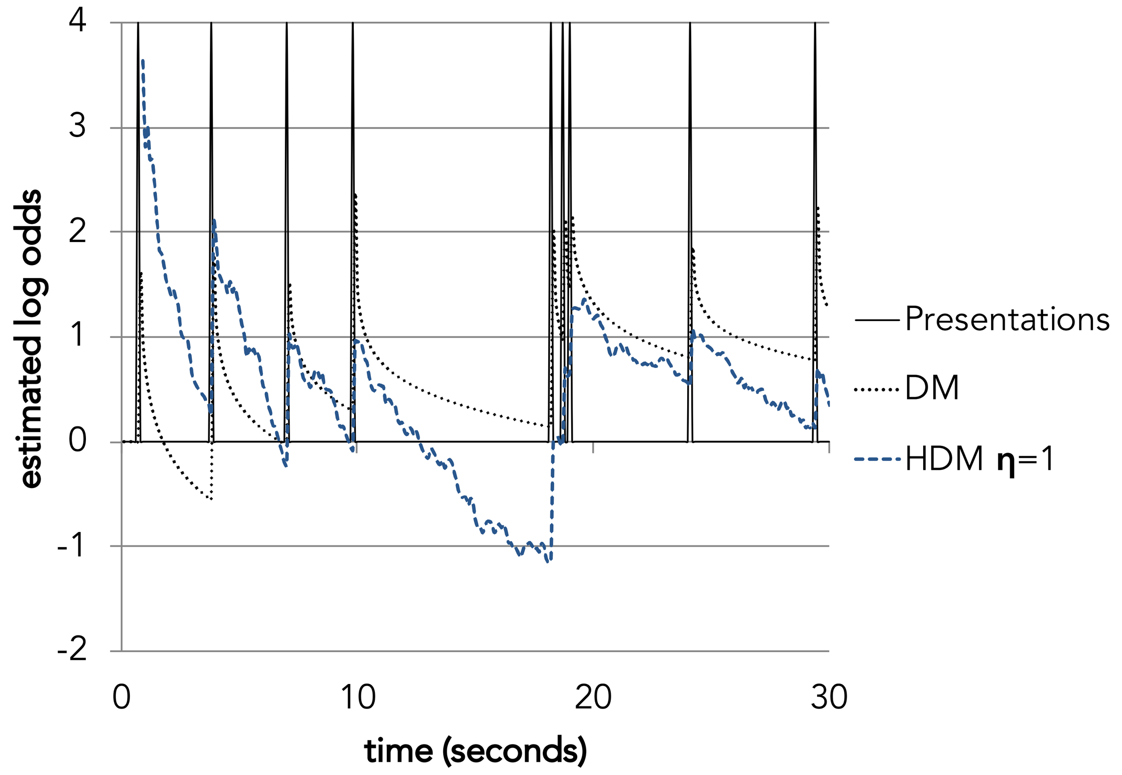
Vectors in HDM

e ***environment vector***, represents an item, randomly generated.
m ***memory vector***, constructed from environment vectors to encode associations between items in the environment.
q ***cue vector***, constructed from environment vectors to encode a question asked of memory.
p ***projection*** of a memory vector onto another memory vector to make probability judgements (quantum model).

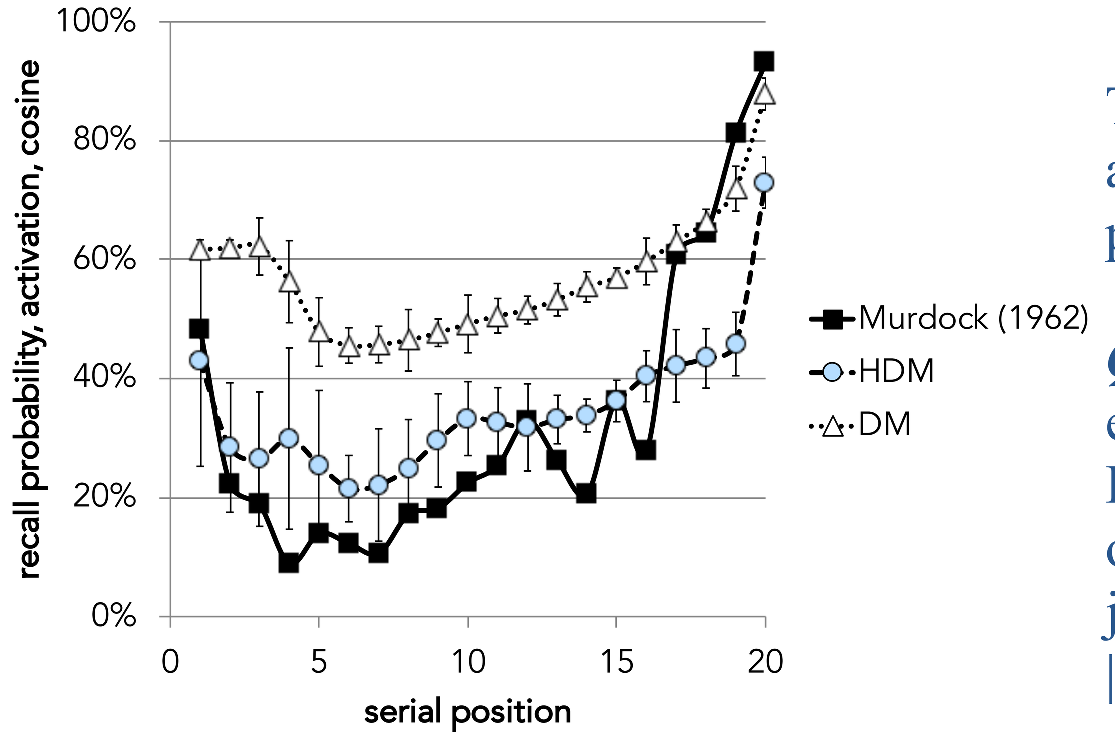
Free Recall

Participants and models were presented 20 words at a rate of one word every 2 seconds. After, participants reported back the list in any order (free recall). For simplicity, in the models, we use the state of memory as a proxy for recall probability. **Noise** η is added to the HDM vectors over time.

Activation of an item in memory over time in **ACT-R DM** and **HDM** as the item is repeatedly presented to the model.



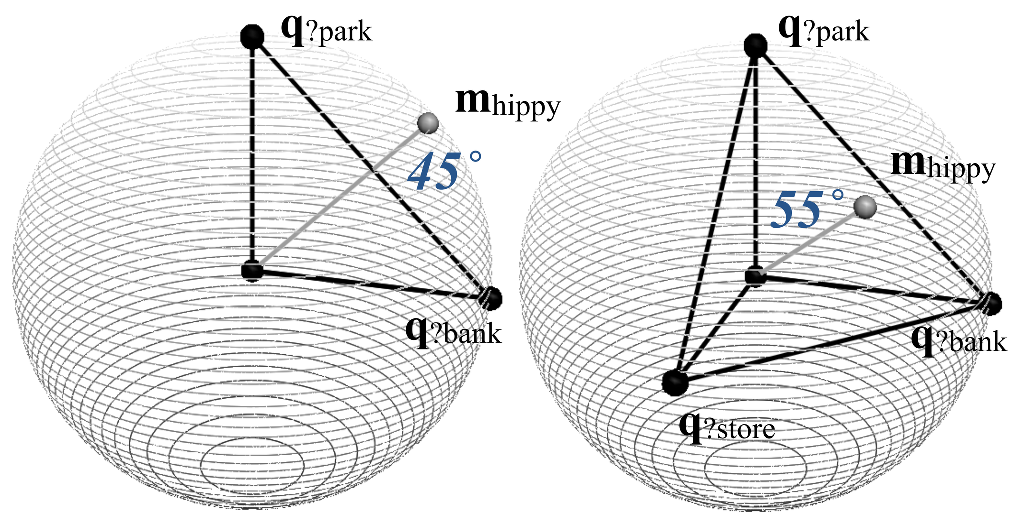
Ease of recall as a function of list position in a **free recall task**. Results averaged across 10 runs of each model. Human data from Murdock (1962).



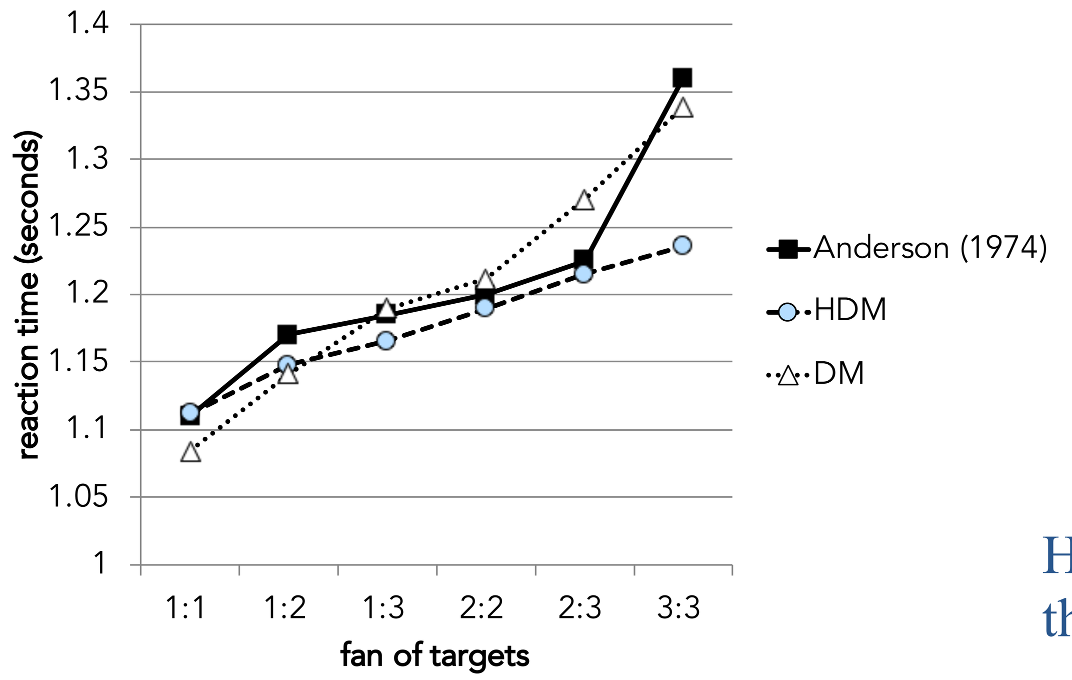
Fan Effect

Study Set: a list of object - location pairs (e.g., ***hippy in park***)
Test Set: participants must determine which pairs were studied.
Effect: participants are ***slower*** to judge pairs that contain items that occur in more pairs in the study set (i.e., have a ***higher fan***).
Theory: ***availability*** of an item in memory with respect to a cue is related to the ***probability*** of the item conditional on the cue.

The fan effect arises from the ***geometry*** of the vector space. The HDM model uses 256 dimensions, but the effect can be illustrated in 3 dimensions.
$\mathbf{m}_{hippy}$ with a fan of 2 (*left*) or 3 (*right*).

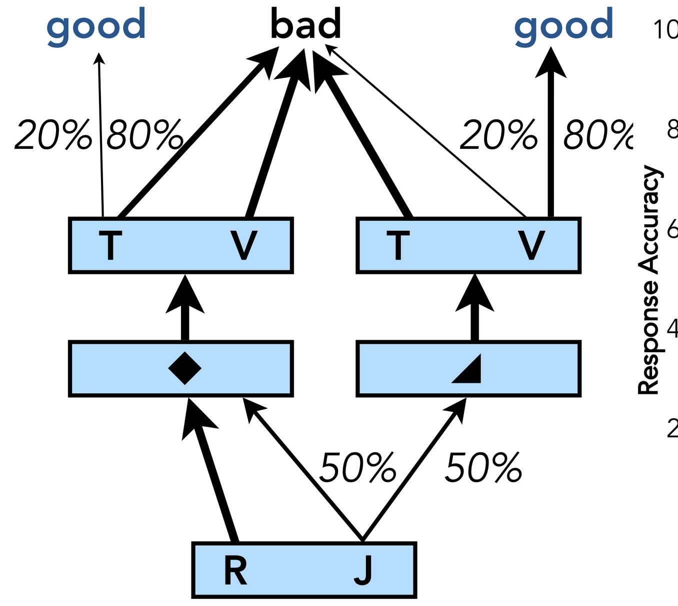


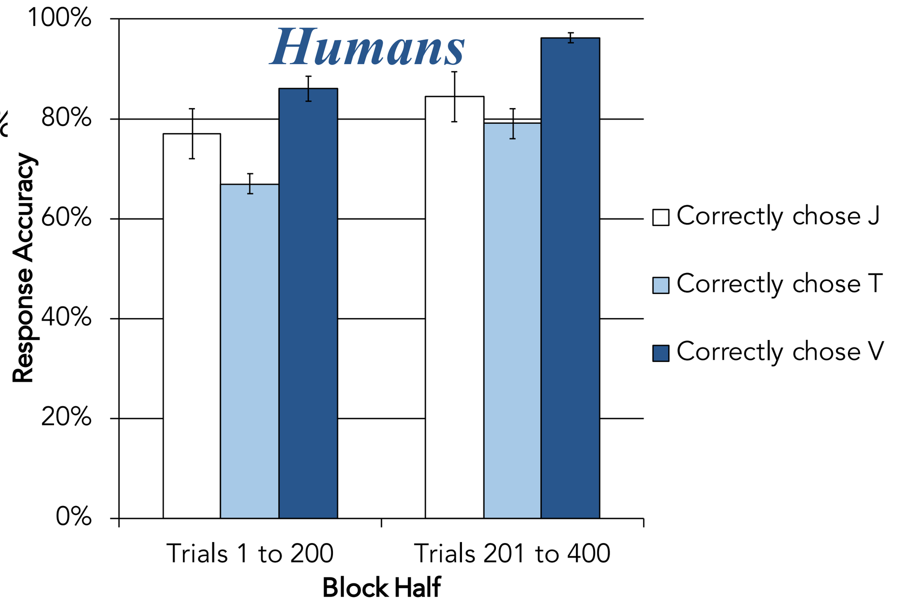
Response time for targets in the fan effect task (*left*). As distance to **q** goes up, retrieval of **m** slows.

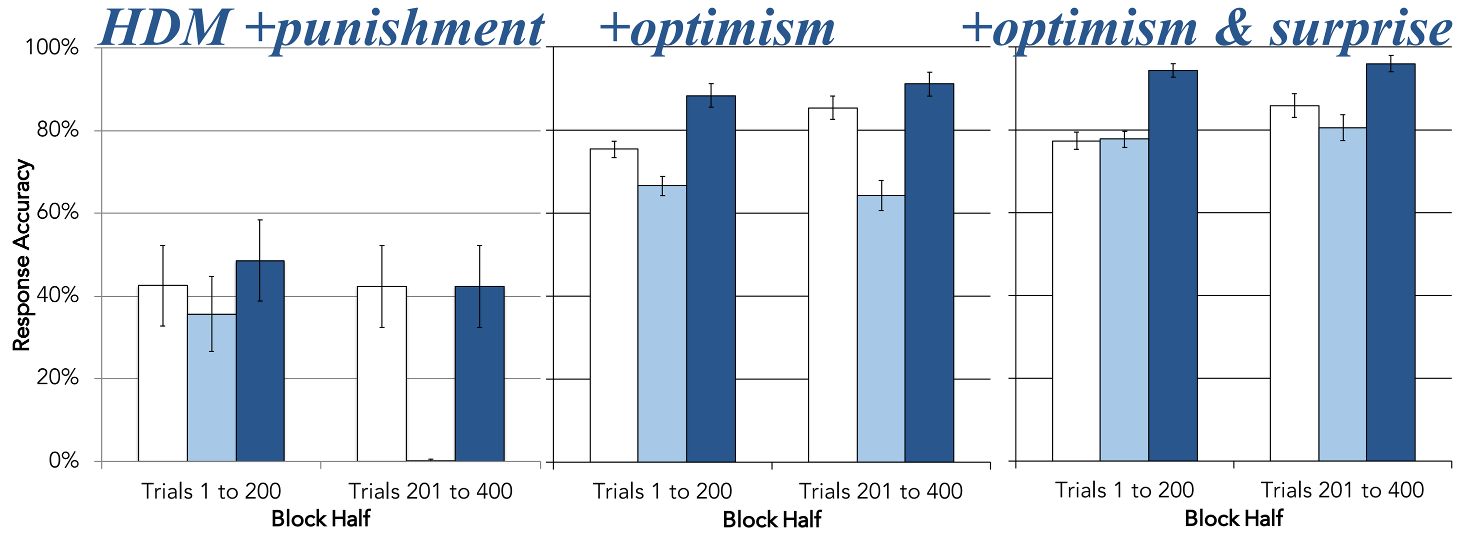


Iterated Decision

Walsh and Anderson (2011)'s iterated binary decision task. Participants gradually learn to make choices between arbitrary symbols that stochastically yield positive feedback.







HDM learns to make the correct decisions when biased to explore the space using ***optimism*** and to self-correct using ***surprise***.

Quantum Probability Judgement

Tversky and Kahneman (1983) have people read a description of a woman named **Linda** and ask them to judge how likely she is to be a **bank teller**, a **feminist**, and **both** a bank teller ***and*** a feminist. In defiance of probability theory, people say she's most likely **both**.

Quantum probability models (Busmeyer et al., 2011) can explain this finding. We can replicate their approach using **HDM** by computing geometric probabilities using projections **p** of memory vectors **m** onto other **m**. Human probability judgements are the squared magnitudes of the projections, e.g., $|\mathbf{p}_{Linda\rightarrow feminist\rightarrow bankteller}|^2 = P$ (feminist∧bankteller | Linda).