Cours 2022-2023:

Quel code neural pour les représentations mentales?
Vector codes and the geometry of mental representations

Stanislas Dehaene
Chaire de Psychologie Cognitive Expérimentale

Cours n°4

**Comment reconnaître une image, prendre une décision,
ou communiquer une information avec des vecteurs neuronaux dynamiques?**

Course 4

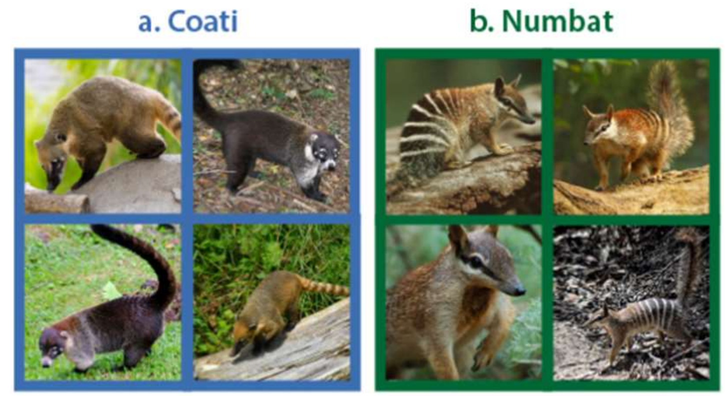
*Image recognition, decision making, and information communication
based on dynamic neuronal vectors*

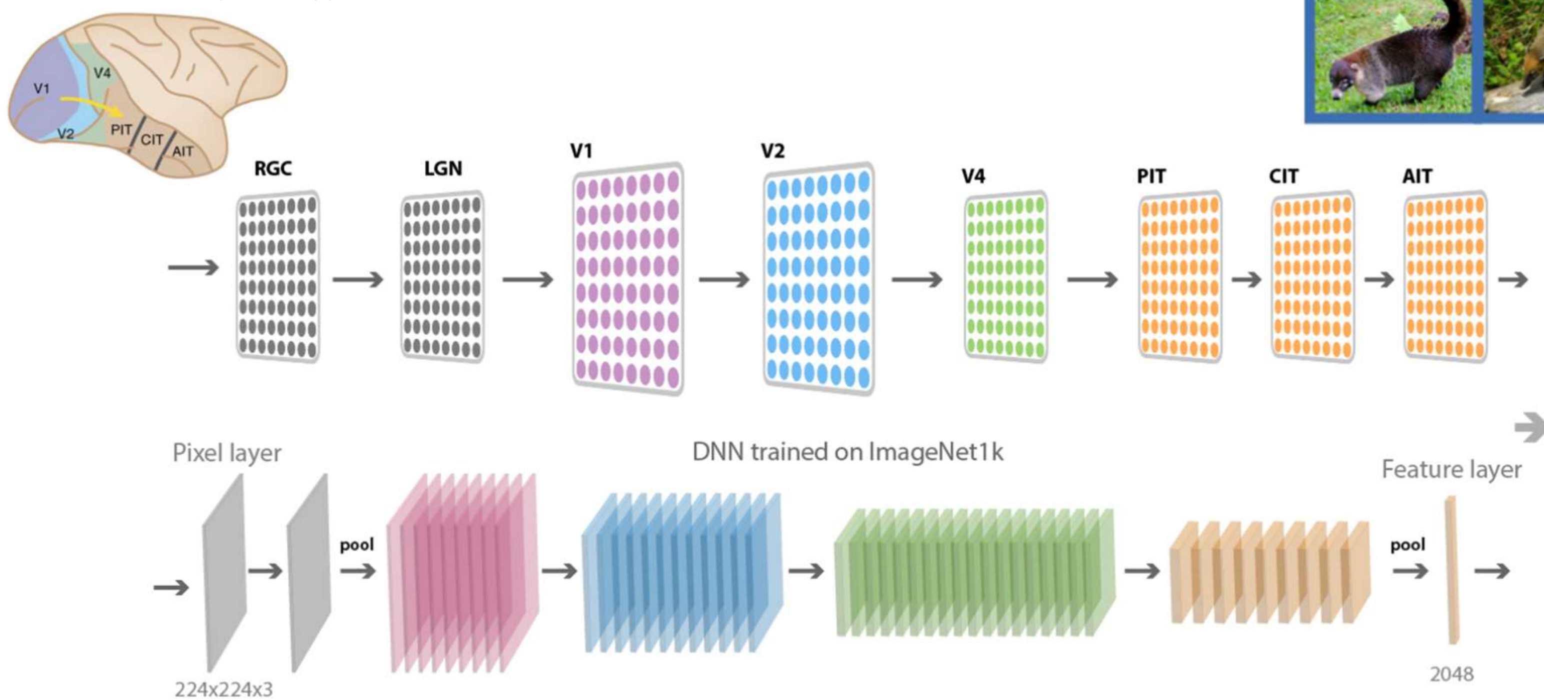
A theory of dimensionality and concept learning

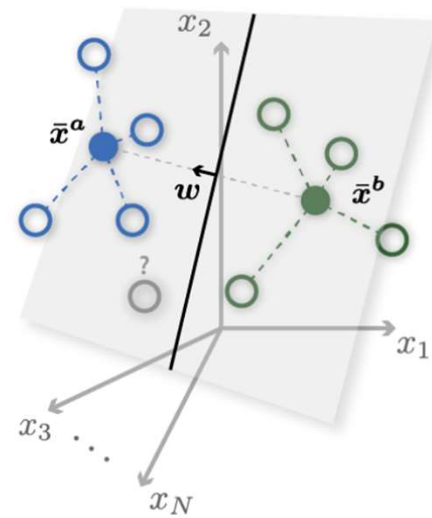
Sorscher, B., Ganguli, S., & Sompolinsky, H. (2022). Neural representational geometry underlies few-shot concept learning. PNAS, 119(43), e2200800119.

Here, the authors propose a general theory of “few shot learning” for image recognition.

1. Prior training has resulted in a tuned high-dimensional vector space for images, which can be used to perform one-shot or few-shot learning of new concepts.
2. Each example image (possibly 1) is encoded in this high-dimensional vector space
3. The **barycenter of examples** defines a **prototypical vector** for the new concept.
4. Classification of new images, or discrimination between two possibilities, is based on the **nearest prototype**







A theory of dimensionality and concept learning

Sorscher, B., Ganguli, S., & Sompolinsky, H. (2022). Neural representational geometry underlies few-shot concept learning. PNAS, 119(43), e2200800119.

Empirical explorations of this scheme:

- Train networks on ImageNet (1000 image categories)
- Test on **binary classification** of all possible pairs of 1000 new images from ImageNet21k

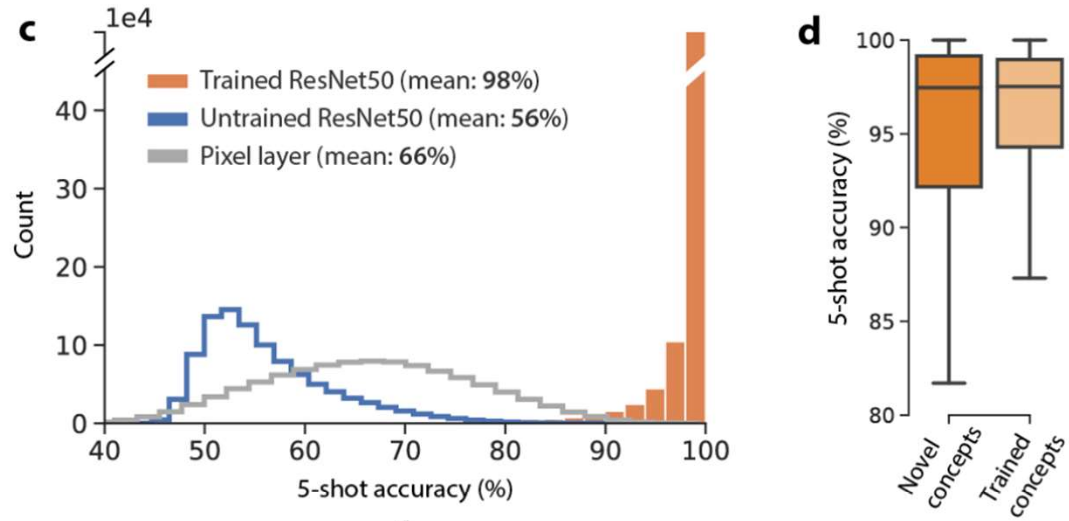
Results:

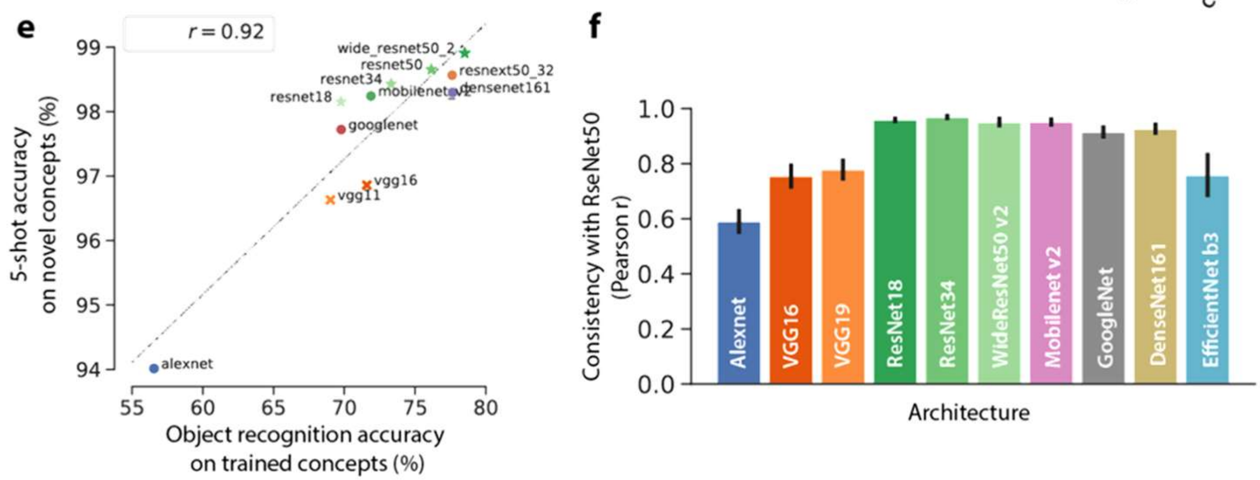
With just 5 examples, prototype learning manages to accurately classify new concepts, with an average of 98.6% correct ! (1-shot learning = 92 %)

All sorts of trained networks work, and their performances are intercorrelated with each other.

Untrained networks, however, do not perform well

→ the vector space must be tuned to pictures.





Geometry explains why some concepts are easier to discriminate than others

Sorscher, B., Ganguli, S., & Sompolinsky, H. (2022). Neural representational geometry underlies few-shot concept learning. PNAS, 119(43), e2200800119.

$$\mathrm{SNR}_a = \frac{1}{2}\frac{\|\boldsymbol{\Delta x_0}\|^2 + (R_b^2 R_a^{-2} - 1)/m}{\sqrt{D_a^{-1}/m + \|\boldsymbol{\Delta x_0}\cdot \boldsymbol{U}_b\|^2/m + \|\boldsymbol{\Delta x_0}\cdot \boldsymbol{U}_a\|^2}}$$

In this vector space, the images for each new concept trace a manifold, which the authors approximate with a high-dimensional ellipsoid.

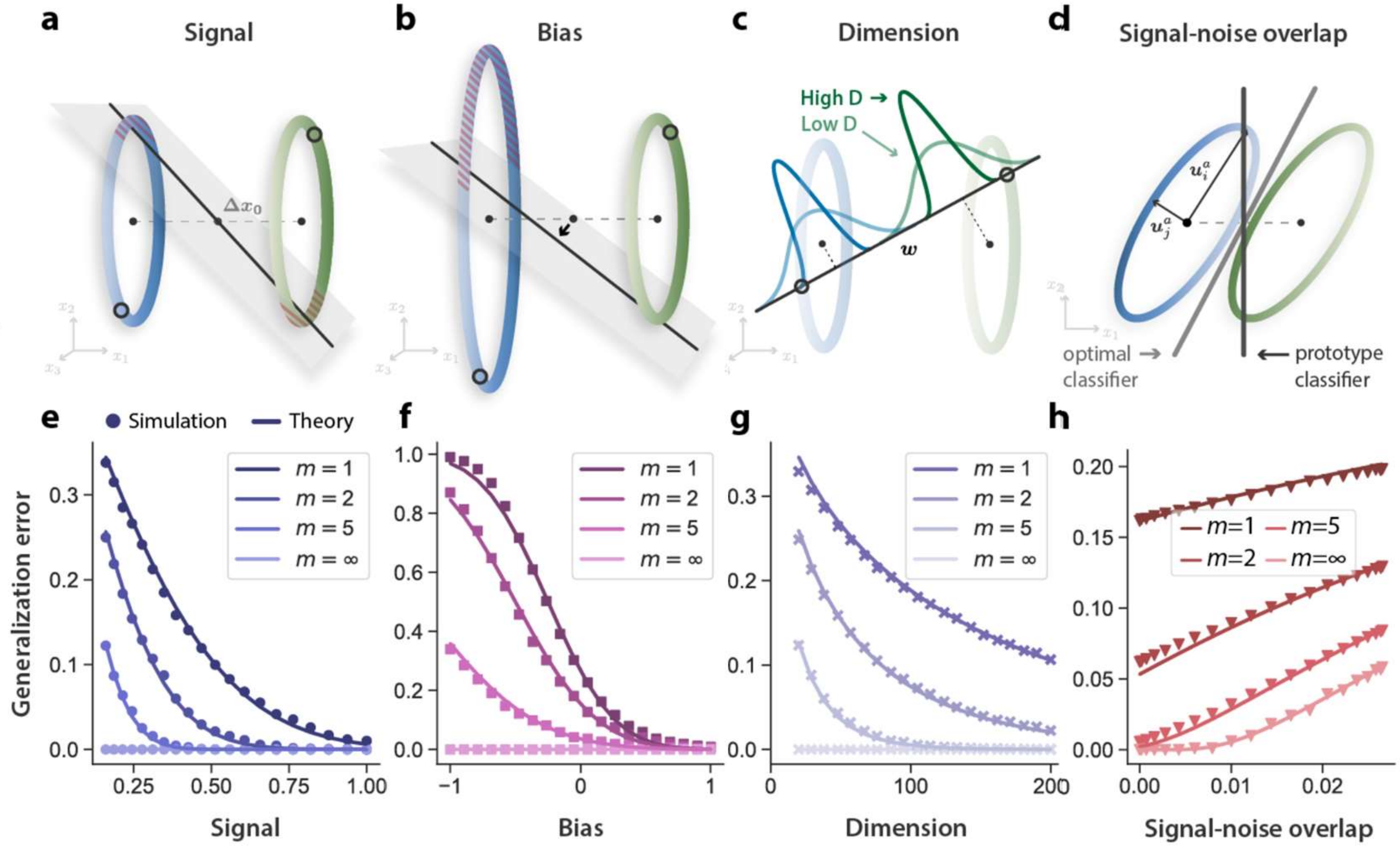
Four sources of errors in classification (m is number of training examples):

a. The **pairwise separation** between the two manifolds (signal) may be weak relative to the noise level.

b. One manifold may have a **larger variance** than the other, resulting in generalization errors and even worse-than chance performance.

c. The **dimensionality** of the manifolds may vary – and here, surprisingly, performance is better in higher dimensions (blessing of dimensionality).

d. **Noise** may vary in the same direction as the signal (the centroid separation vector). Note that here, the asymptote does not go to zero… unless a more optimal classifier is used.



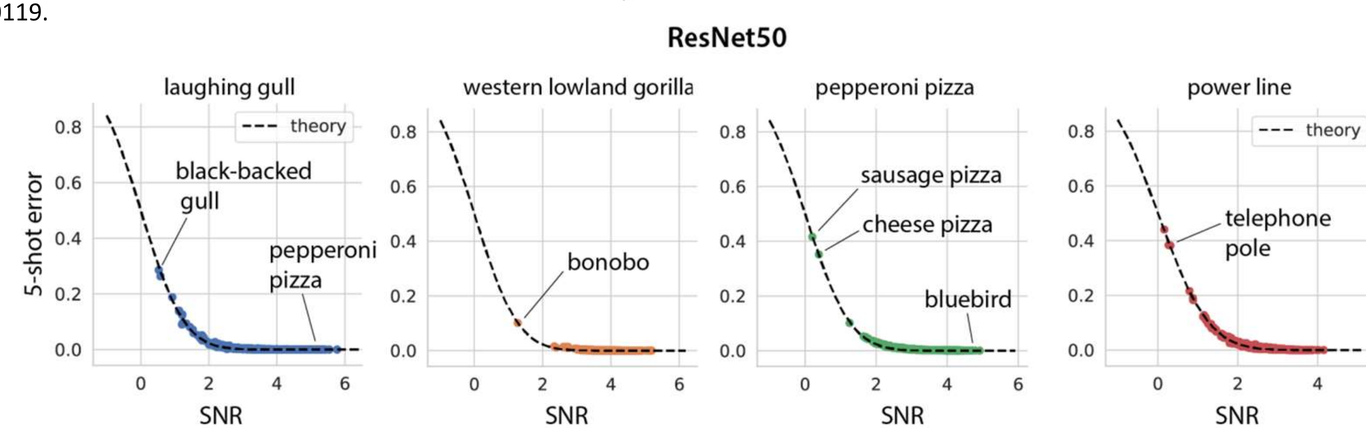
The dark line shows a fit of the author’s equation to simulations of ellipsoid categories

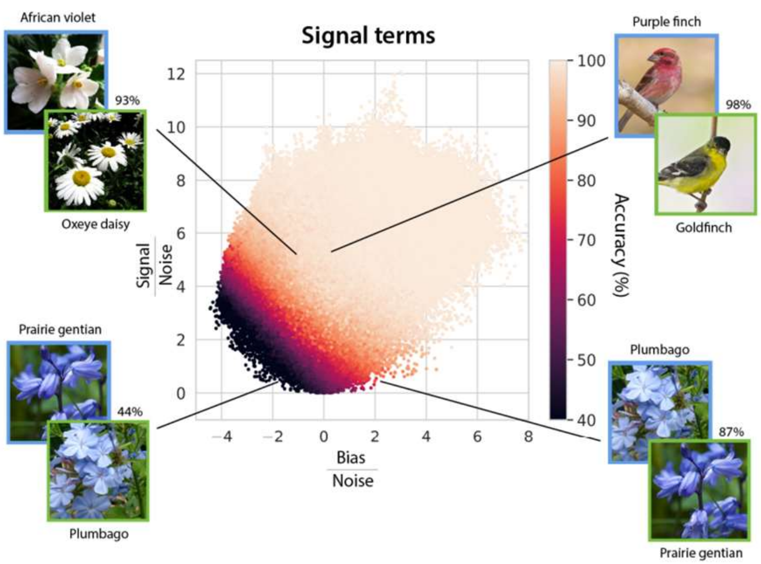
A theory of dimensionality and concept learning

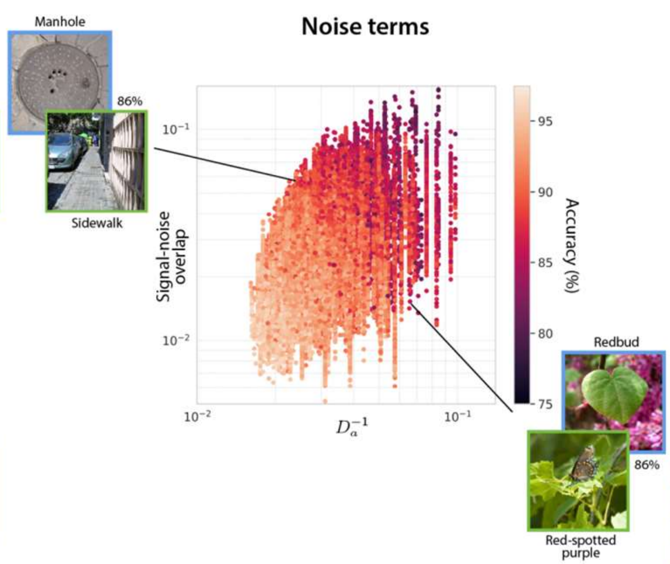
$$\mathrm{SNR}_a = \frac{1}{2}\frac{\|\boldsymbol{\Delta x_0}\|^2 + (R_b^2 R_a^{-2} - 1)/m}{\sqrt{D_a^{-1}/m + \|\boldsymbol{\Delta x_0}\cdot \boldsymbol{U}_b\|^2/m + \|\boldsymbol{\Delta x_0}\cdot \boldsymbol{U}_a\|^2}}$$

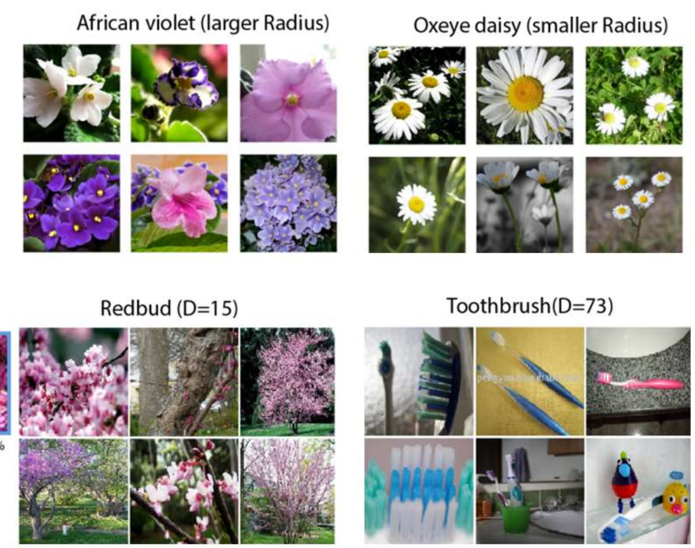
Sorscher, B., Ganguli, S., & Sompolinsky, H. (2022). Neural representational geometry underlies few-shot concept learning. PNAS, 119(43), e2200800119.

More important: the SNR equation fits the binary classification of **actual images.**
The analysis therefore provides four different geometric quantities that characterize category learning: **signal, bias, dimension, and signal-noise overlap**.
Examples show how pairs of categories may differ on all those parameters, and explain why they are more or less difficult to separate.









Coding in higher dimensions facilitates learning and generalization

Elmoznino, E., & Bonner, M. F. (2022). High-performing neural network models of visual cortex benefit from high latent dimensionality (p. 2022.07.13.499969). bioRxiv. https://doi.org/10.1101/2022.07.13.499969

Networks with higher "effective dimensionality" are better at learning to classify new categories according to a prototype rule.

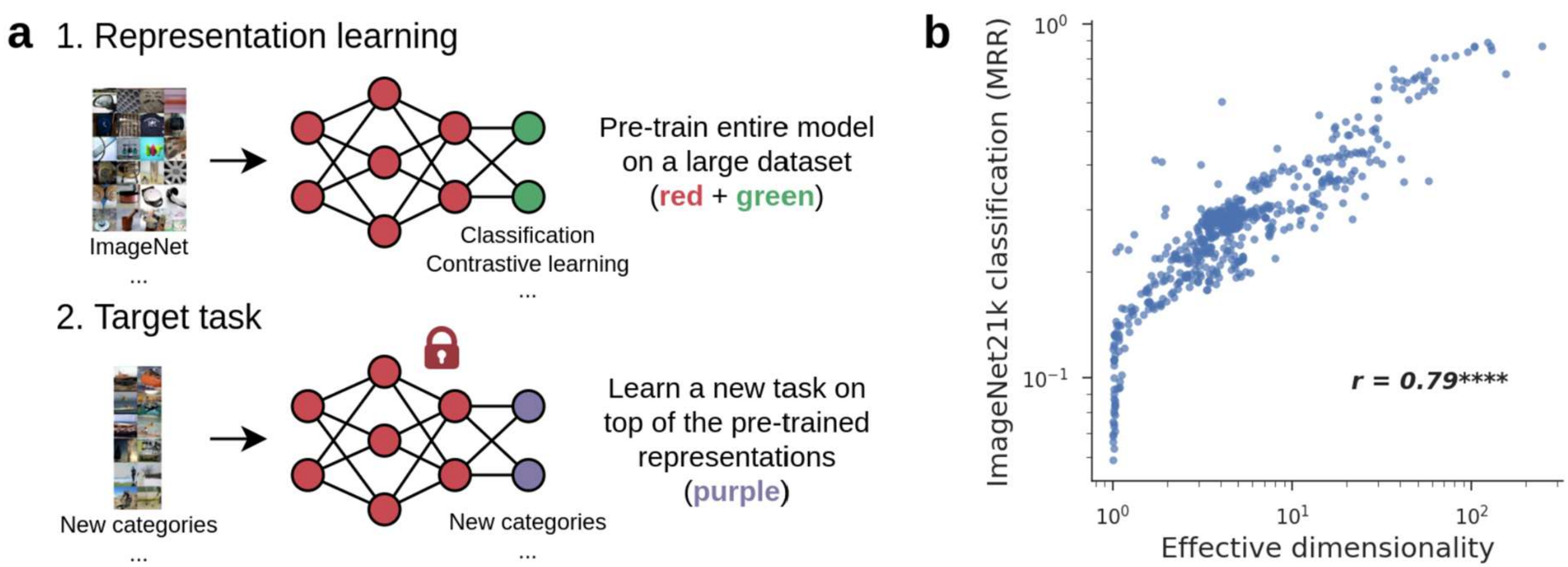


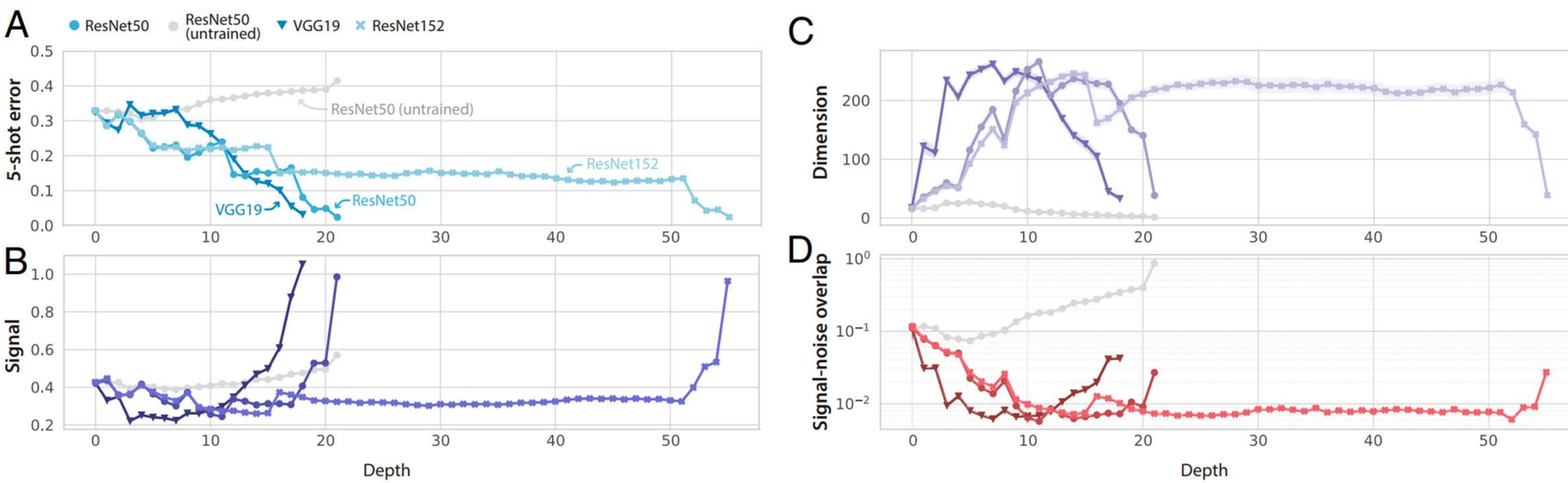
Figure 4: The computational benefit of high effective dimensionality in generalization to new object categories. We examined the hypothesis that high-dimensional representations are better at learning to classify new object categories (Sorscher et al., 2021). **a.** We tested this theory using a transfer learning paradigm, where our pre-trained model representations were fixed and used to classify novel categories through a prototype learning rule. **b.** Our high-dimensional models achieved substantially better accuracy on this transfer task, as measured using the mean reciprocal rank (MRR).

A theory of dimensionality and concept learning

Sorscher, B., Ganguli, S., & Sompolinsky, H. (2022). Neural representational geometry underlies few-shot concept learning. PNAS, 119(43), e2200800119.

The theory also permits to analyze **the intermediate layers** of any artificial neural network for visual recognition.
All networks tested seem to improve image discrimination across successive layers (A) by

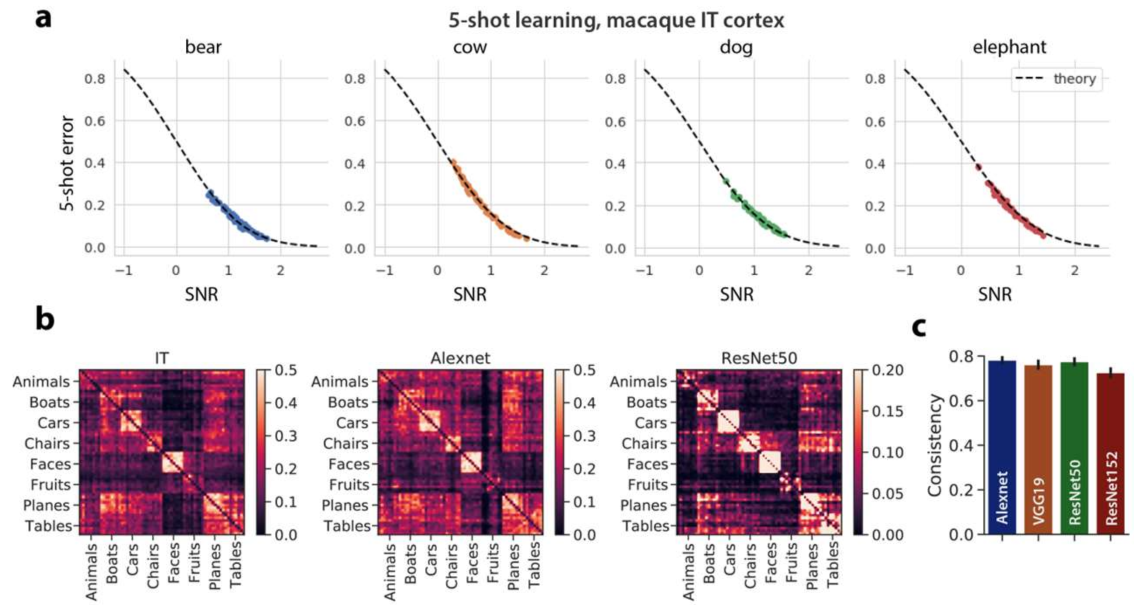
- Expanding the number of dimensions, then compressing it (C)
- Thereby dramatically enhancing the signal in the last layers (B)
- And progressively separately the signal from the noise (D), yet with a surprising increase at the end.

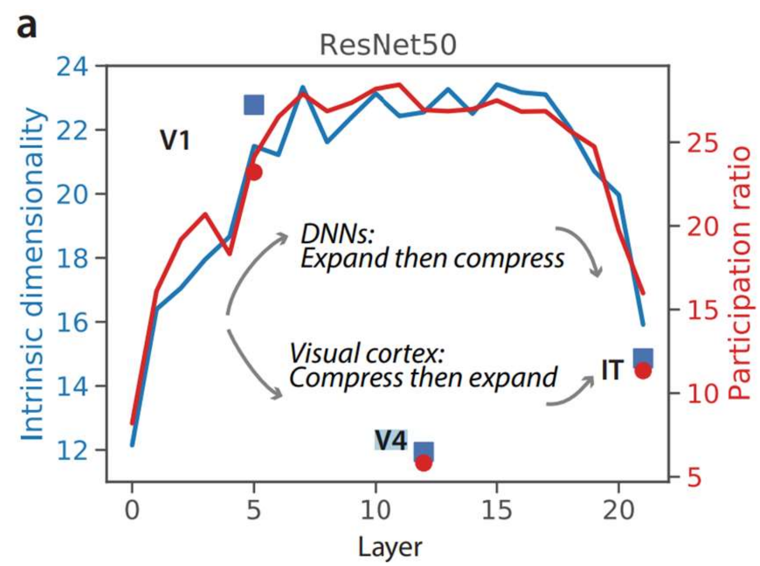


A theory of dimensionality and concept learning

Sorscher, B., Ganguli, S., & Sompolinsky, H. (2022). Neural representational geometry underlies few-shot concept learning. PNAS, 119(43), e2200800119.

The theory can also be used to analyze the geometry of representations in monkey visual cortex (V4 and IT).
It predicts how well a given concept can be discriminated from others.
The errors in classifying concepts from IT neurons and from various artificial CNNs are similar…
But surprisingly, the underlying parameters do not vary in the same way across the hierarchy: the strategy across layers, and in V4 in particularly, does not seem to match that of artificial neural networks.





The geometric theory of prototypes can even support zero-shot verbal learning

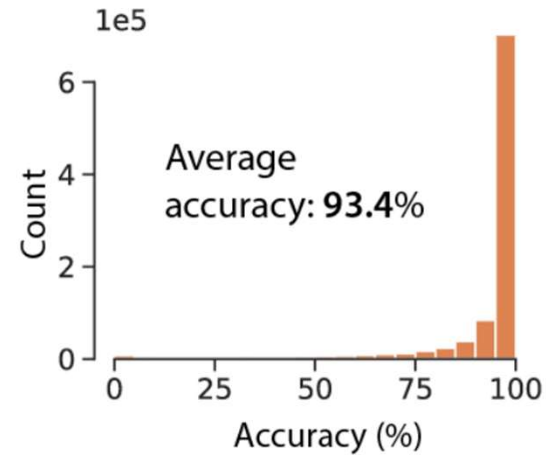
Sorscher, B., Ganguli, S., & Sompolinsky, H. (2022). Neural representational geometry underlies few-shot concept learning. PNAS, 119(43), e2200800119.

Suppose you only had a verbal description of the concepts. The words in those descriptions can be assigned vector representations, based on their statistics of cooccurrence with other words (this will be detailed in course 5).
The average vector for this "bag of words" is taken as the verbal representation of the concept.
This **verbal vector space** can then be aligned onto the **picture vector space** by using the 1000 images used for training + their verbal descriptions (an isometry is learned using Procrustes alignment).
Prototypes for new concepts can then be acquired solely by their verbal description… and then used to classify pictures, with 93.4 % accuracy !
Remarkably, this figure is slightly better than that obtained with a single picture (m=1, 92% correct). Thus, it is not true that a picture is worth a thousand words – on the contrary!
We shall now examine two broad consequences of the same geometrical framework:

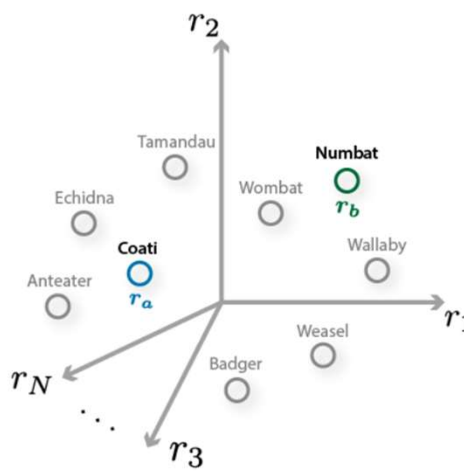
- Shepard's exponential law of generalization
- King & Dehaene's explanation of subliminal versus conscious processing.

I. coati, coati-mondi, Nasua narica
ring-tailed mammal with a slender head and an elongated nose

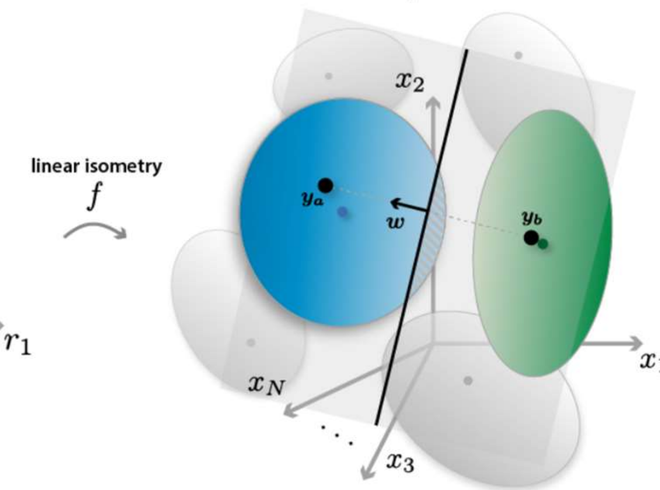
II. numbat, banded anteater
striped marsupial with a finely pointed muzzle and a bushy tail



Language representation



Visual representation



Towards a universal law of generalization

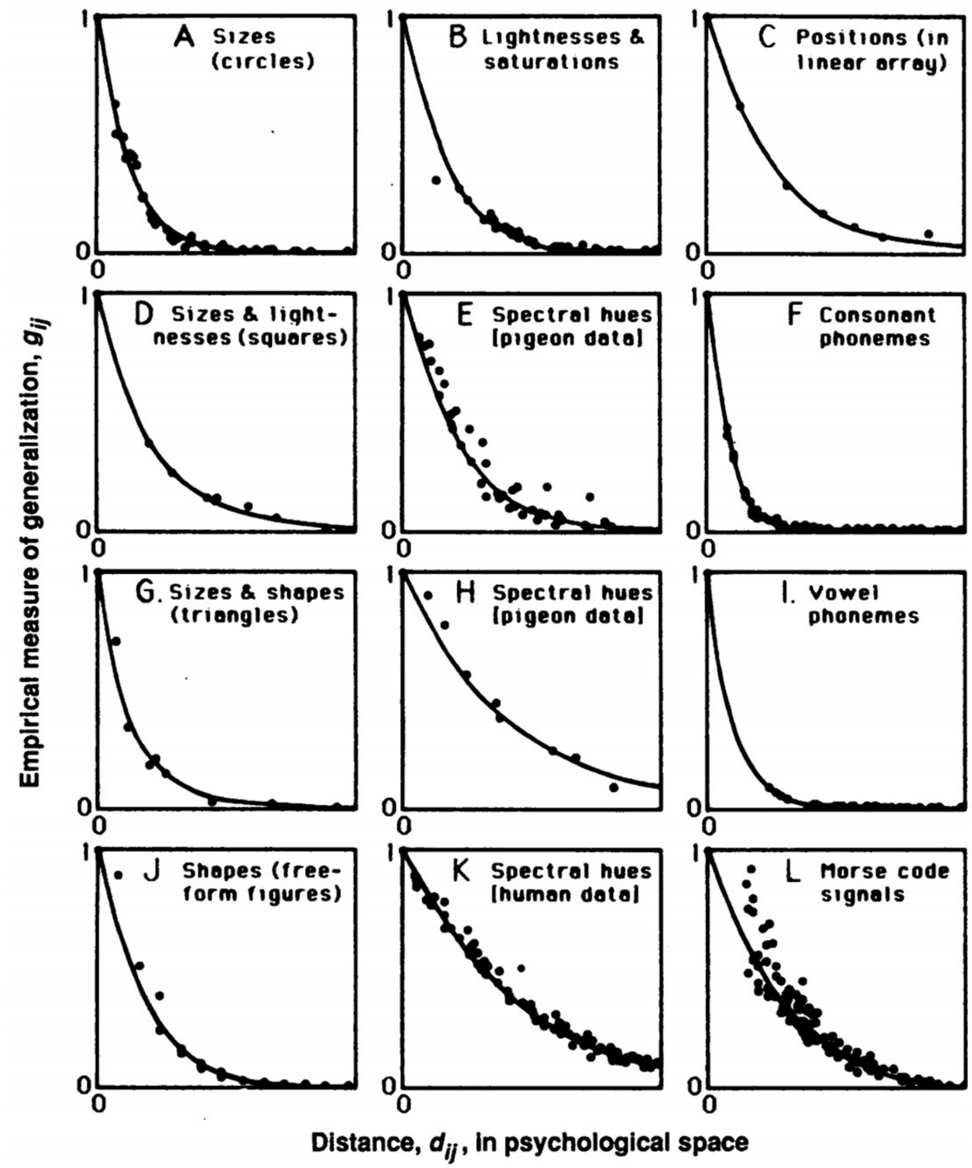
Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. Science, 237(4820), 1317-1323.

In an influential paper, Shepard shows how the generalization of a learned concept follows a universal exponential law.
Note that this is ***not*** true if performance is plotted as a function of distance in the original stimulus space, for instance:

- the psychological similarity between notes decreases as their difference in pitch increases, but increase again for octaves.
- The similarity of colors does not decrease with wavelength distance, but forms a circle, first discovered by Newton.

The key is to plot the data in **psychological space.**
The dimensionality and organization of that space can be determined by a mathematical procedure called **non-metric multidimensional scaling**, while solely supposes that the distance between points must be monotonically related to psychological similarity judgments.

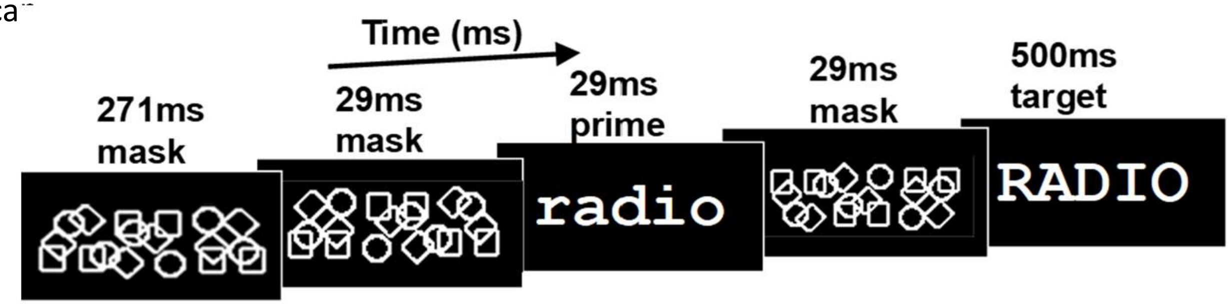
Shepard finds that it is always possible to find a small vector space, such that increasing distance between concepts on that space (either L1 or L2) predicts an exponential decrease in similarity judgments.
The Sorscher et al. prototype theory can explain this law for pictures. It predicts that the decrease should be a specific function (the integral of a Gaussian) of distance, ponderated by bias and by the various noise terms.
The Shepard findings suggests that many other domains (e.g. pitch) may be explained by similar internal spaces.



Explaining the enigmas of subliminal processing and conscious access as categorical decisions in a vast representational space

King, J.-R., & Dehaene, S. (2014). A model of subjective report and objective discrimination as categorical decisions in a vast representational space. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1641), 20130204. https://doi.org/10.1098/rstb.2013.0204

When stimuli are flashed and masked, categorization can be above-chance even though the participants deny seeing the stimuli.
A geometrical vector view, **extended with Bayesian decision making**, can explain this paradox – indeed, it results from a perfectly rational strategy.



We can explain all of the following phenomena:

- Stimuli which are subjectively reported as '**unseen**' can nevertheless be objectively discriminated **above chance** in a two-alternative forced-choice task.
- Discrimination performance is typically **better on seen than on unseen trials**, even when sensory stimuli are physically identical.
- Experimental paradigms can be designed in which **objective discrimination performance is identical, while subjective visibility differs**.
- Subjective reports vary **nonlinearly** as a function of sensory strength. For instance, brief or faint visual stimuli are generally reported as 'completely unseen', but once their duration or contrast reaches a threshold level, subjects tend to report items as 'clearly seen'.
- **Prior knowledge** increases the subjective visibility of physically identical stimuli.
- **Attention** generally increases subjective visibility but has also been found to decrease it.

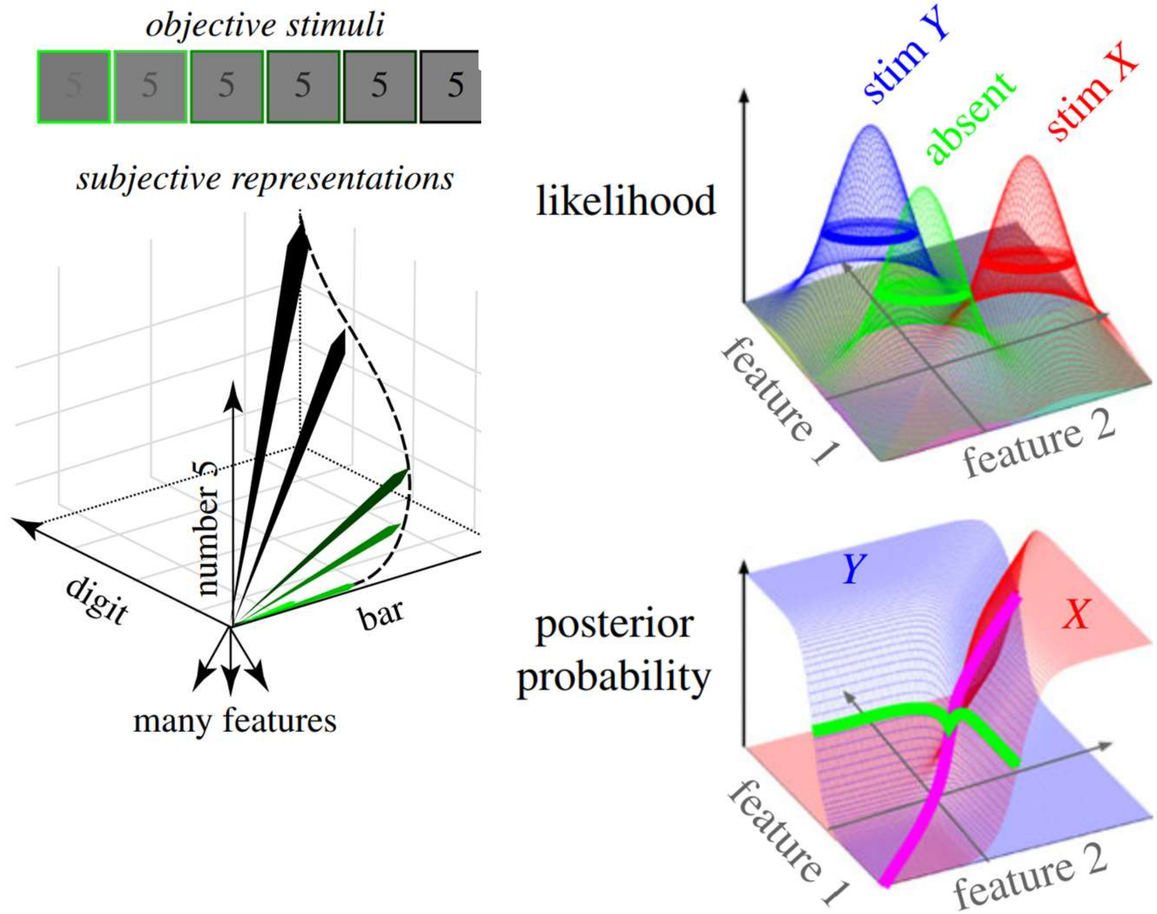
Explaining the mysteries of subliminal processing and conscious access using vector spaces

King, J.-R., & Dehaene, S. (2014). A model of subjective report and objective discrimination as categorical decisions in a vast representational space. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1641), 20130204. https://doi.org/10.1098/rstb.2013.0204

Like Sorscher et al., we assume that each incoming stimulus is encoded in a high-dimensional vector space.
We suppose that increasingly visibility translates into a lengthening of the vector along some of its axes.
This vector space is huge, but we are going to simplify it by projecting it along two main axes (e.g. vectors X and Y).

Crucially, instead of supposing that subjects base their judgement on the closest of two prototypes (Sorscher et al.), we assumed a **Bayesian decision-making system** that
1. Has a set of task-driven categories of stimuli
2. Estimates the likelihood that the observed vector came from each possible category.
3. Combines this likelihood with a (possibly flat) prior to compute the posterior probability of each category, and
4. chooses the most likely one (MAP strategy, *maximum a posteriori*).

This view generalizes to non-binary classification tasks. Note that while forced-choice tasks may be binary, conscious perception is a choice amongst a much richer array of possibilities (including “no stimulus was presented”).



Subliminal processing and conscious access : a perspective from vector spaces

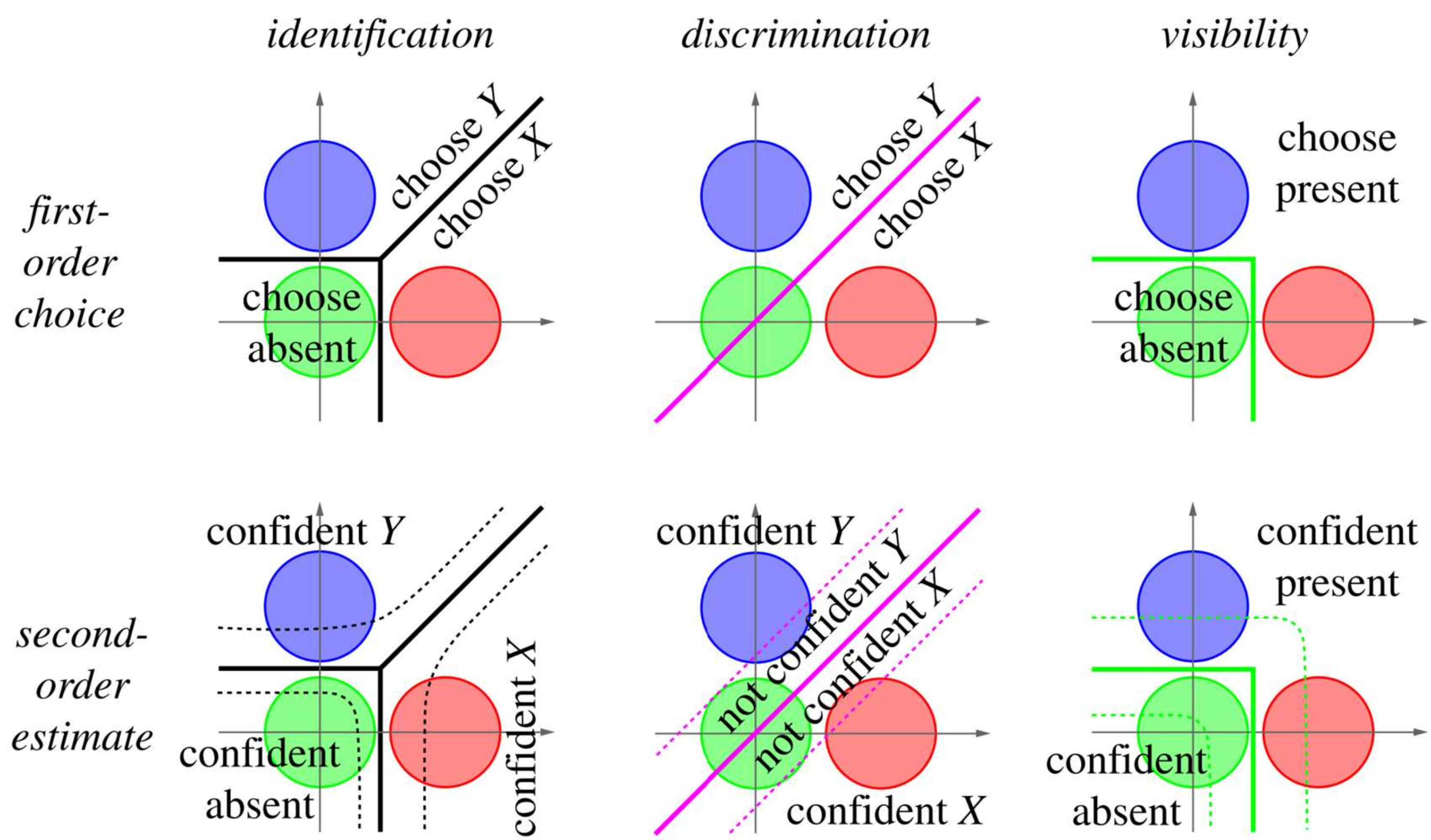
King, J.-R., & Dehaene, S. (2014). A model of subjective report and objective discrimination as categorical decisions in a vast representational space. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1641), 20130204. https://doi.org/10.1098/rstb.2013.0204

Each task judgement corresponds to a different layout of the response categories.

In particular,

- conscious identification consists in finding the MAP category among all possible classes of responses, including “stimulus absent”.
- Forced-choice discrimination is a much simpler task which merely consists in deciding which is more likely, X or Y? (linear decision boundary, similar to choosing the nearest prototype).
- Visibility judgment is another binary task, but with a more complex decision boundary.

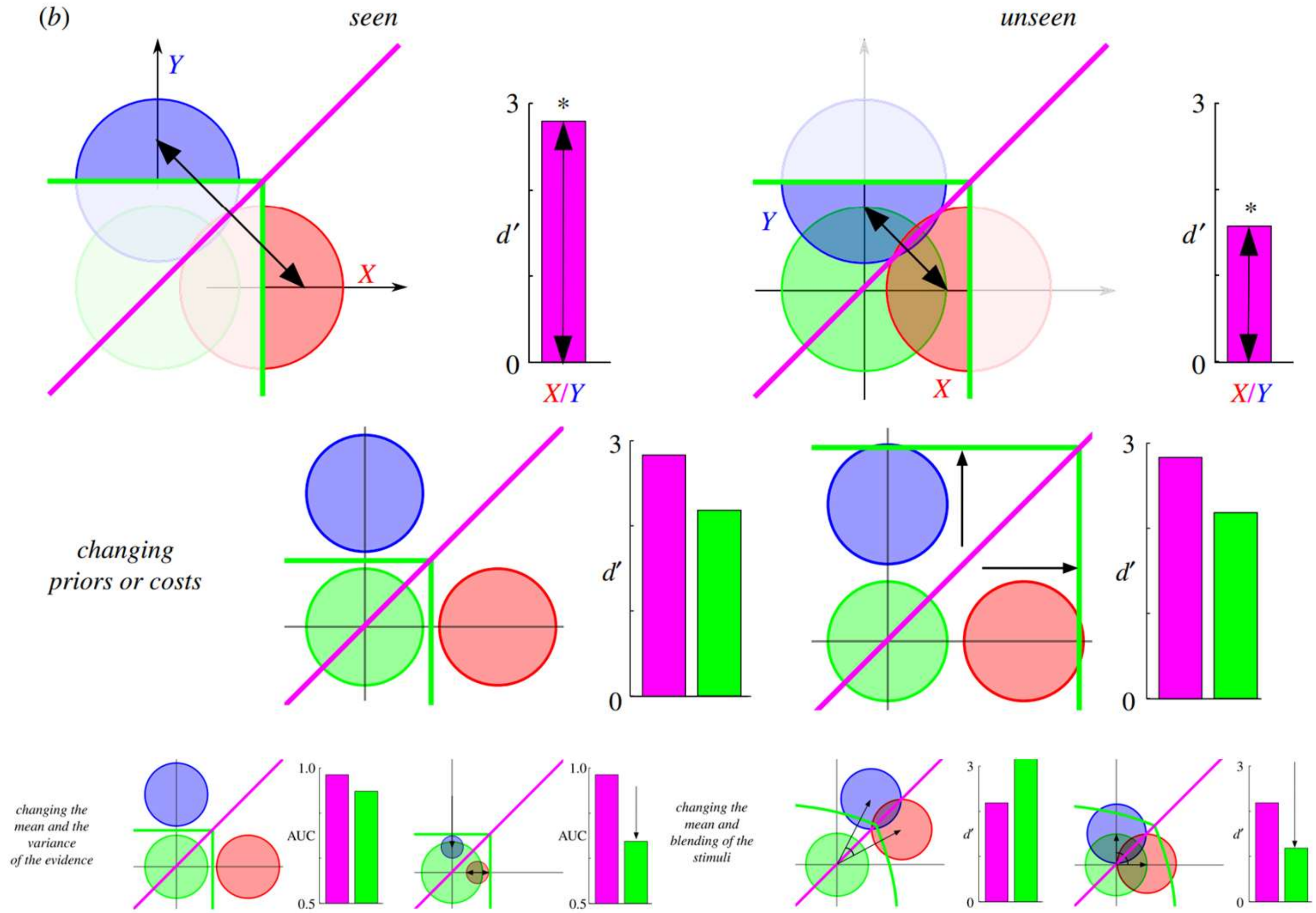
Note that confidence judgments are also distinct tasks, not reducible to visibility or conscious identification tasks.



Subliminal processing and conscious access : a perspective from vector spaces

King, J.-R., & Dehaene, S. (2014). A model of subjective report and objective discrimination as categorical decisions in a vast representational space. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1641), 20130204. https://doi.org/10.1098/rstb.2013.0204

- Stimuli which are subjectively reported as '**unseen**' can nevertheless be objectively discriminated **above chance** in a two-alternative forced-choice task.
- Discrimination performance is typically **better on seen than on unseen trials**, even when sensory stimuli are physically identical.
- Experimental paradigms can be designed in which **objective discrimination performance is identical, while subjective visibility differs**.



Subliminal processing and conscious access : a perspective from vector spaces

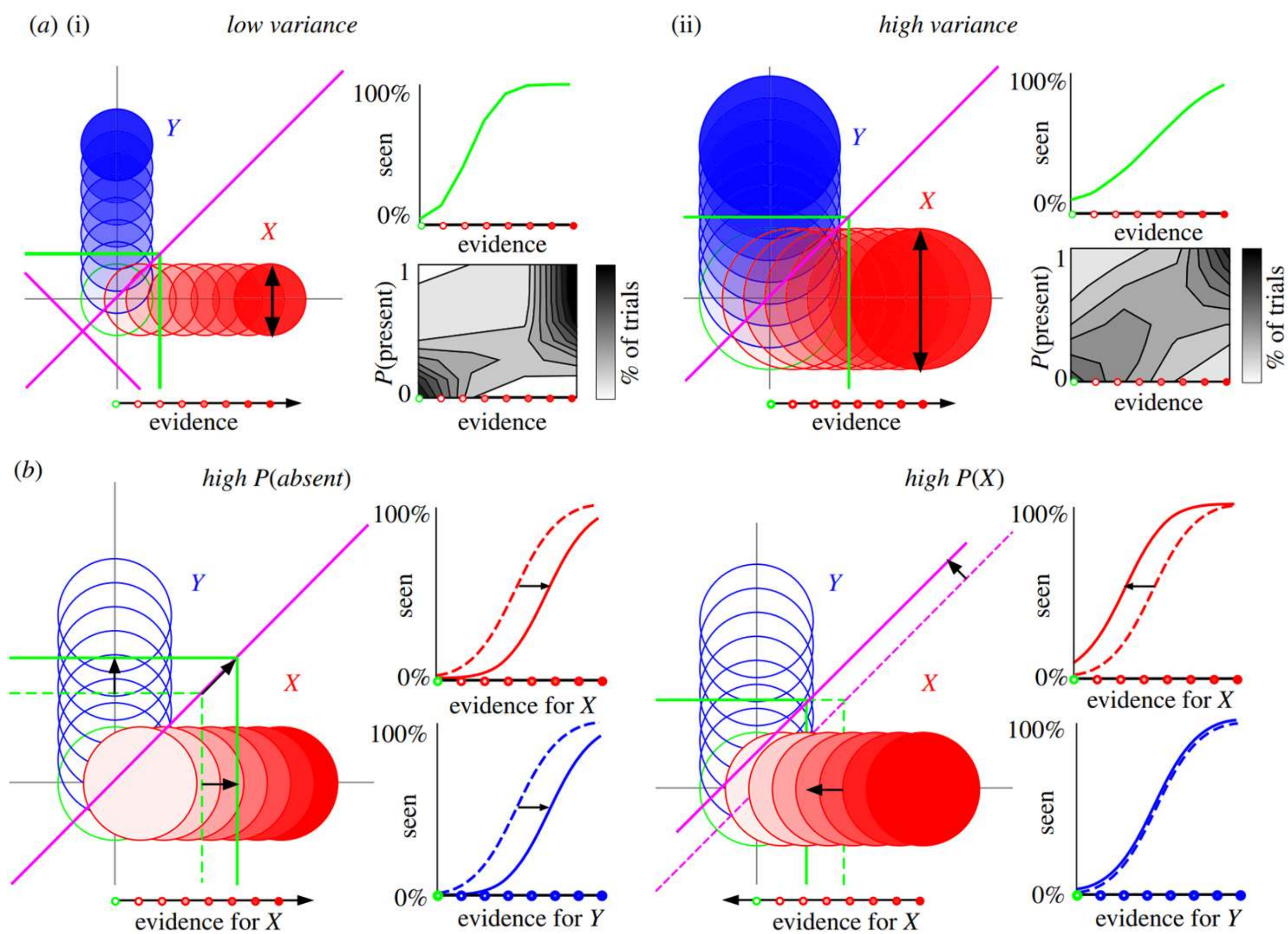
King, J.-R., & Dehaene, S. (2014). A model of subjective report and objective discrimination as categorical decisions in a vast representational space. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1641), 20130204. https://doi.org/10.1098/rstb.2013.0204

Explaining the **non-linear threshold** for conscious perception

- Subjective reports vary **nonlinearly** as a function of sensory strength.
- The non-linearity becomes less steep as the variance increases (or equivalently, as visibility decreases).

The non-linear threshold also varies according to the prior for present/absent or for X versus Y

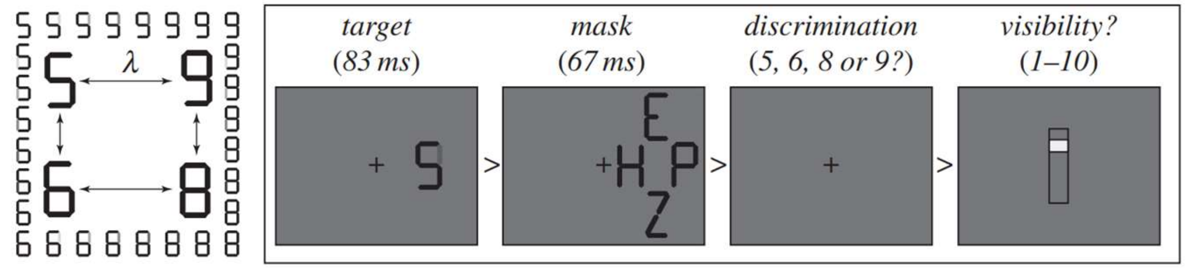
- A higher frequency of absent trials reduces the subjective visibility of physically identical stimuli.
- If we manipulate the prior for just one of the stimuli, we see that its visibility increases, while the visibility of other stimuli does not change – mimicking e.g. hysteresis in conscious perception (Melloni et al. 2011).



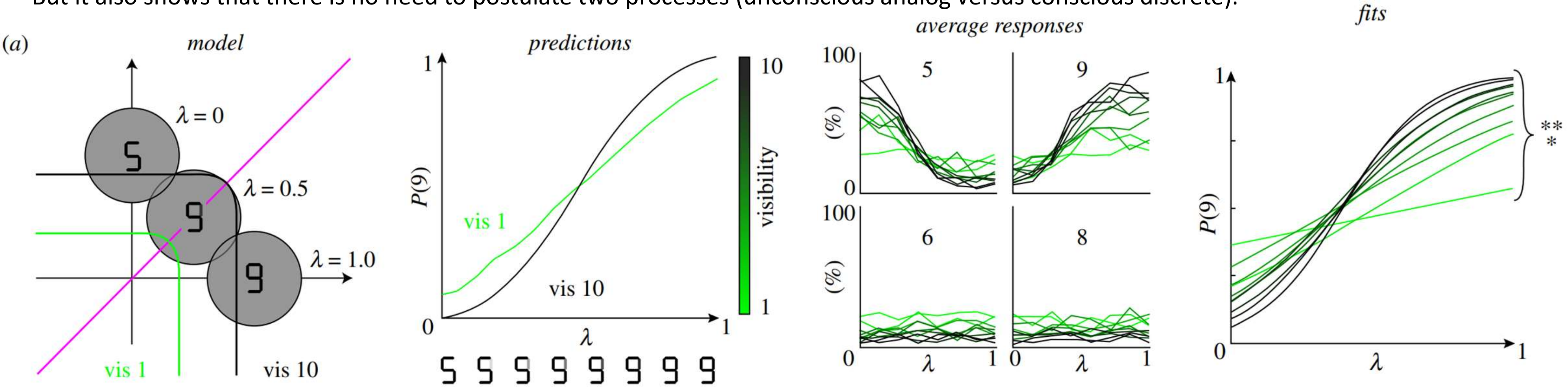
Subliminal processing and conscious access : a perspective from vector spaces

King, J.-R., & Dehaene, S. (2014). A model of subjective report and objective discrimination as categorical decisions in a vast representational space. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1641), 20130204. https://doi.org/10.1098/rstb.2013.0204

An empirical test of these ideas: subjective perception of masked digits, at the threshold of perception, while varying the evidence for one digit versus another.
The model predicts that the forced-choice discrimination profile **of physically identical stimuli** will become more nonlinear as visibility increases.
Results: sigmoidal functions that become increasingly steep and non-linear as subjective visibility increases from 1 to 10.



This experiment replicates and extends earlier work by De Gardelle, Charles and Kouider (2011): linear profile of forced-choice discrimination for masked stimuli, non-linear sigmoid for seen stimuli.
But it also shows that there is no need to postulate two processes (unconscious analog versus conscious discrete).



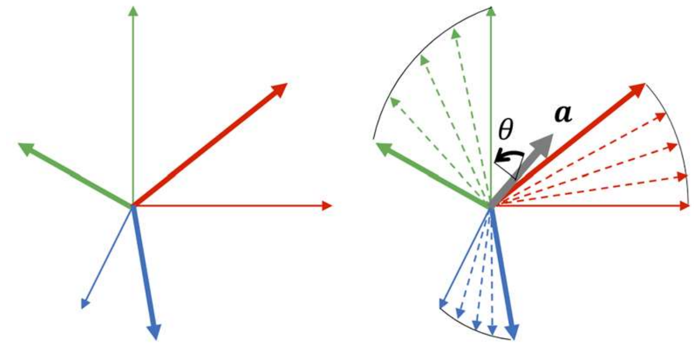
Dynamics of decision making : neuronal vectors change over time

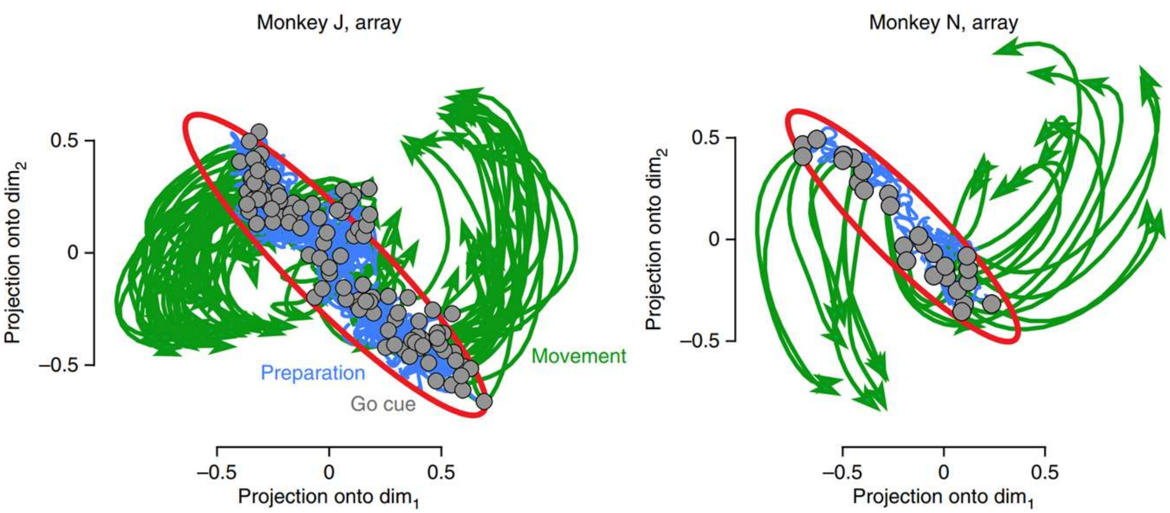
1. Information is typically not encoded by a single fixed vector, but by multiple vectors, a neural code that changes over time.

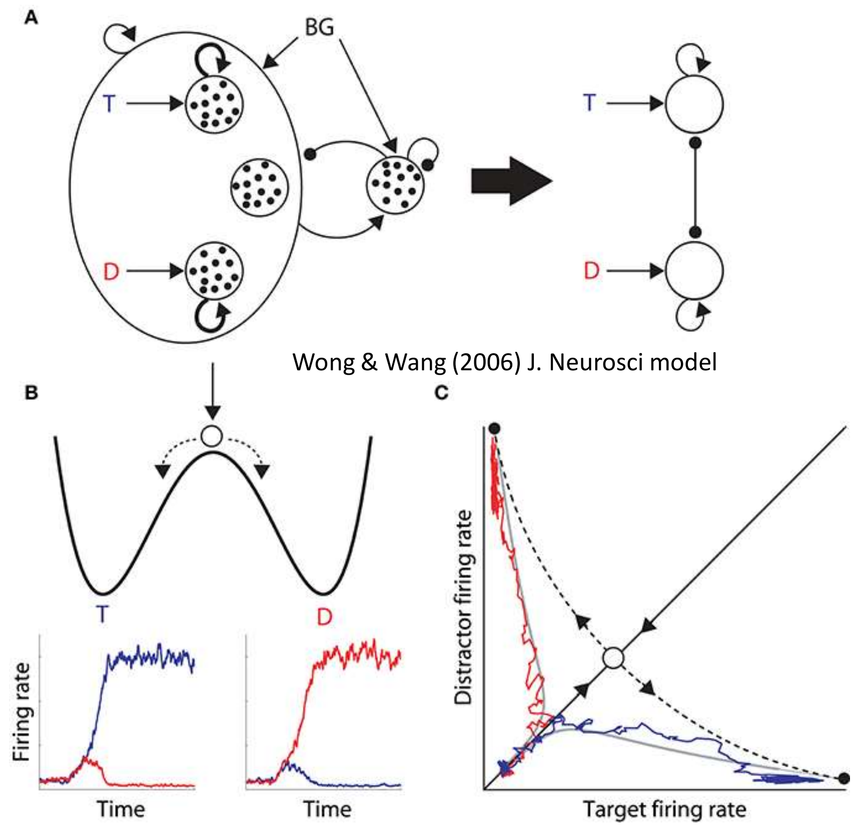
2. Vector change over time may reflect
- the passing of a neural representation from one area to the next
- A computation, i.e. a transformation of the content, for instance from perception to decision

3. Computation is seen as a vector flow

4. The vector perspective helps to visualize such transformations, which would be very hard to see at the single neuron level.







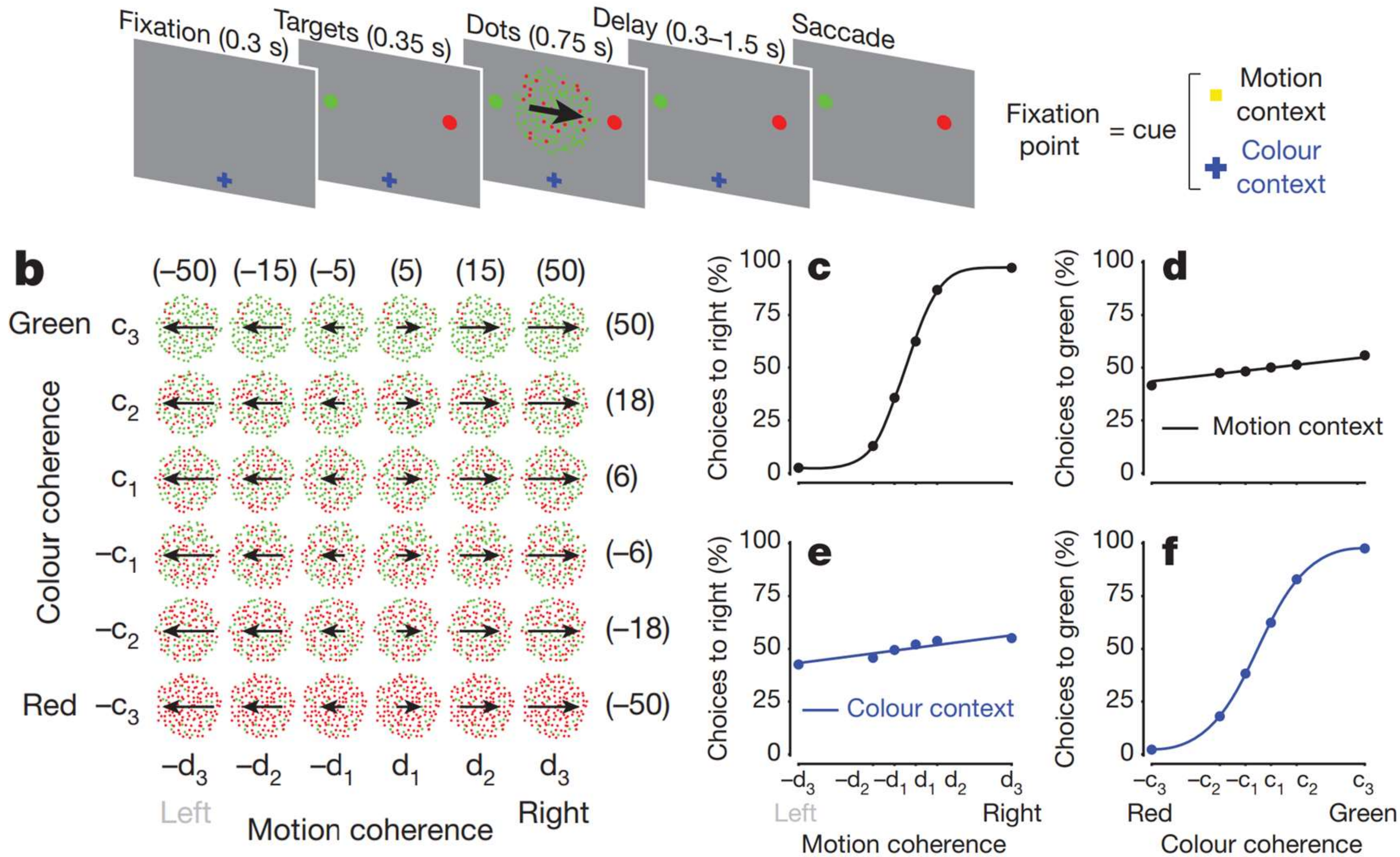
Wong & Wang (2006) J. Neurosci model

Dynamics of decision making in prefrontal cortex

Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature, 503(7474), 78-84. https://doi.org/10.1038/nature12742

Two monkeys were trained to perform two distinct tasks on the same stimuli: a color task and a motion task.
Behavior indicates that monkeys attend to the relevant dimension.

Question: how is the relevant dimension "routed" to output?
Is it "gated" by attention?
No! both dimensions are coded, but only one is transmitted downstream.



Dynamics of decision making in prefrontal cortex

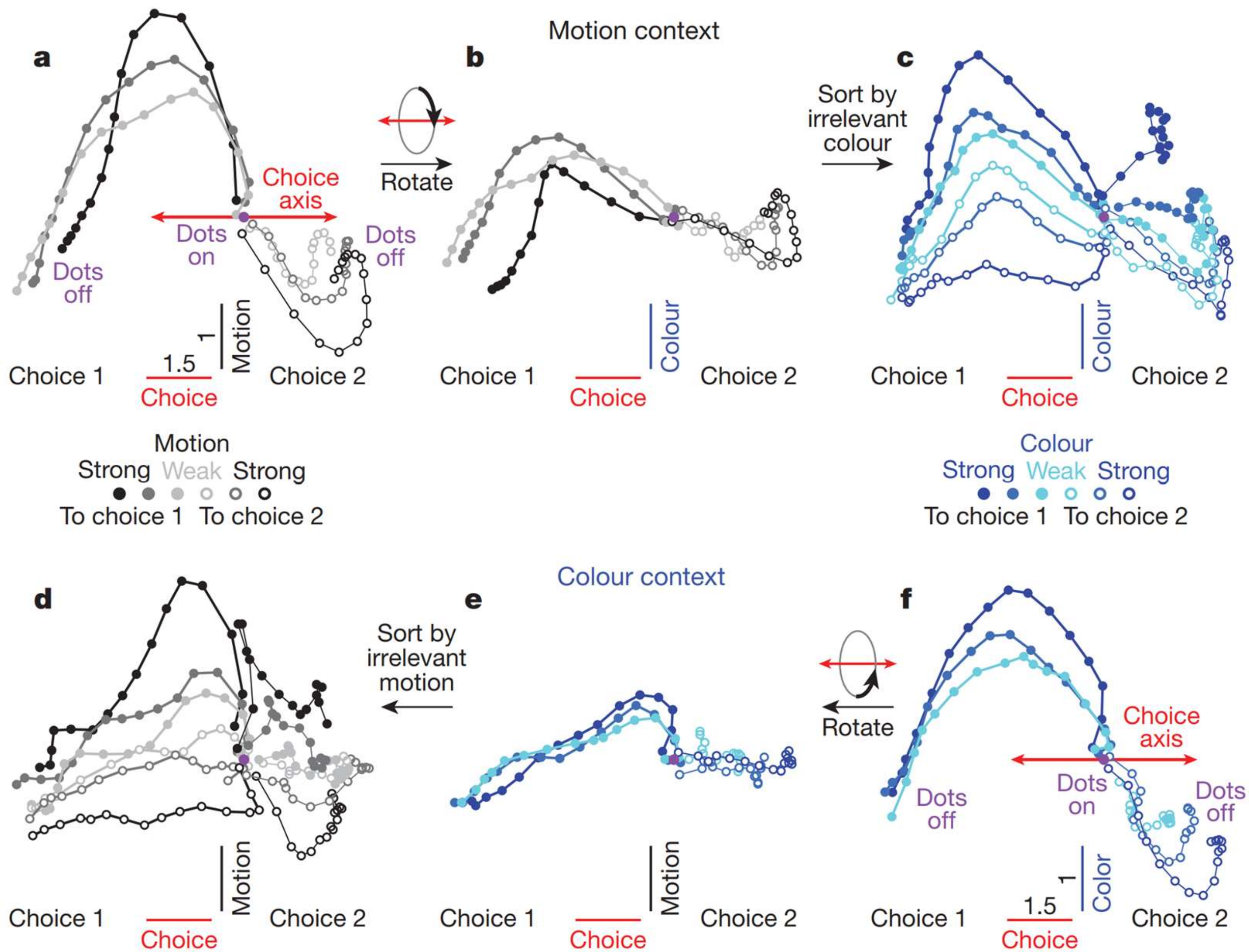
Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013) Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature, 503(7474), 78-84.

The authors reduce their large recordings (388 single-units and 1014 multi-unit)

- First to 12 dimensions by PCA
- Then to 4 dimensions by regression onto choice, motion, color, and context.

Each curve shows the average dynamics on a given trial type, defined by stimulus strength and ultimate choice. What do we see?

1. **Gradual accumulation of evidence** along the choice axis (horizontal).
2. Relevant sensory dimensions (motion and color) produce very different dynamics, as shown by the curved trajectory. They indicate the **strength of momentary evidence.**
3. **Irrelevant sensory dimensions** are also represented in PFC, with very little or no reduction in activity !
4. The context axis (not shown here) shows an overall displacement in neural space, but perceptual and choice axes stay the same.



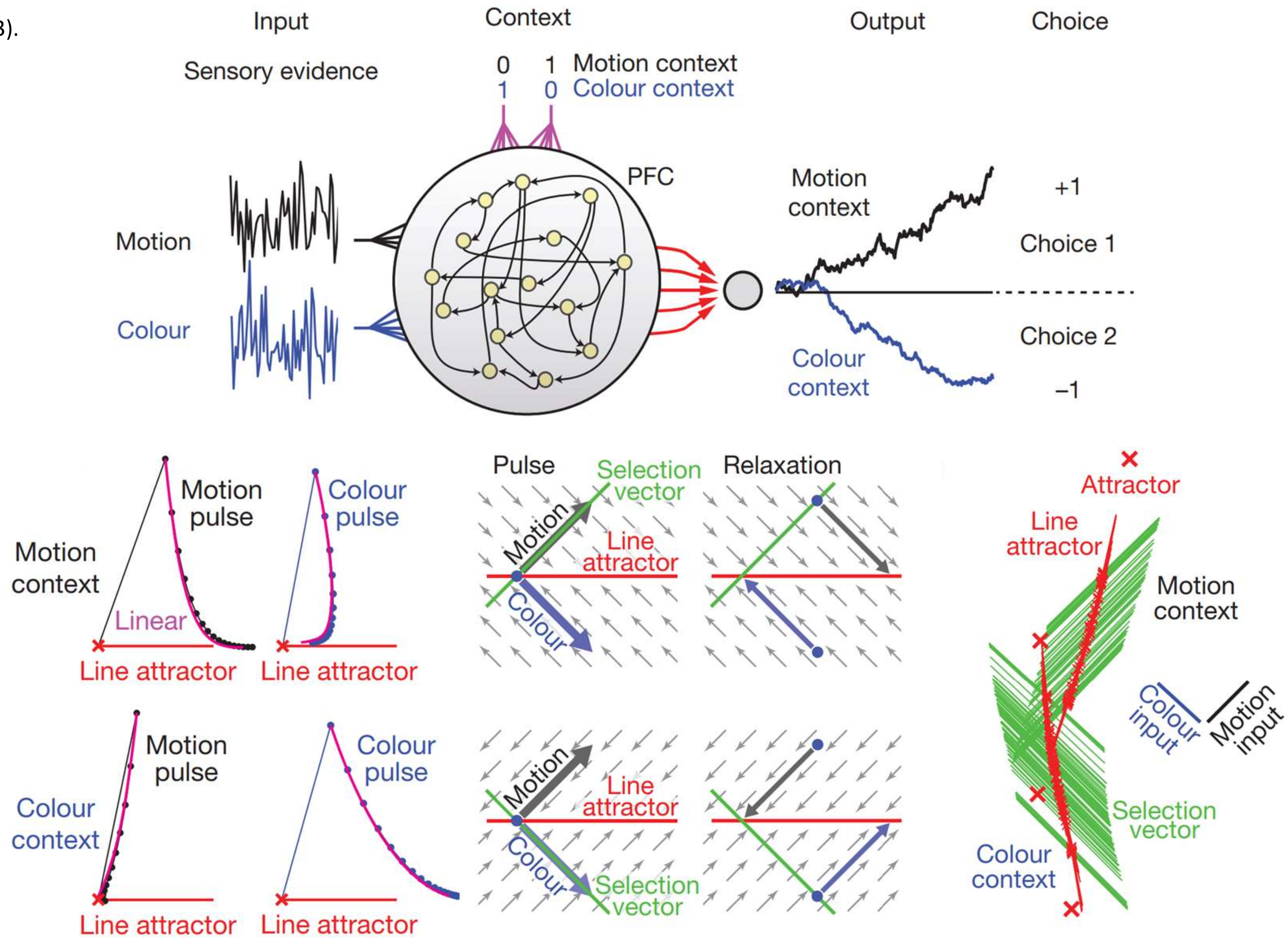
Dynamics of decision making in prefrontal cortex

Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature, 503(7474), 78-84.

A simulation (with random synapses and backprop training) helps understand what is going on.

1. The model develops a “line attractor” with stable fixed points, along which the model accumulates evidence for a given choice.
2. Perceptual evidence sends the activity away from this line attractor, and then activity relaxes towards the appropriate decision (for the relevant dimension) or towards zero (for the irrelevant dimension).
3. The dynamics changes according to context. The projection towards the decision axis is always perpendicular to a **selection axis**, whose orientation varies with context.

This is a nice example of how the dynamics of vector states implements collective decision making – imagine trying to understand this at the single-neuron level!



Computing in the “null space”

Kaufman, M. T., Churchland, M. M., Ryu, S. I., & Shenoy, K. V. (2014). Cortical activity in the null space : Permitting preparation without movement. Nature Neuroscience, 17(3), 440-448. https://doi.org/10.1038/nn.3643

The concept of “vector subspace” can shed light on a mystery:

When monkeys are instructed to prepare a movement, there is considerable activity in premotor (PMd) and motor cortex (M1) – yet without any muscle activity until the go cue.

This is true even though there are direct projections from both of these areas to the spinal chord.

Maybe the activity is “below threshold”? But preparatory activity is not a weaker version of motor activity, nor is there any sign of a “gate”.

Proposed solution: Motor activity M is some function of neural activity N – suppose for simplicity that this relation is linear: M = W N

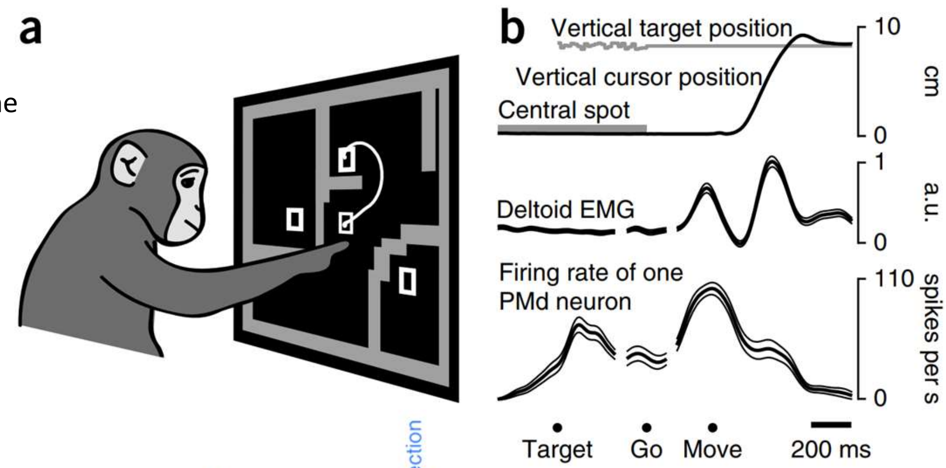
Then, because there are fewer muscles than neurons, there must be many combinations of neural firing that do not change anything at the muscle level – the “**null space**”.

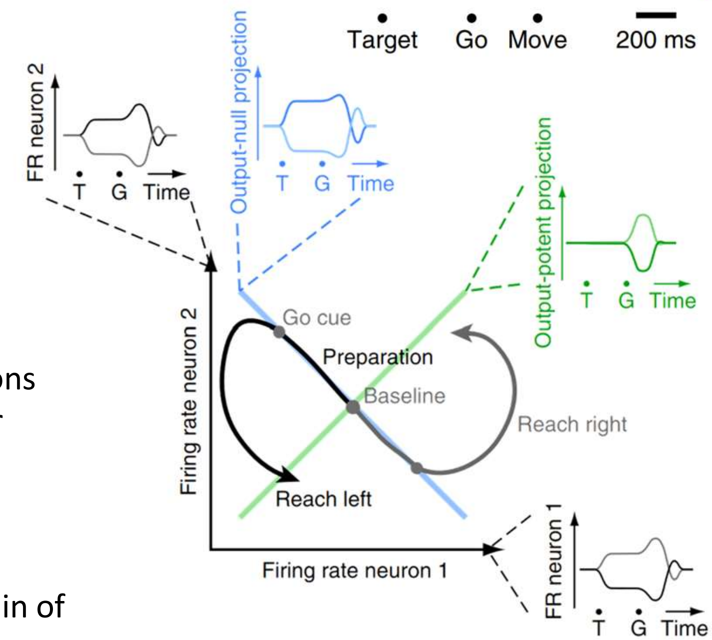
For instance, imagine that the only thing that affects the muscles is the sum of two neurons (green line). Then their difference (perpendicular blue line) does not matter, as long their sum stays constant.

Neurons can change their activity along the blue line (null space) in order to prepare for movement, and then rotate or project this activity onto the output space in order to act.

Mathematically, the null space or kernel (“noyau”) of a linear transformation is the domain of the source space which is mapped to the zero vector:

Two vectors have the same output *iff* they differ by a vector that belongs to the null space.

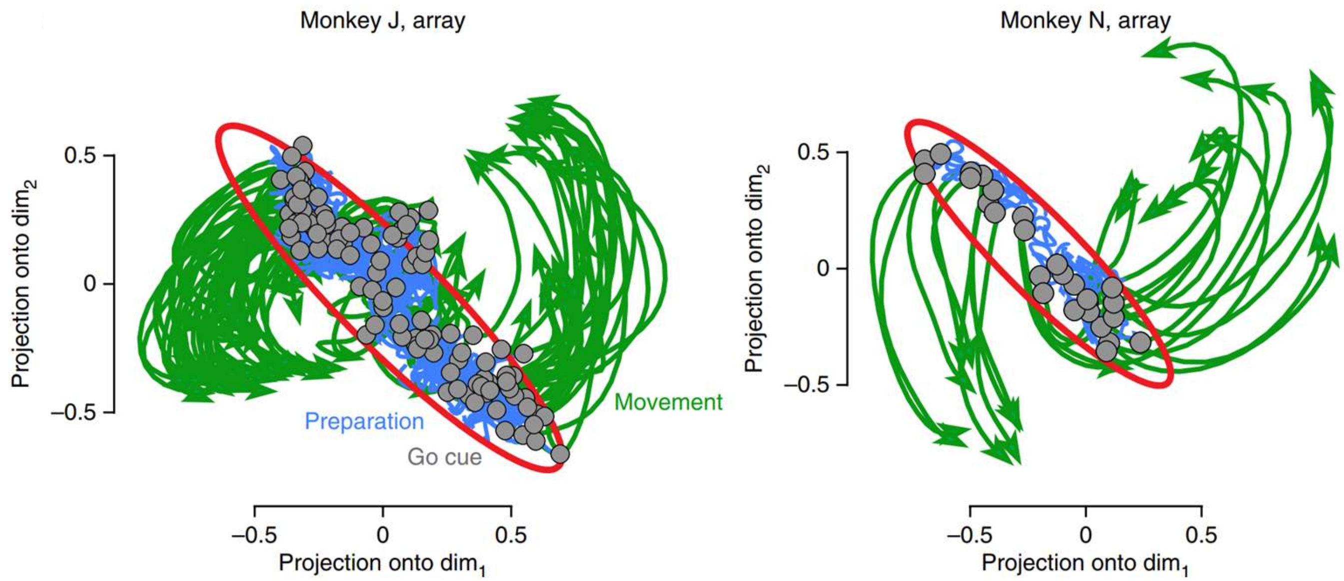




$$\ker(L) = \{\mathbf{v} \in V \mid L(\mathbf{v}) = \mathbf{0}\}$$

Movement preparation in the null space

Kaufman, M. T., Churchland, M. M., Ryu, S. I., & Shenoy, K. V. (2014). Cortical activity in the null space : Permitting preparation without movement. Nature Neuroscience, 17(3), 440-448. https://doi.org/10.1038/



Is this story true? To test it, the authors record from tens of neurons in PMd and M1.

It is possible to find weighted sums of those data such that the preparatory activity cancels out.

For instance, in the graph at right, each gray dot represents one average condition of movement through the maze. Along the blue axis, all of their preparatory activity projects identically – but during the movement, their activity vectors rotate and predict movement.

Note that preparatory activity is far from random or noisy – it is predictive of the amplitude and orientation of the upcoming movement.

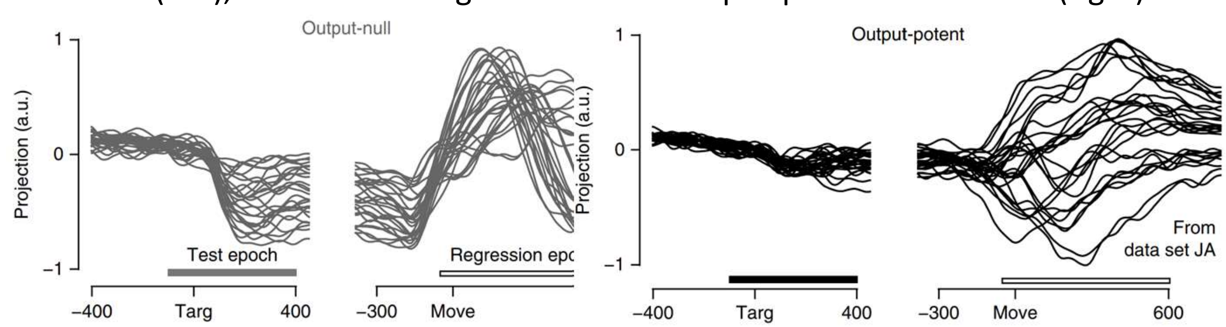
How can we identify the output space and null space?

First, perform data reduction using principal component analysis, thus reducing muscle activity M to $\widetilde{M}$ (with 3 dimensions), and neural activity N to $\widetilde{N}$ (with 6 dimensions).

Then use regression to solve for $\widetilde{W}$ such that $\widetilde{M} = \widetilde{W}\widetilde{N}$

And finally project the entire trajectory onto (1) the null space of $\widetilde{W}$, and (2) the other orthogonal space.

Results: potent preparatory activity in the output-null dimensions of PMd and M1 (left), 3 to 8 times larger than in the output-potent dimensions (right).



Movement preparation in the null space

Kaufman, M. T., Churchland, M. M., Ryu, S. I., & Shenoy, K. V. (2014). Cortical activity in the null space : Permitting preparation without movement. Nature Neuroscience, 17(3), 440-448. https://doi.org/10.1038/nn.3643

Remarkably, the same mechanism accounts for communication between PMd and M1.

Preparatory activity in PMd occurs primarily in the null space for activating M1.

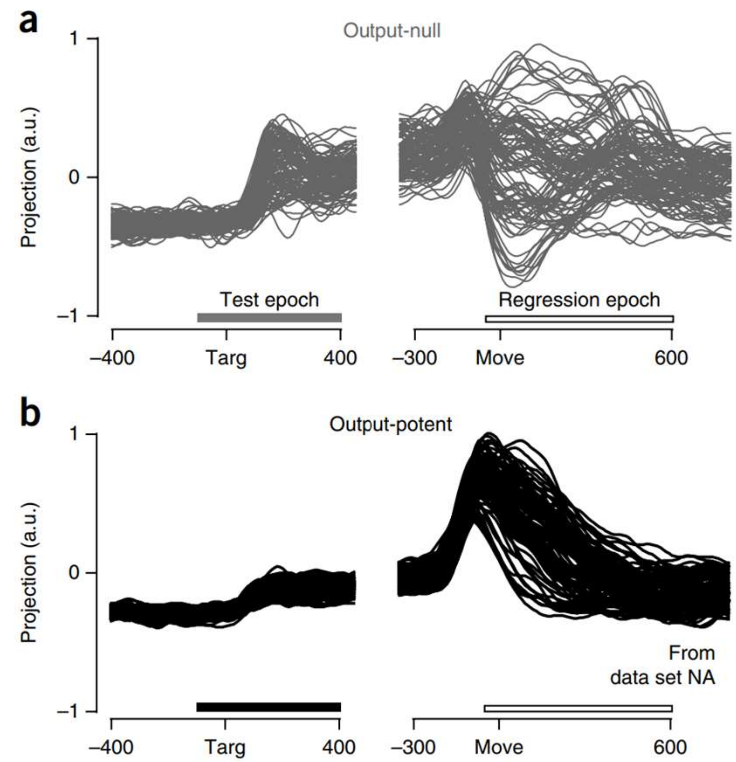
This reduction (not complete cancellation) can explain why M1 activity is smaller during preparation than during movement.

Controls:

- The converse is not true (when analyzing in the M1→ PMd direction)
- The conclusions do not depend on the choice of dimensionality
- The result is not due to a segregation of neurons specialized for preparation versus execution, but to different linear combinations of the same neurons.

Conclusion:

This could be a general mechanism for cortical communication – only a subpart of the activity is transmitted to another downstream area (or to muscles).

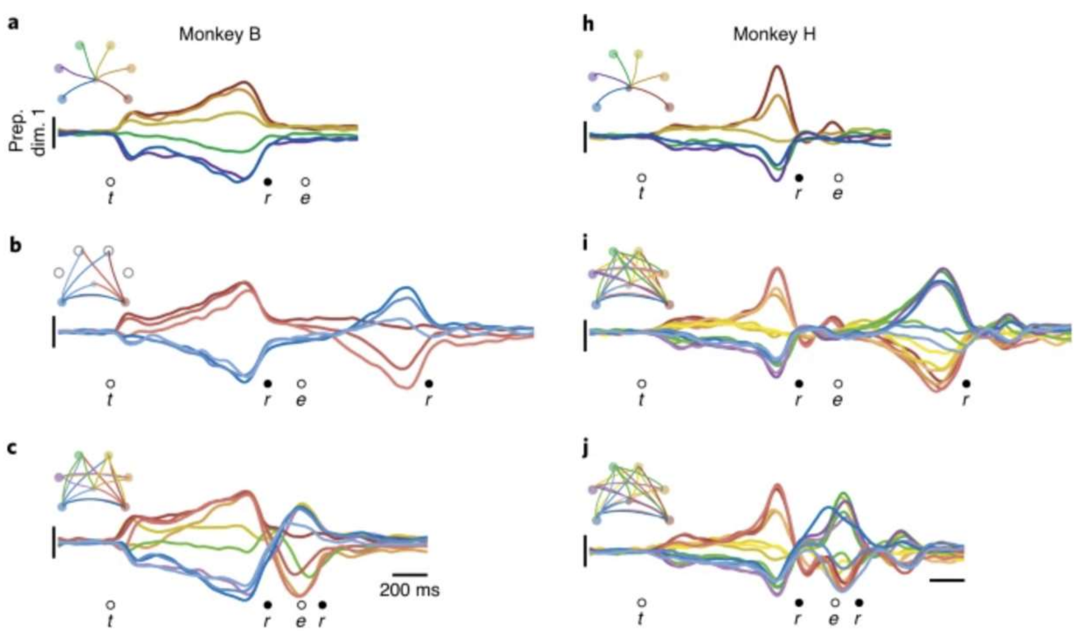


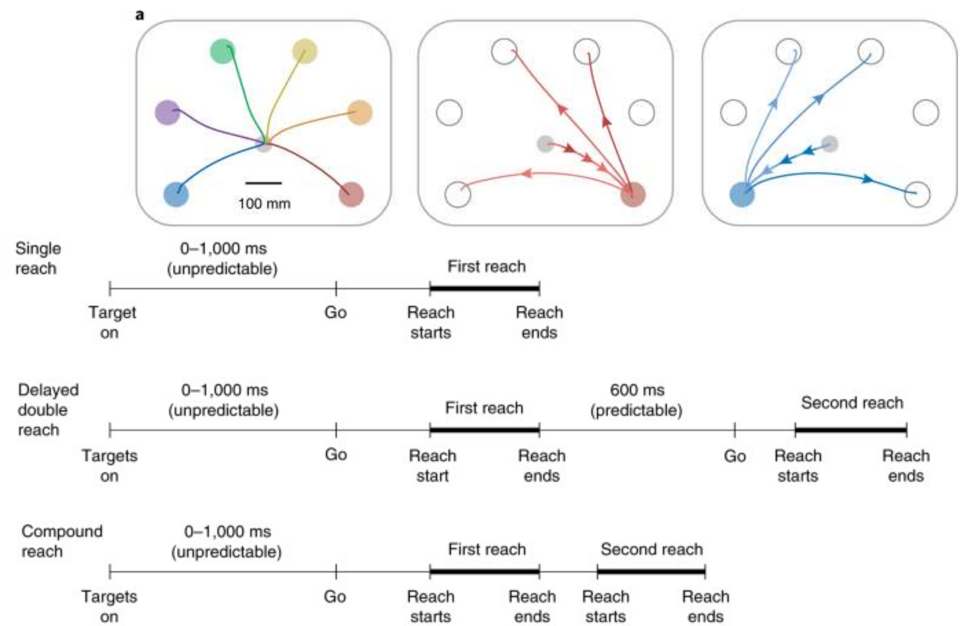
What happens during a fast movement? Parallel, independent subspaces for preparation, triggering and execution

Zimnik, A. J., & Churchland, M. M. (2021). Independent generation of sequence elements by motor cortex. *Nature Neuroscience*, *24*(3), 412-424.

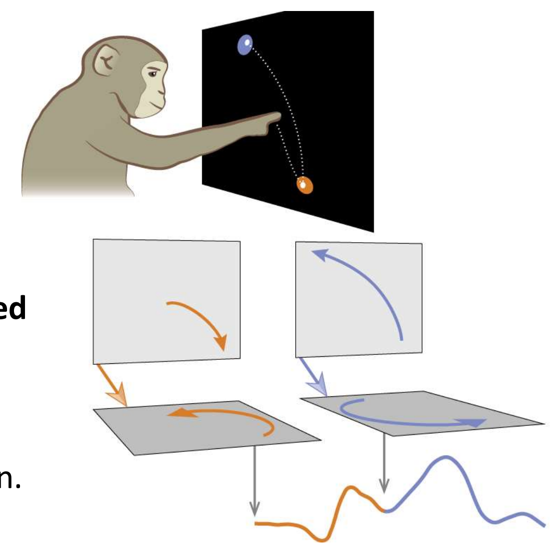
Monkeys trained to performed either a single reach or a double reach (spaced or immediate).
Results: the same component subspaces for preparation, triggering and execution are activated, only faster and partially in parallel with each other.

Fig. 4: Time-course of activity in preparatory dimensions.





The results allow to reject a model where speeded compound movements are generated by ad-hoc holistic neural activity. Activity can be **factorized** into independent subspaces for preparation, then triggering and execution.



Routing of neural information using neural subspaces

Semedo, J. D., Zandvakili, A., Machens, C. K., Yu, B. M., & Kohn, A. (2019). Cortical Areas Interact through a Communication Subspace. Neuron, 102(1), 249-259.e4. https://doi.org/10.1016/j.neuron.2019.01.026

Key proposal : **“The ability of a source area to communicate only certain activity patterns while keeping others ‘‘private’’ could be a means for the selective routing of signals between areas.”**

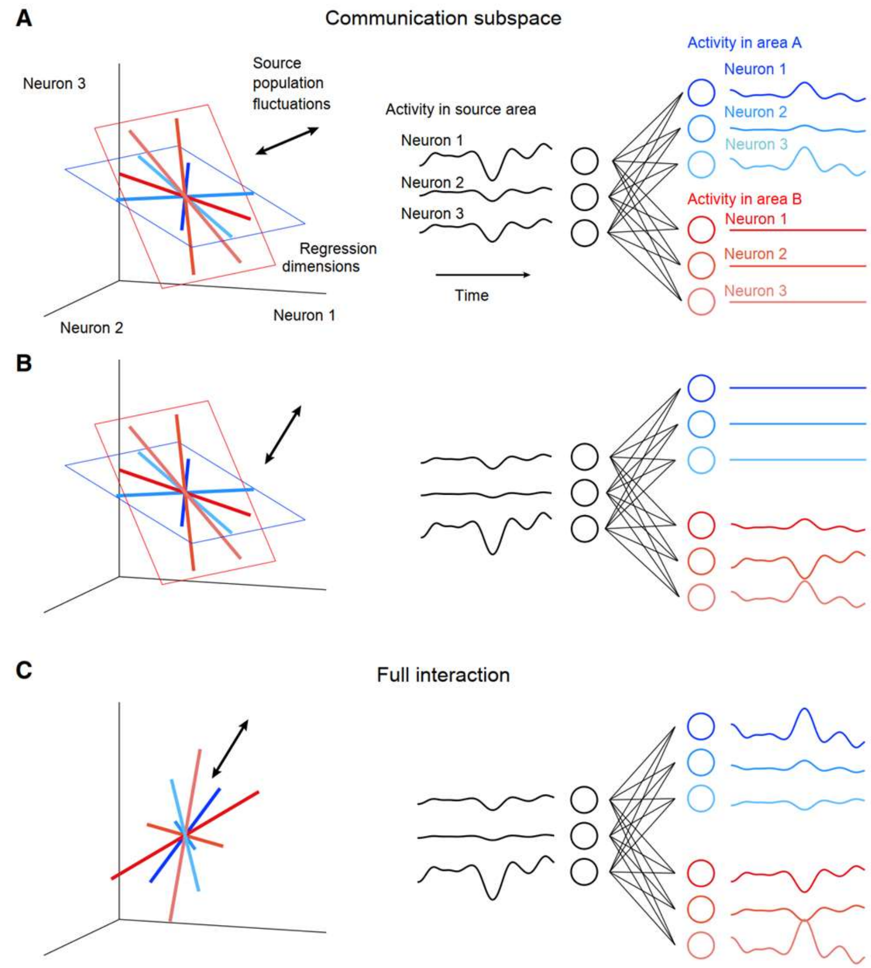
→ Completely different mechanism than the “communication through coherence” hypothesis (no need to set-up oscillations jointly in both areas)

→ More flexibility: different subspaces could be used to route information to different areas.

“The selective routing allowed by the communication subspace could be adjusted dynamically, allowing moment-to-moment modulation of interactions between cortical areas. Dynamic routing could be accomplished by altering the structure of population activity in a source area; it need not involve changing the communication subspace itself”

What the authors mean here is that a **projection** or **rotation** could be used to bring information to the appropriate subspace, thus opening or closing communication channels at will.

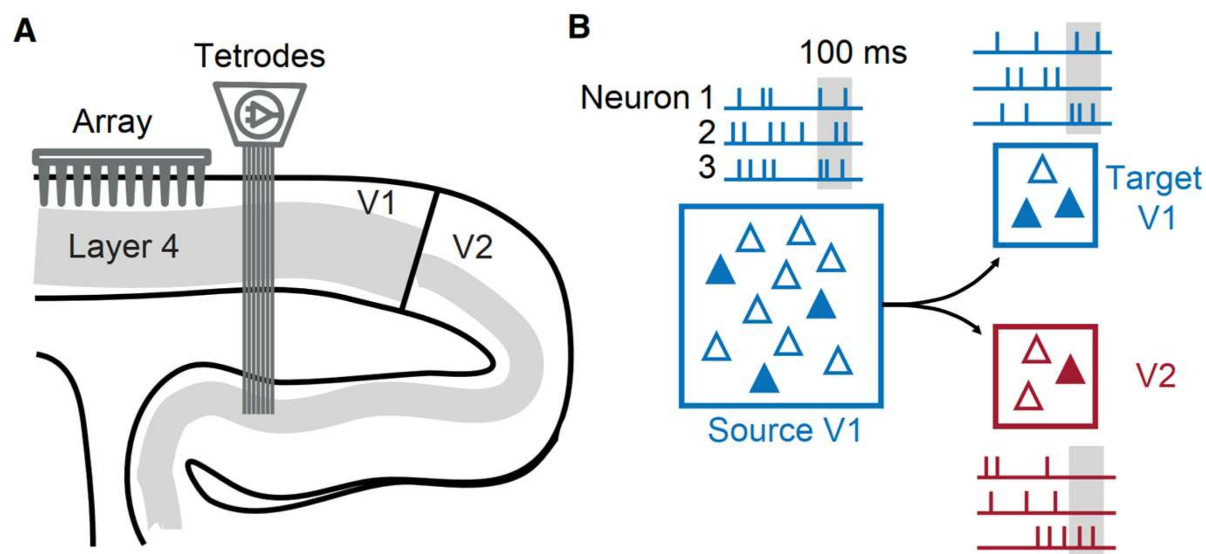
Meanwhile, the *private* dimensions could be used to perform covert computations.



Neuronal subspaces for communication between visual areas

Semedo, J. D., Zandvakili, A., Machens, C. K., Yu, B. M., & Kohn, A. (2019). Cortical Areas Interact through a Communication Subspace. Neuron, 102(1), 249-259.e4. https://doi.org/10.1016/j.neuron.2019.01.026

- Goal = measure the communication of dense information from V1 to V2, using simultaneous recordings
- In V1: Recordings from layer 2-3 projection cells, 88 to 159 neurons
- In V2: movable tetrodes, recordings for middle layers, 24 to 37 neurons
- V1 and V2 sites share overlapping receptive fields.
- The monkey is anesthetized.
- Gratings are presented.
- The mean activity (PSTH = peri-stimulus time histogram) is subtracted, and they study the transmission of trial-by-trial variability.



Neuronal subspaces for communication between visual areas

Semedo, J. D., Zandvakili, A., Machens, C. K., Yu, B. M., & Kohn, A. (2019). Cortical Areas Interact through a Communication Subspace. Neuron, 102(1), 249-259.e4. https://doi.org/10.1016/j.neuron.2019.01.026

First finding:

The correlation between pairs of neurons is low and similar within V1 and between V1 and V2.

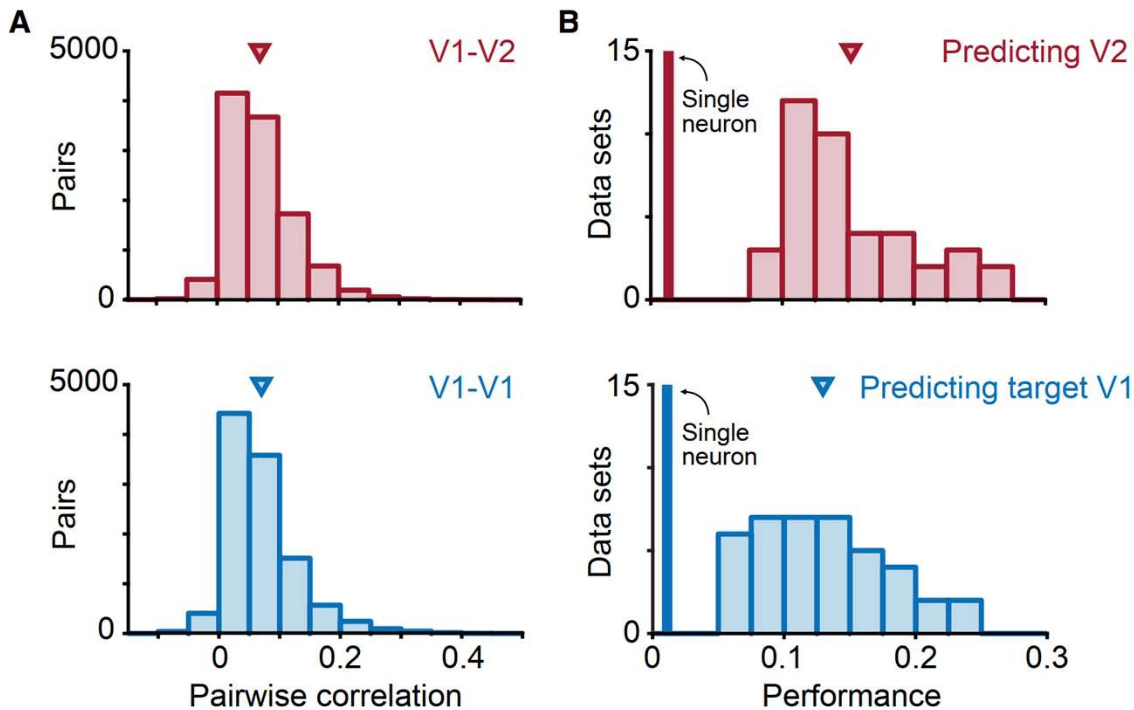
Single neurons can only explain ~1% of the variance in V2.

However, multivariate linear regression can explain ~13-15 % of the variance

(a level of performance which, according to simulations, is in line with Poisson noise in V2, or with subsampling of the true causal population in V1)

→ Populations of neurons in V1 influence V2. This is entirely expected, but…

- What is the size of the predictive population?
- Could a subset of V1 neurons suffice to predict V2 activity?
- Can we distinguish *private* V1 activity from *predictive (or public)* activity?

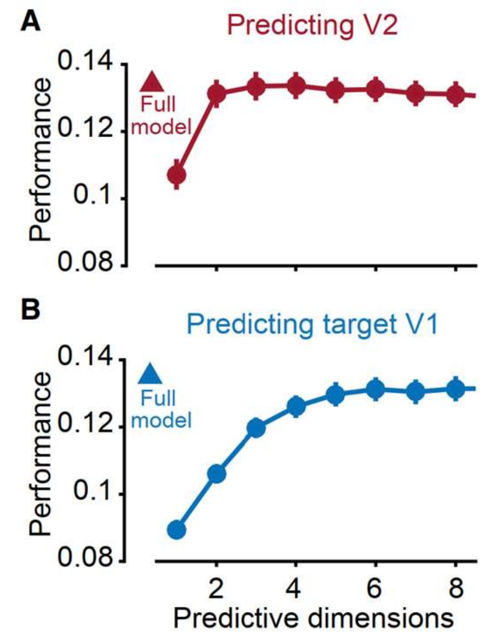


r

Cross-validated R^2

Neuronal subspaces for communication between visual areas

Semedo, J. D., Zandvakili, A., Machens, C. K., Yu, B. M., & Kohn, A. (2019). Cortical Areas Interact through a Communication Subspace. Neuron, 102(1), 249-259.e4. https://doi.org/10.1016/j.neuron.2019.01.026



Key Idea: estimate how many dimensions suffice to account for V1-V2 interactions.

Reduced-rank regression, a variant of linear regression in which the regression dimensions are constrained to lie in a low-dimensional subspace, can find a subset of *predictive* dimensions that are equally good at predicting V2 activity.

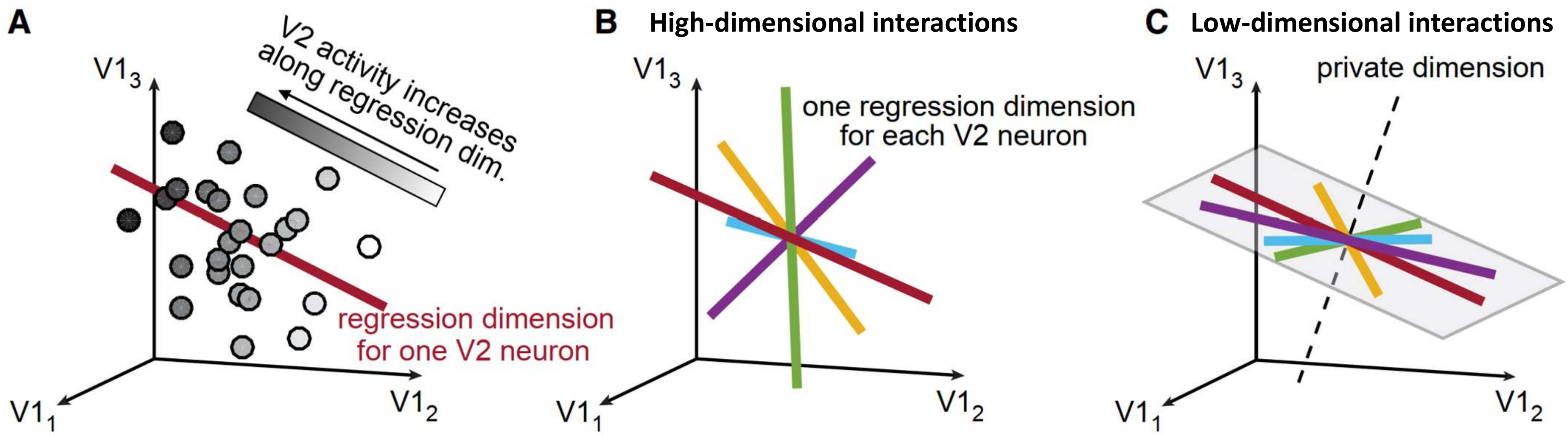
Results:

A small number of V1 dimensions (2) suffices to explain V2 activity

A larger number (5) is needed to explain V1-V1 interactions.

This finding is not due to the intrinsic dimensionality, which is actually larger for V2 data than for V1 data → existence of a limited *communication subspace* between V1 and V2.

Similar findings for V1-V4 interactions in awake macaque monkeys.



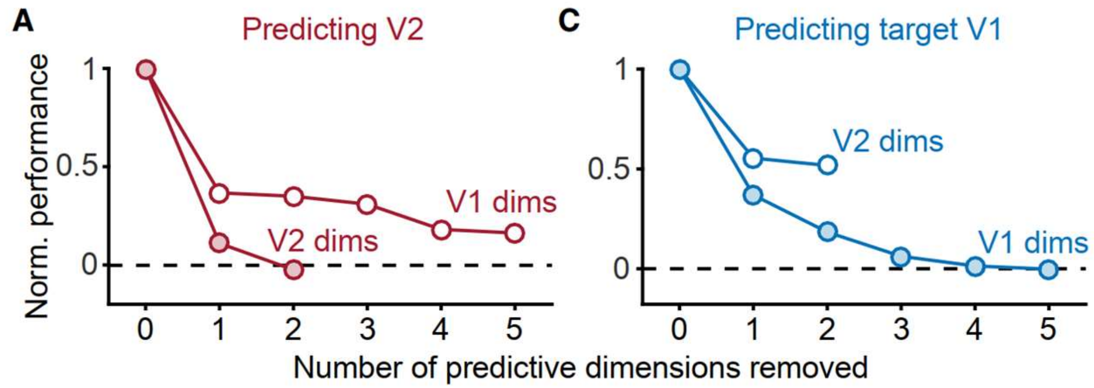
Neuronal subspaces for communication between visual areas

Semedo, J. D., Zandvakili, A., Machens, C. K., Yu, B. M., & Kohn, A. (2019). Cortical Areas Interact through a Communication Subspace. Neuron, 102(1), 249-259.e4. https://doi.org/10.1016/j.neuron.2019.01.026

Private and public spaces are different within area V1.

This can be seen by selectively removing some of the axes from the V1 population:

- Removing private V1 dimensions does not (fully) prevent predicting V2 activity, while removing public ones does.
- The converse is true for predicting V1 activity.



Additional controls:

- A similar communication subspace is found when analyzing responses to a single input grating. It is not identical, but generalizes to other gratings (75% of the original variance) and to movie-induced brain activity.
- The communication subspace cannot be identified by principal component analysis (PCA) – it does not correspond to the maximum variance of V1 activity (but V1-V1 predictive dimensions do)
- The article proves the existence of a restricted communication subspace, but probably does not correctly estimate its true number of dimensions (larger datasets, with more receptive fields, would probably lead to a larger number of dimensions).

Conclusions

A **low-dimensional representation** in inferotemporal cortex can support **very fast encoding of new categories.**

Decisions correspond to **classifications** in this space, which can be seen as the **projection or rotation** of the internal coding vector onto the response categorization axis.

The decision boundaries can explain the psychophysics of **conscious and unconscious decision making**

Only a small number of neurons may communicate their outputs to other brain regions.

This constraint creates an opportunity :

- existence of a **vast "null space" for covert internal computations**.
- **Selective communication** between any two areas by projection onto the relevant subspace

Vendredi 6 Janvier
COURS : **Vecteurs neuronaux ou cellules grand-mère : les représentations mentales sont-elles localisées ou distribuées ?**
SÉMINAIRE : **L'intelligence artificielle peut-elle modéliser le langage mathématique ?** — François Charton (FAIR Paris)

Vendredi 13 Janvier
COURS : **Géométrie des représentations visuelles : chaque visage est un vecteur**
SÉMINAIRE : **Commonsense Physical Reasoning in man and machine** — Ernest Davis (NYU, par zoom)

Vendredi 20 Janvier
COURS: **Exploiter la factorisation et les sous-espaces vectoriels pour coder l'information et communiquer entre aires cérébrales**
SÉMINAIRE : **Number symbols in the brain and mind** — Daniel Ansari (University of Ontario)

Vendredi 27 Janvier
COURS : **Comment prendre une décision ou faire des calculs avec des vecteurs dynamiques?**
SÉMINAIRE : **Comment se développent les réseaux cérébraux associés aux concepts mathématiques ?** — Marie Amalric (Université de Trento, Italie)

Vendredi 3 Février
COURS : **La représentation vectorielle des mots et des concepts**
SÉMINAIRE : **Les succès et les nouveaux défis de l'intelligence artificielle en mathématiques** — Léon Bottou (FAIR, New York)

Vendredi 10 Février
COURS : **La représentation vectorielle du langage : Comment représenter une phrase ?**
SÉMINAIRE : **Intuitions of mathematics and their refinement with age and education** — Manuela Piazza (Université de Trento, Italie)