Cours 2025-2026:

Qu'est-ce que la conscience et quels sont ses mécanismes cérébraux ?

What is consciousness, and what are its neuronal mechanisms?

Stanislas Dehaene
Chaire de Psychologie Cognitive Expérimentale

Cours n°2

Profondeur et limites du traitement non-conscient : données récentes

Depth and limits of non-conscious processing : an update

Most processing in the human brain is nonconscious… but we are not conscious of it

A truism : « we cannot be conscious of what we are not conscious of ».

Deep consequences : Our conscious introspection only gives us access to a narrow layer of conscious processing.

When Alan Turing introspected on his mental activity as a mathematician, he came up with the "Turing machine" – a slow, serial, rule-based system which he saw it as a simplified view of 'a man in the process of computing a real number' and whose 'human memory is necessarily limited', such as he could only operate on a single symbol at a time.

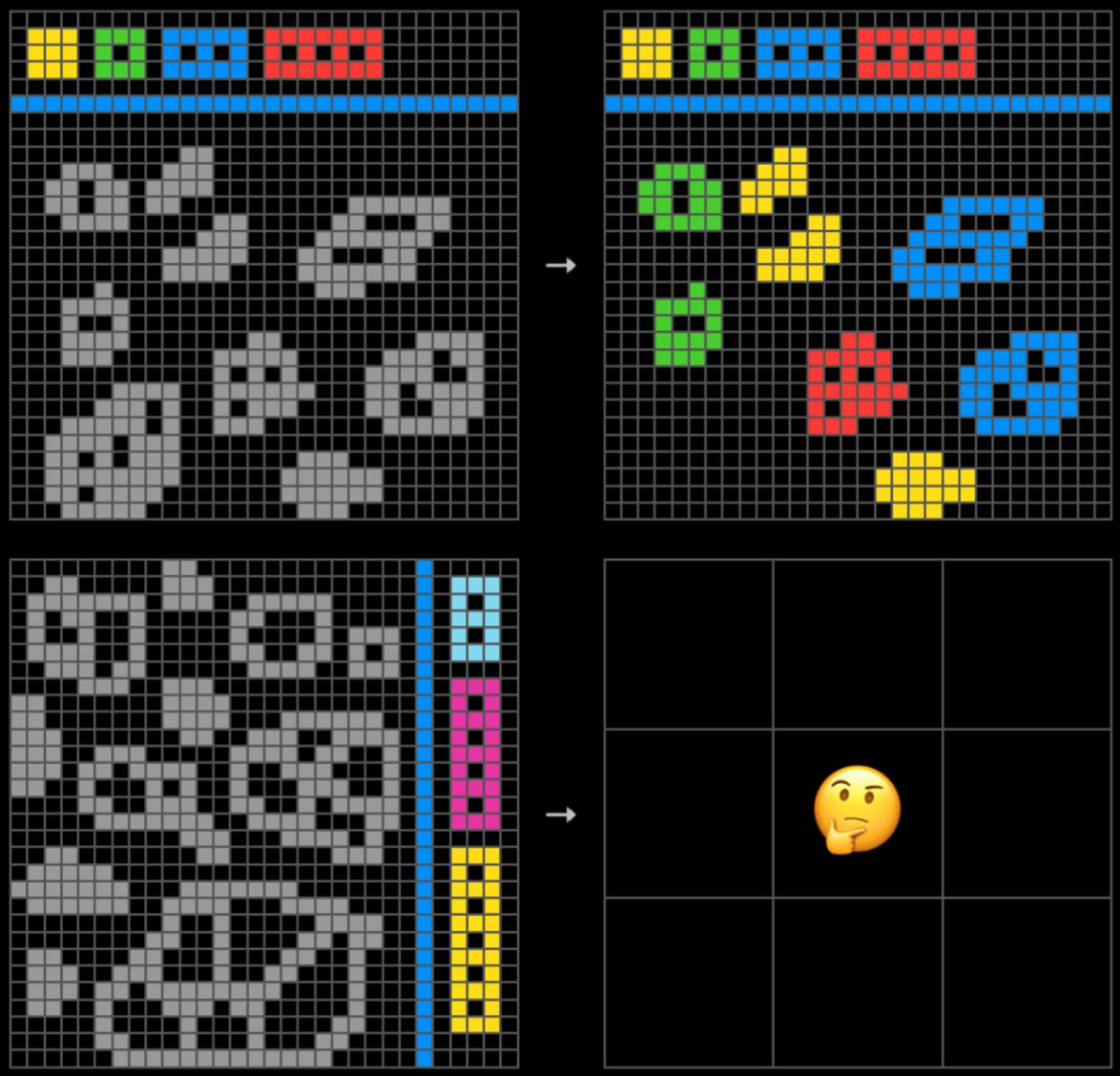
The Turing's machine metaphor may still hold for our conscious computations, which are indeed slow and serial (Kahneman's system 2), but Turing was unable to discover, by introspection alone, that, in the underlying hardware, the brain is actually a massively parallel analog machine – its computations are not accessible to introspection.

E.g. we cannot introspect how we recognize a face or compare two numbers (system 1).

but note that today, both systems are being increasingly captured by AI -- this week, Gemini 3 DeepThink (Google DeepMind) just reached 85% success in the ARC-AGI-2 challenge.

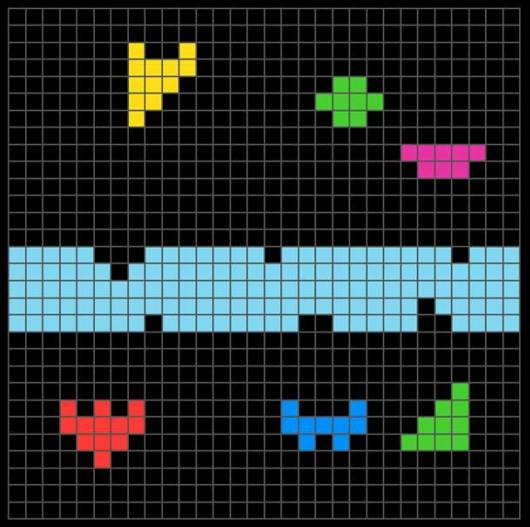
Nous avons dit qu'un nombre est calculable lorsque son expression décimale est calculable avec des moyens finis, définition que nous allons maintenant expliciter, sans tenter de la justifier avant la section 9. En attendant, je me contenterai de suggérer que cette définition est justifiée par le fait que la mémoire humaine est nécessairement limitée.

Un homme en train de calculer la valeur d'un nombre réel peut être comparé à une *machine* susceptible de se trouver dans un nombre fini d'états $q_1, q_2, \ldots, q_R$, que nous appellerons ses *m-configurations*. La machine est alimentée avec une *bande* (analogue au papier qu'utilise l'homme), divisée en

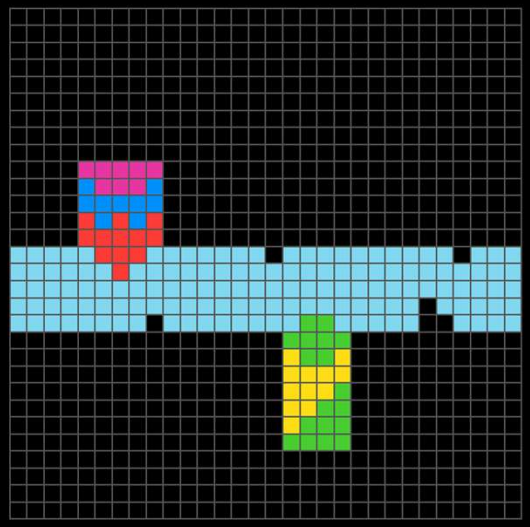


https://arcprize.org/

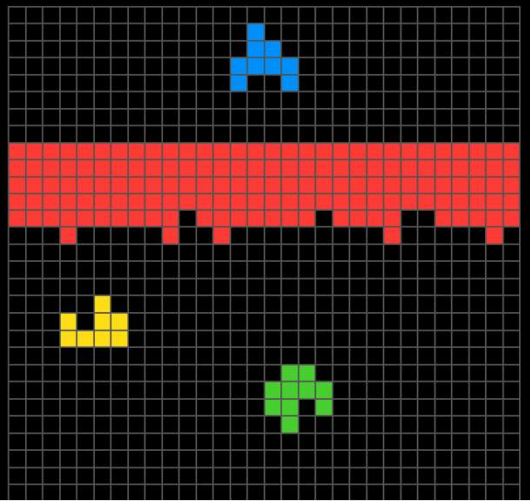
Ex.1 Input (30x30)



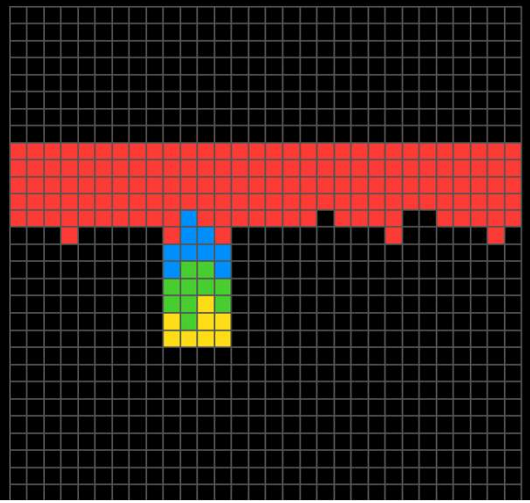
Ex.1 Output (30x30)



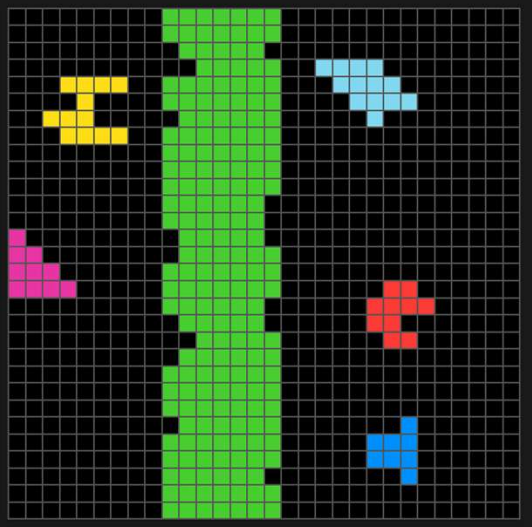
Ex.2 Input (30x30)



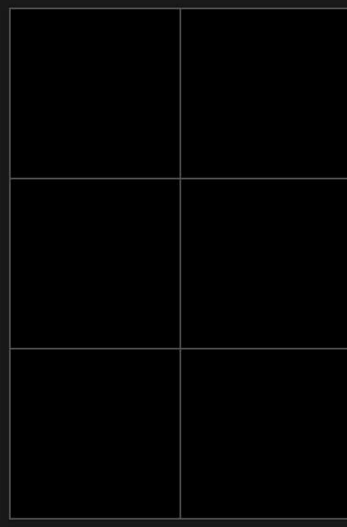
Ex.2 Output (30x30)



Input (30x30)



Output



1. Configure your output grid:

3x3 Resize Copy from input C

2. Edit your output grid cells:

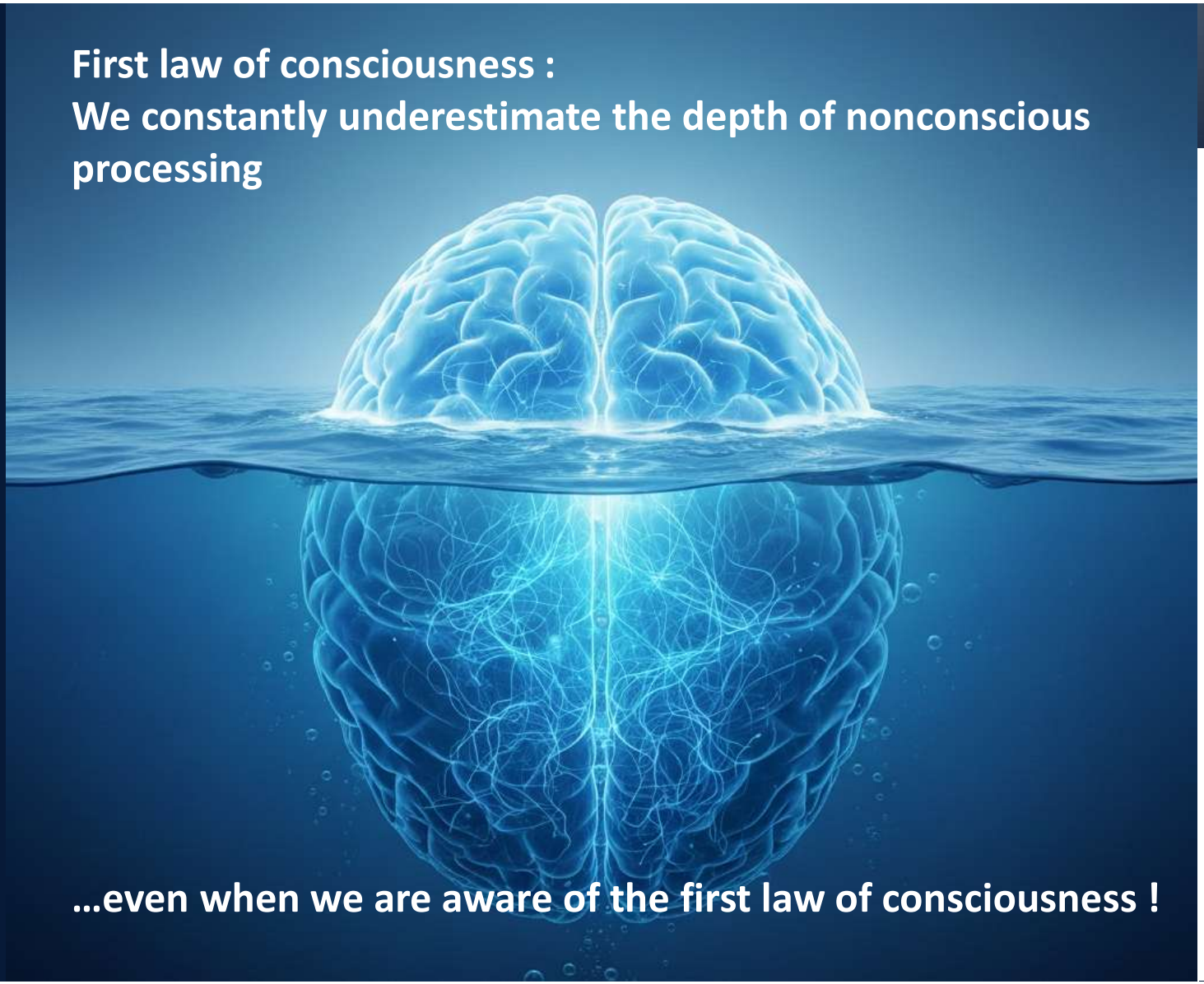
Edit Select Fill

3. See if your output is correct:

Submit solution Error loading task.

https://arcprize.org/

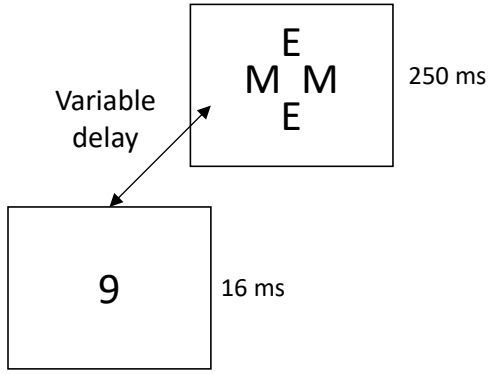
First law of consciousness :
We constantly underestimate the depth of nonconscious processing



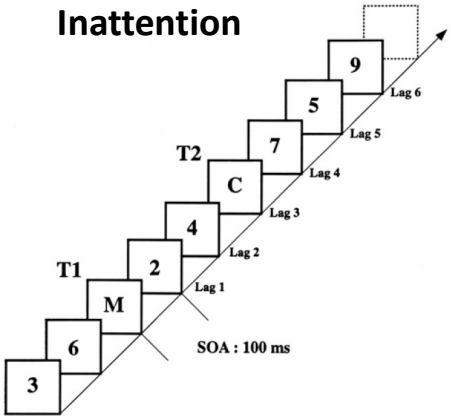
…even when we are aware of the first law of consciousness !

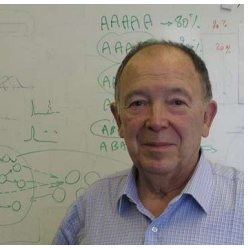
We need scientific paradigms to study the fate of nonconscious stimuli.

Masking



Inattention





Jean-Pierre Changeux

The global neuronal workspace (GNW) hypothesis

Dehaene & Changeux, PNAS 1998 ; Dehaene & Naccache, Cognition 2001

Lionel Naccache

Virtually every automatized chain of cortical processors can operate nonconsciously.

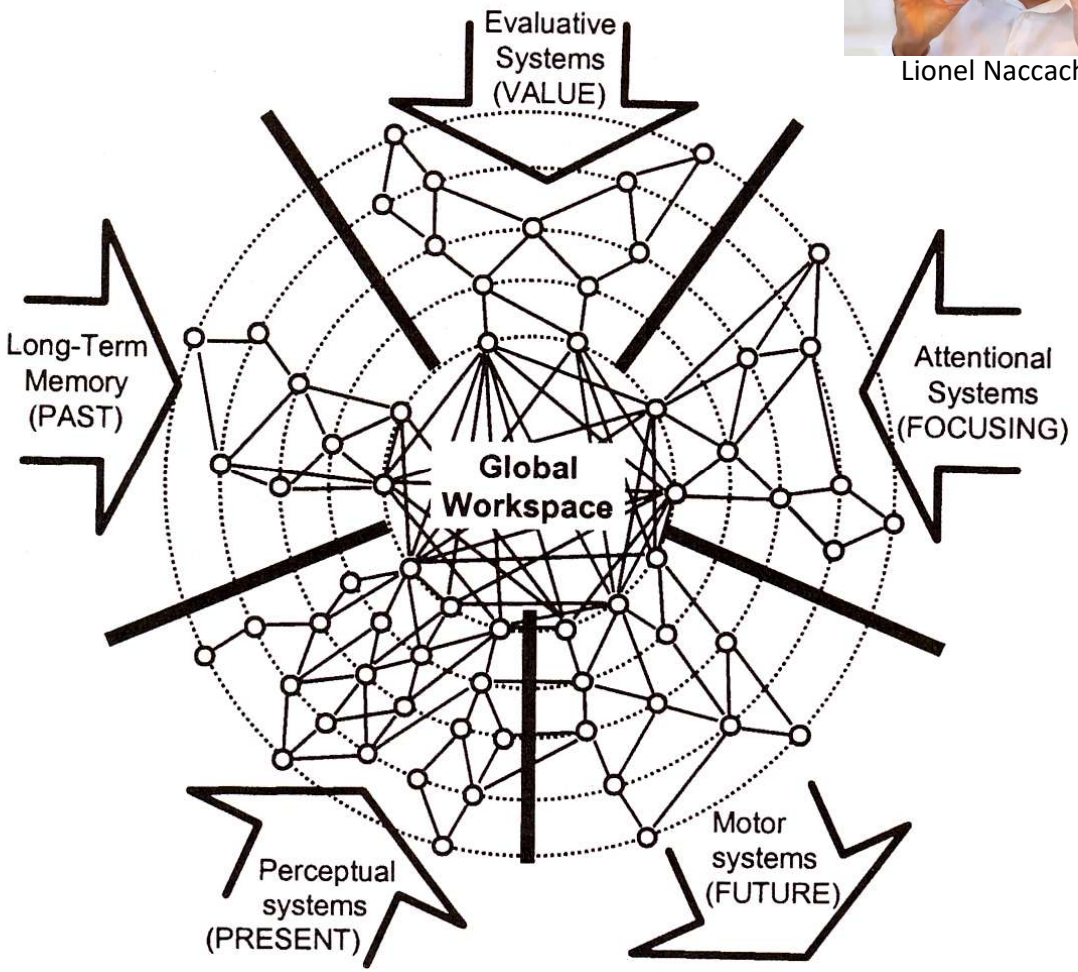
→ Visual recognition

→ Massively **parallel processing** of non-conscious stimuli.

→ **Language processing** : Up to what depth?

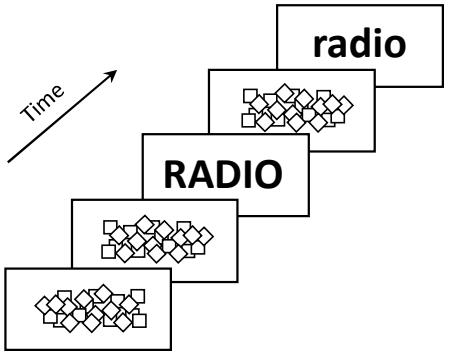
→ Possible limits on nonconscious computations:

- We may not be able to nonconsciously
- - **integrate information** flexibly (e.g. integrate several words into a sentence)
- Reflect upon ourselves (**metacognition**)

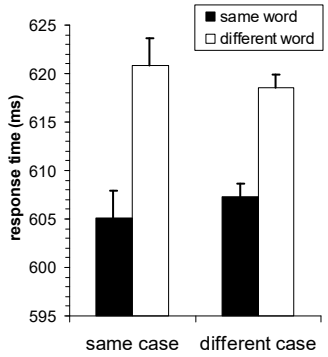


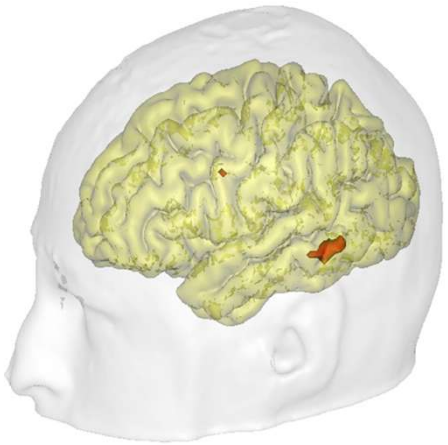
Brain imaging studies of nonconscious word recognition

Dehaene et al., Nature Neuroscience, 2001

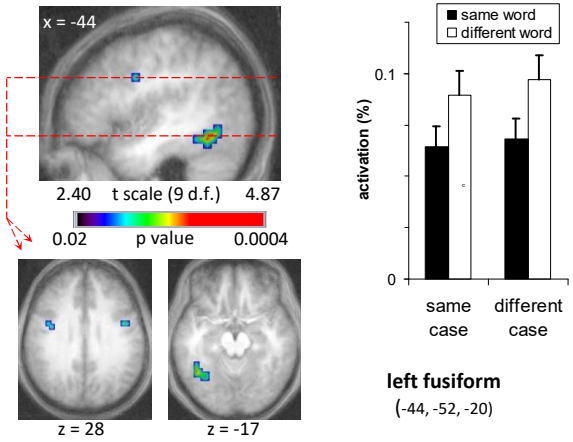


Behavior: case-independent repetition priming

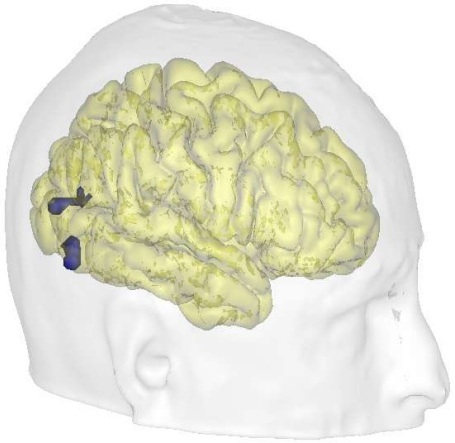




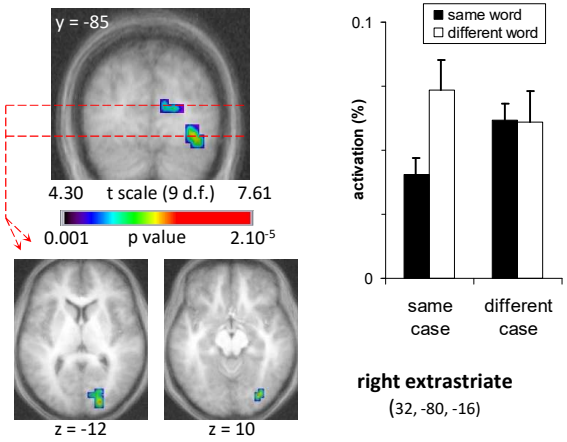
VWFA: case-independent priming



left fusiform
(-44, -52, -20)



Occipital cortex: case-specific priming



right extrastriate
(32, -80, -16)

A taxi conversation with Dan Dennett

FUNDAÇÃO ORIENTE

FUNDAÇÃO
CALOUSTE
GULBENKIAN

**INSTITUTO
GULBENKIAN
de CIÊNCIA**

CONVENTO DA ARRÁBIDA

September 5 to September 7 1999

***« Stan, are we so sure that
there is no unconscious binding? »***

Does letter binding occur without consciousness?

Dehaene et al., *Psychological Science,* 2004

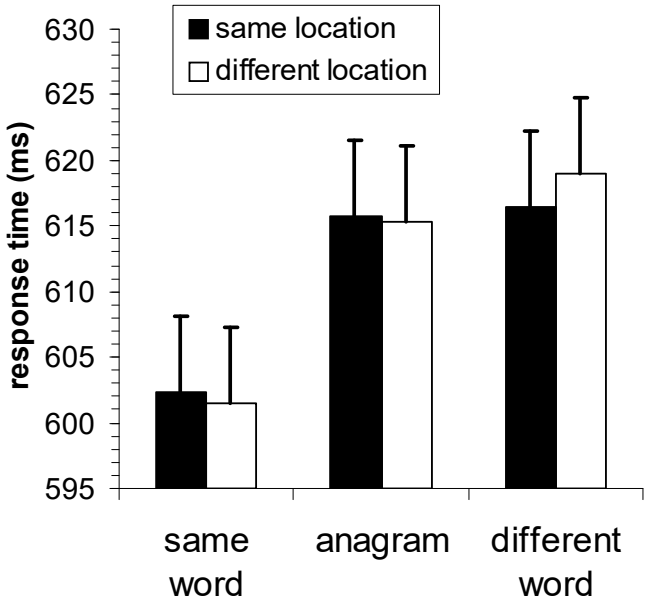
What counts as a repetition?
Same letters? Same words?

ANGER

word locations: **same** / **different**

Prime-Target Relation

	same word	anagram	different word
same	#REFLET #reflet	#TREFLE #reflet	#PATERE #reflet
different	REFLET# #reflet	TREFLE# #reflet	EPATER# #reflet



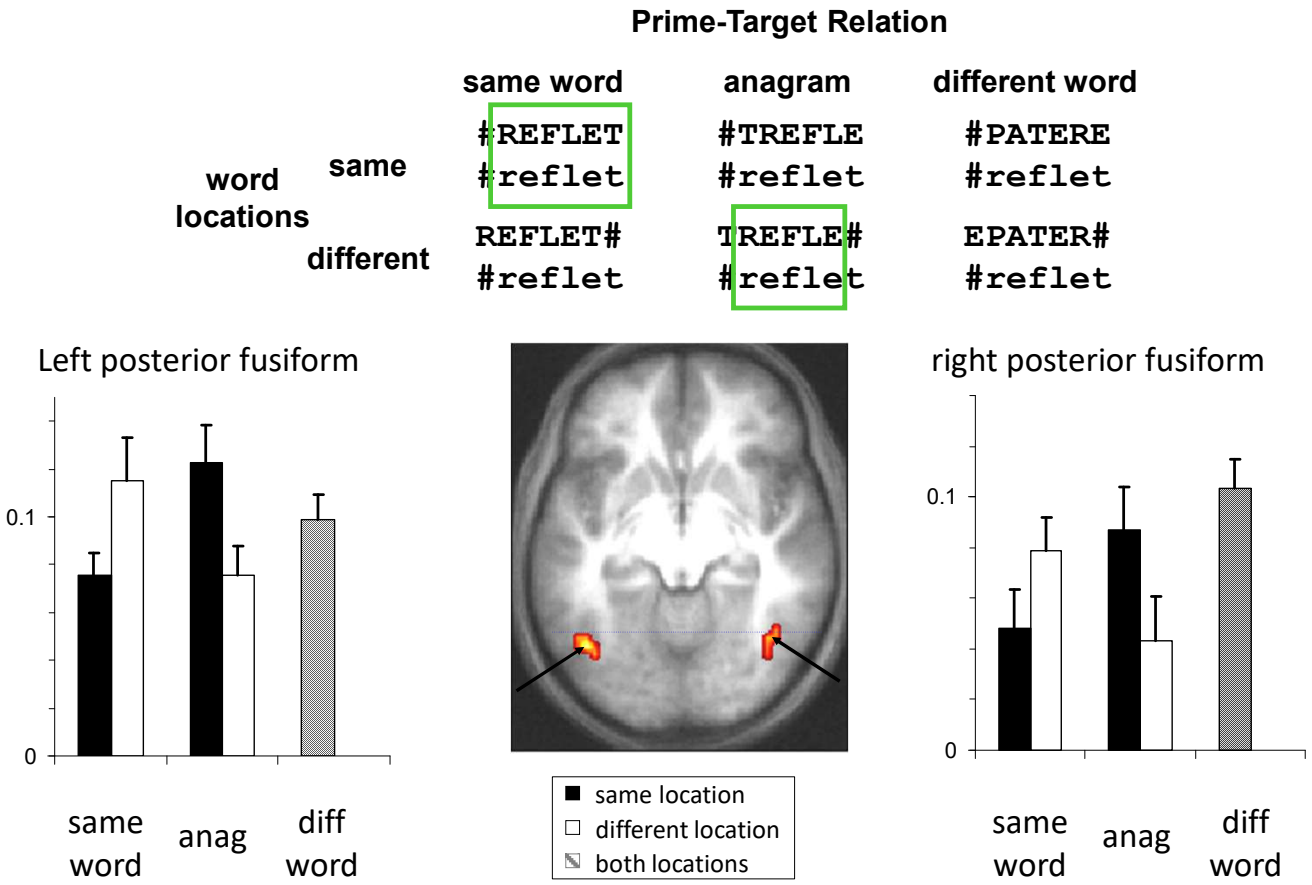
Task = bisyllabic word or not ?

Behavioral priming arises only from the whole-word level.

No priming is observed when the same letters, presented at the same location, form a different word.

Retinotopic letter priming in the bilateral posterior VWFA

Dehaene et al., *Psychological Science,* 2004



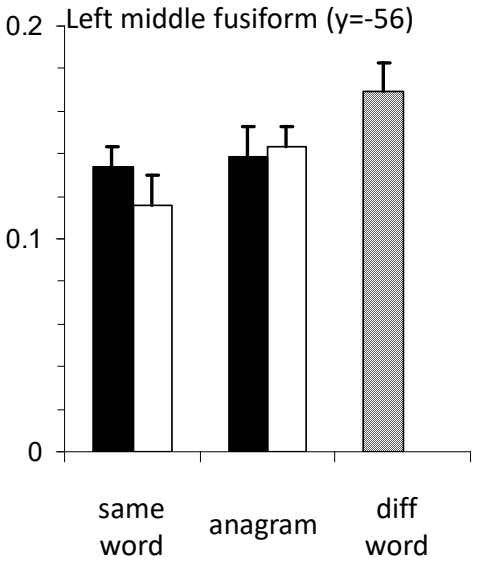
A neural code for individual letters

An ordinal code in the anterior VWFA

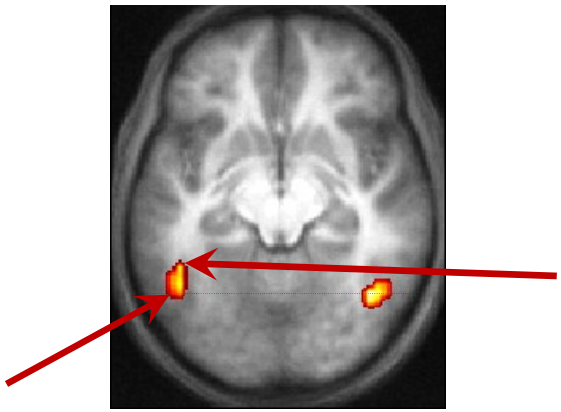
Dehaene et al., *Psychological Science,* 2004

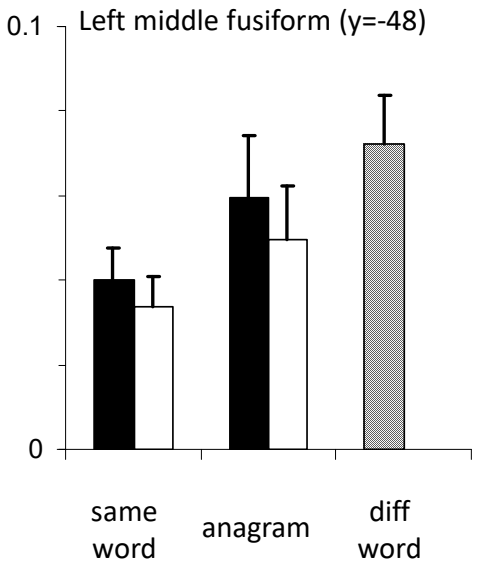
Prime-Target Relation

		same word	anagram	different word
word locations	same	#REFLET #reflet	#TREFLE #reflet	#PATERE #reflet
	different	REFLET# #reflet	TREFLE# #reflet	EPATER# #reflet



A letter code with positional invariance



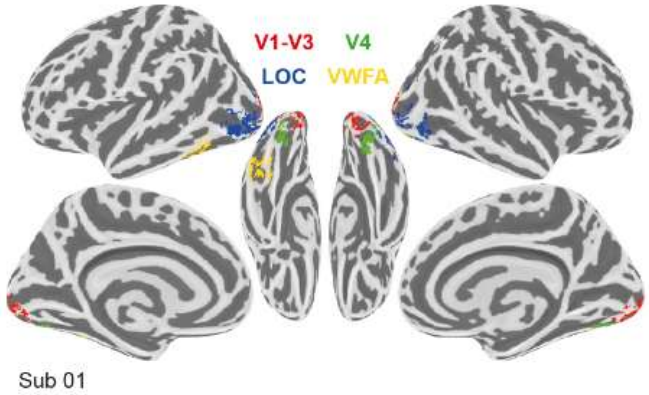


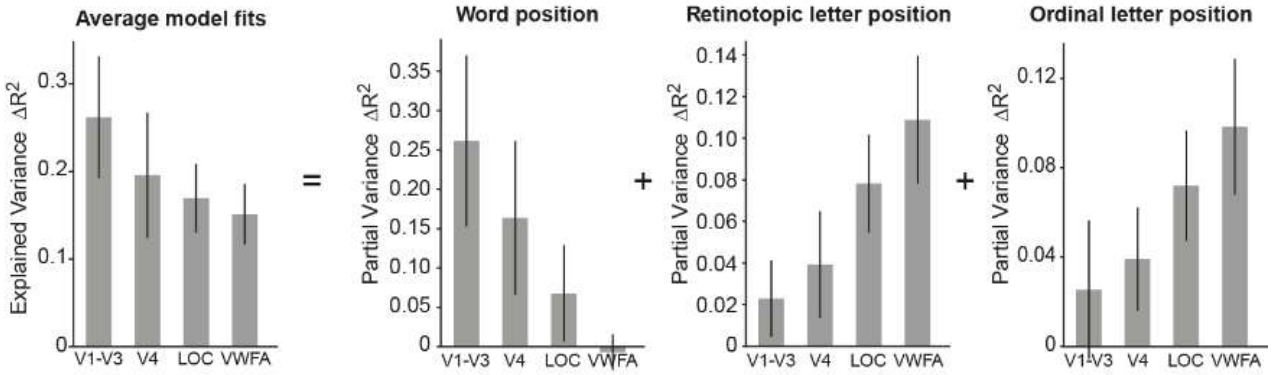
An ordinal letter code

The neural code for written words: an empirical test with 7T fMRI and MEG

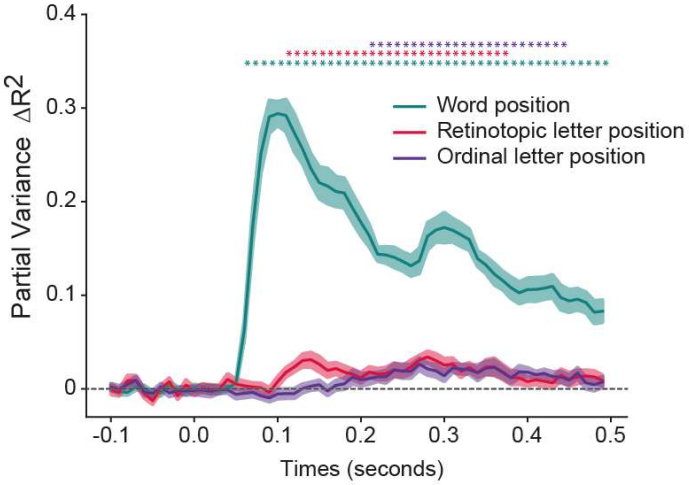
Agrawal & Dehaene, PNAS (2025)

Fitting 7T MRI signals with a forward model with various units

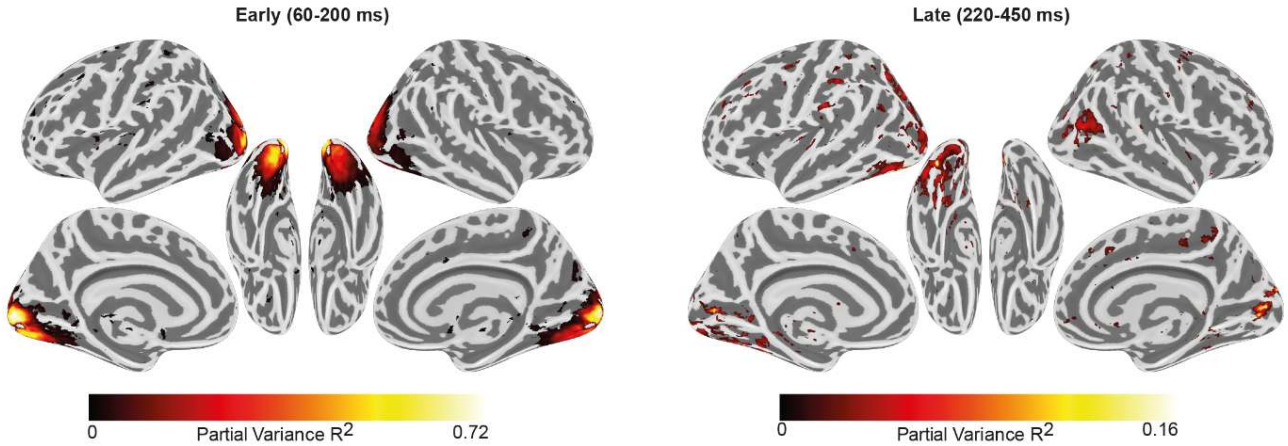




Fitting MEG signals with the same model

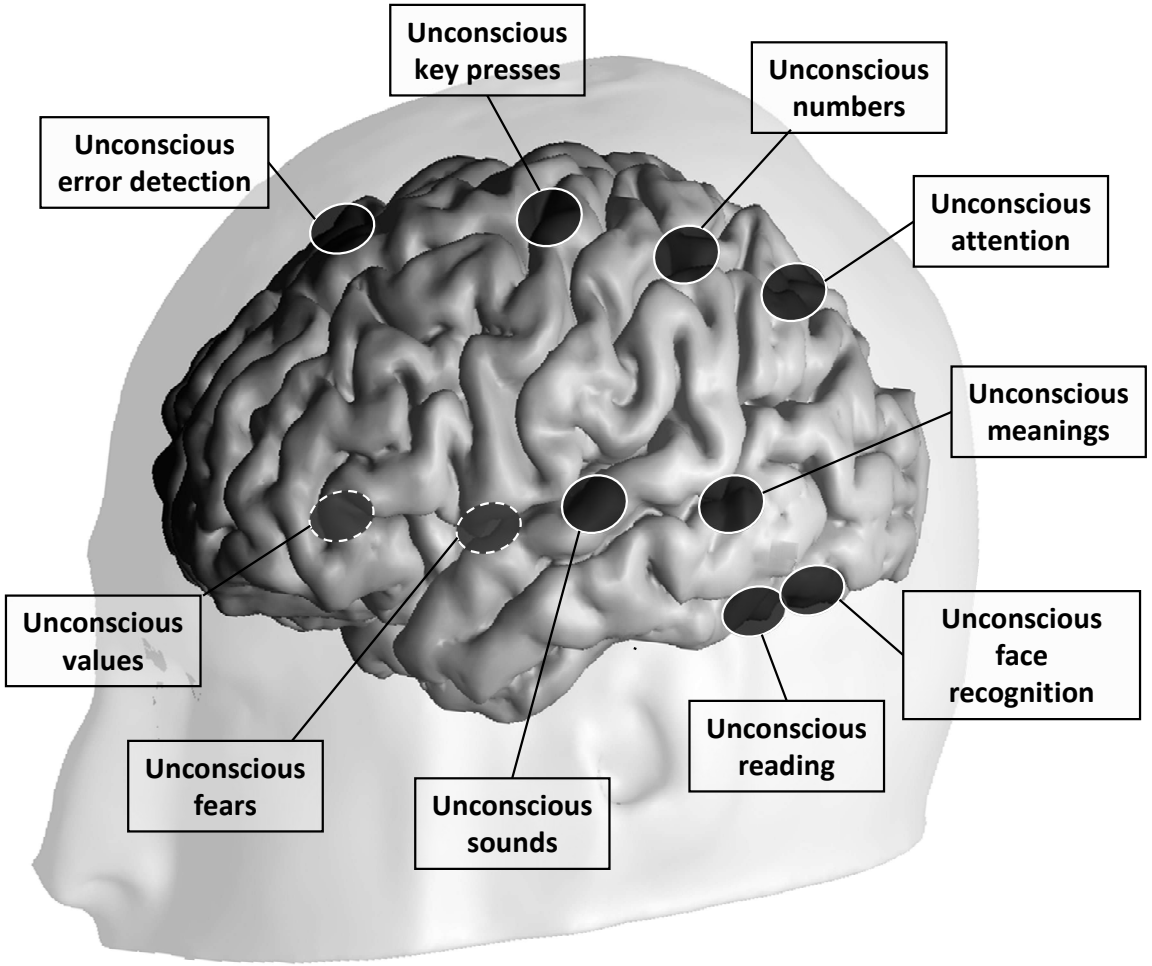


MEG to fMRI : fusion with searchlight representational similarity analysis (RSA)



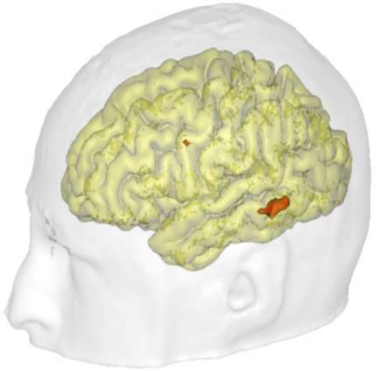
Considerable amounts of nonconscious processing

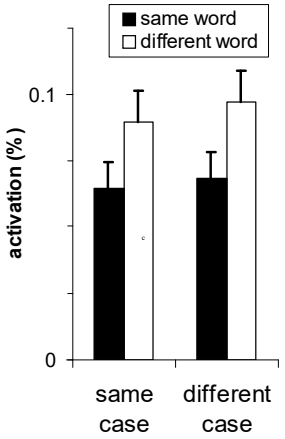
Virtually any brain region can process **some** information nonconsciously.



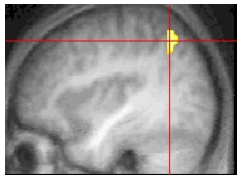
E.g. what happens to a subliminal word?

Orthographic priming



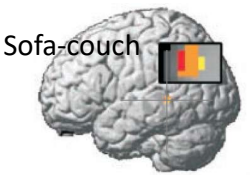


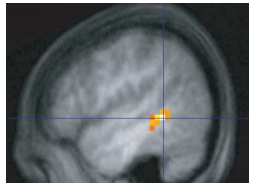
Number words



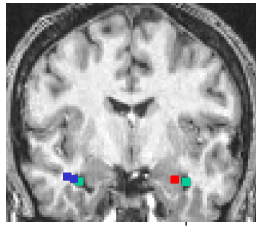
Synonym words

Sofa-couch



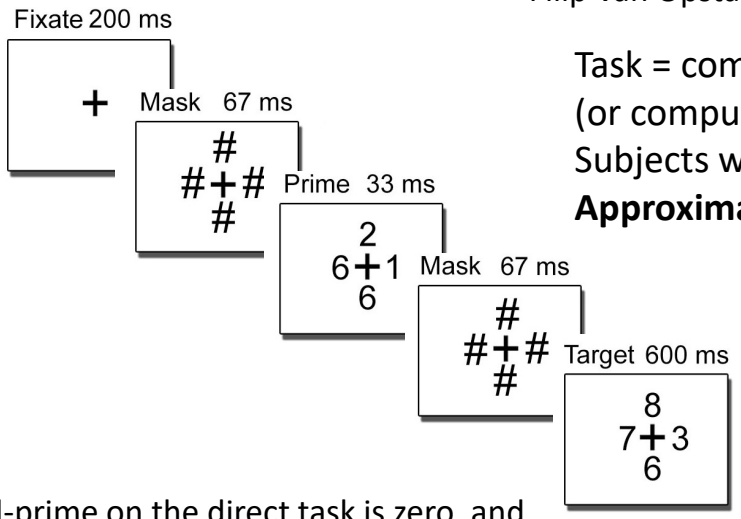


Fearful words



Unconscious processing of several simultaneous symbols ?

Filip Van Opstal, Floris de Lange, Stanislas Dehaene, *Cognition*, 2011



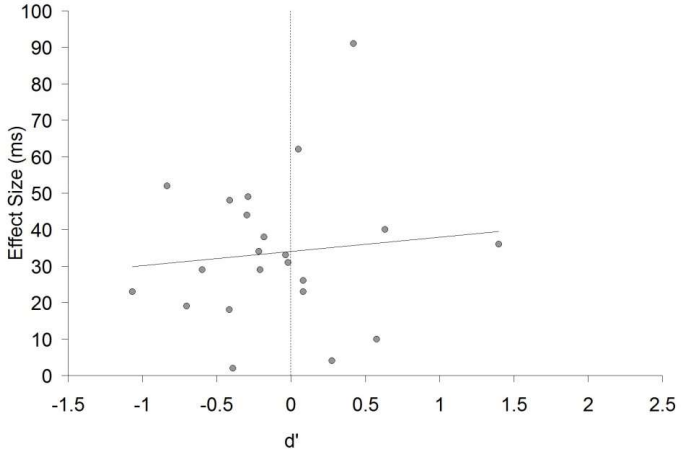
Task = compute the sum of **four target digits** and decide if it is larger or smaller than 20 (or computer the mean and compare it to 5)

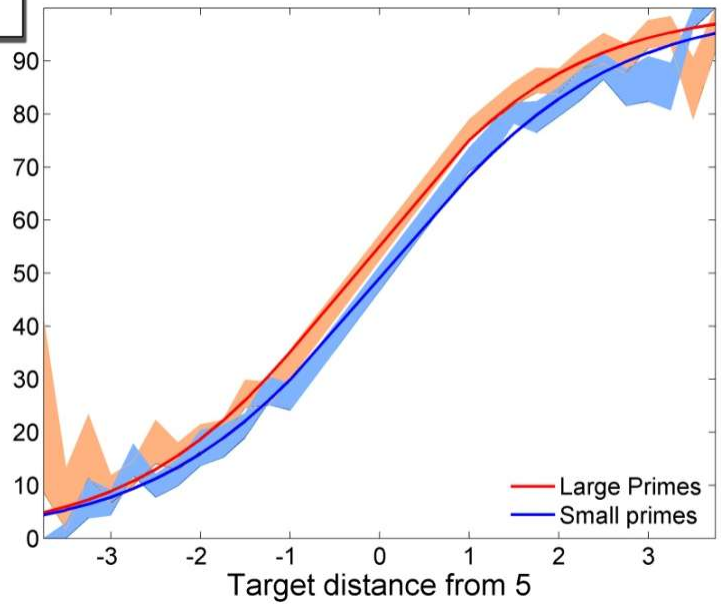
Subjects were influenced by all eight numbers, including the **four subliminal primes**

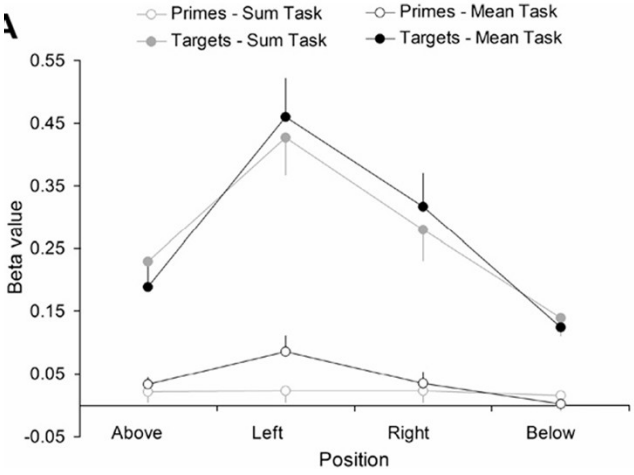
Approximate arithmetic is possible without consciousness

$$y = 1/(1+e^{-z}),$$
$$with\ z = (\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots \beta_8 x_8)$$

d-prime on the direct task is zero, and uncorrelated with the priming effect.







Sentence processing : Can multiple words be combined nonconsciously?

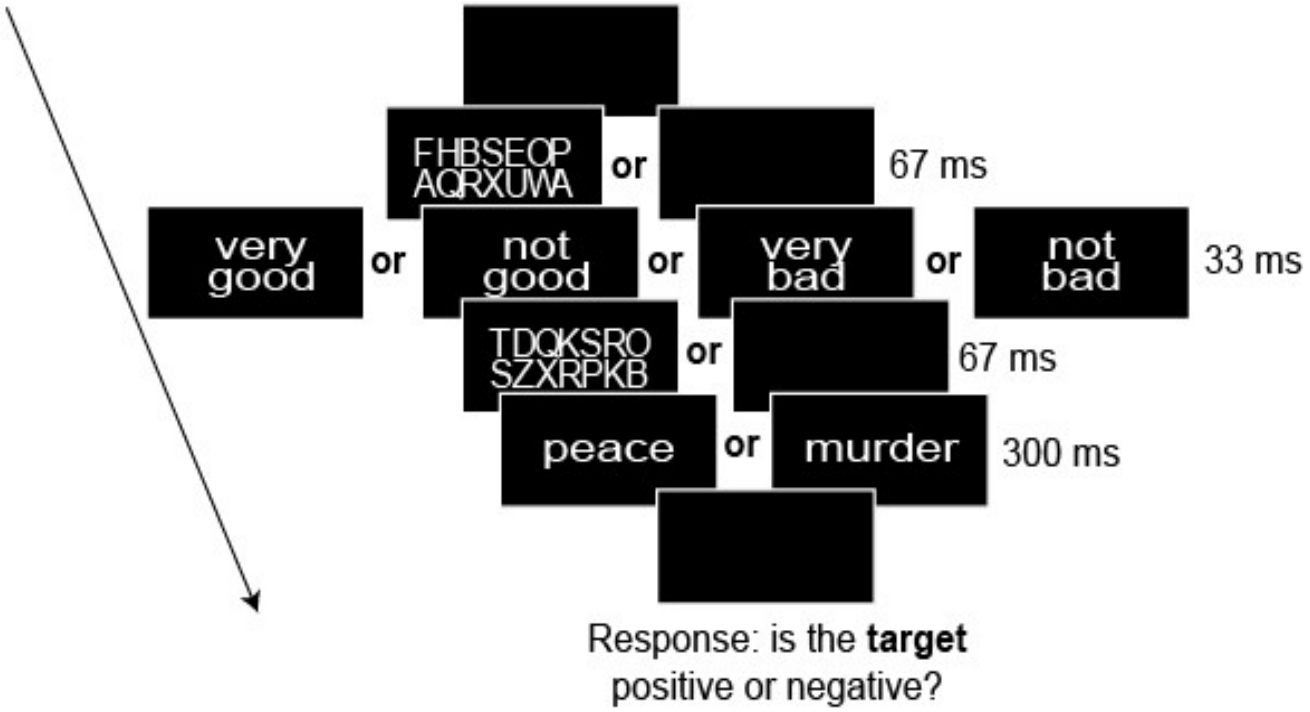
Simon Van Gaal et al., *Proc Royal Society 2014*

Question : can non-conscious processing go beyond a mere **addition** of evidence ?

Can the meaning of **combinations of words** be computed non-consciously ?

« Happy war » → N400 evoked by (nonconscious) semantic incongruity

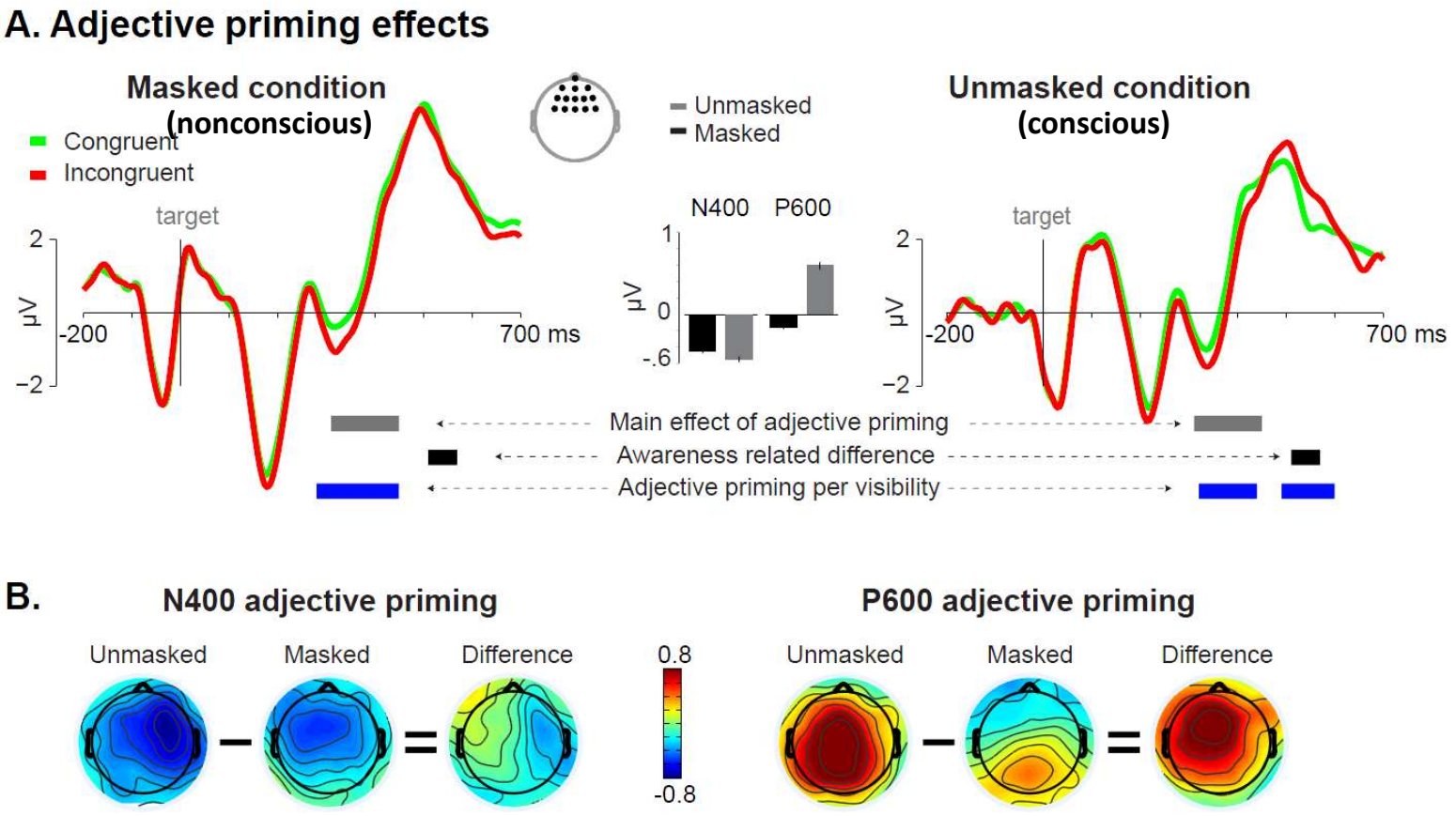
« Very happy war » versus « Not happy war » → modulation of N400?



The N400 is modulated by the adjective…

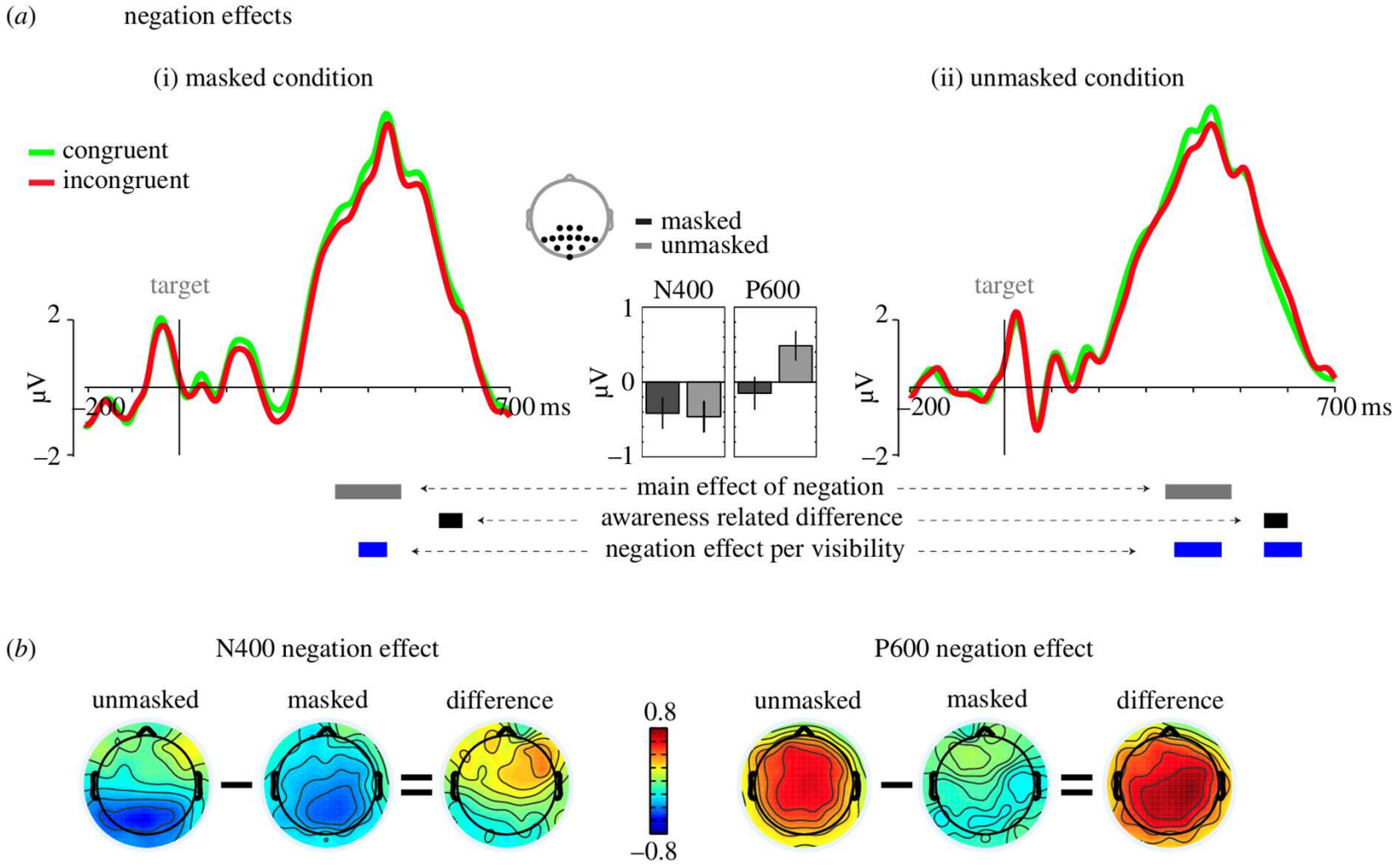
Simon Van Gaal et al., *Proc Royal Society 2014*

Neural correlates of the adjective priming effect: nonconscious N400, conscious P600



The N400 is also modulated by the modifier!

Simon Van Gaal et al., *Proc Royal Society 2014*

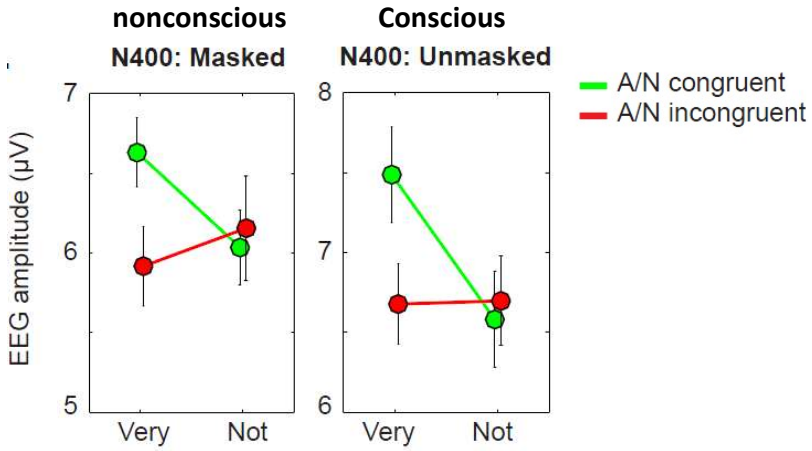


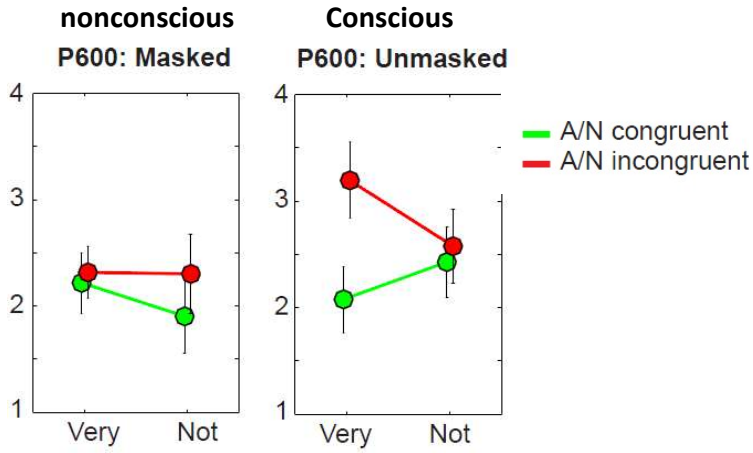
Nonconscious N400, conscious P600

Simon Van Gaal et al., *Proc Royal Society 2014*

The N400 reflects a purely nonconscious stage of word processing:

It is not even *affected* by whether the word is or is not conscious.





The P600 reflects a purely conscious stage of processing:
It is strongly modulated by a conscious modifier, but not at all by a nonconscious one.

Conclusion :

The meaning of three words forming a noun phrase can be nonconsciously integrated.

An XOR operation can be computed nonconsciously within the semantic system.

Can syntactic unification between two consecutive words occur nonconsciously?

Berkovitch and Dehaene, Cognitive Psychology, 2019

Lucie Berkovitch

We all possess a syntactic lexicon that specifies:

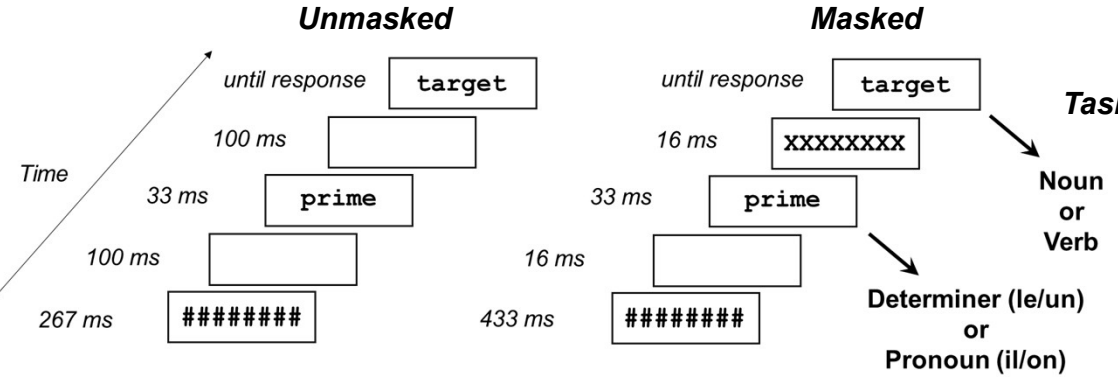
- The part of speech to which a word belongs (e.g. noun)
- Other syntactic features (e.g. syntactic number and gender)

eats =
eat+s
+verb +present + singular

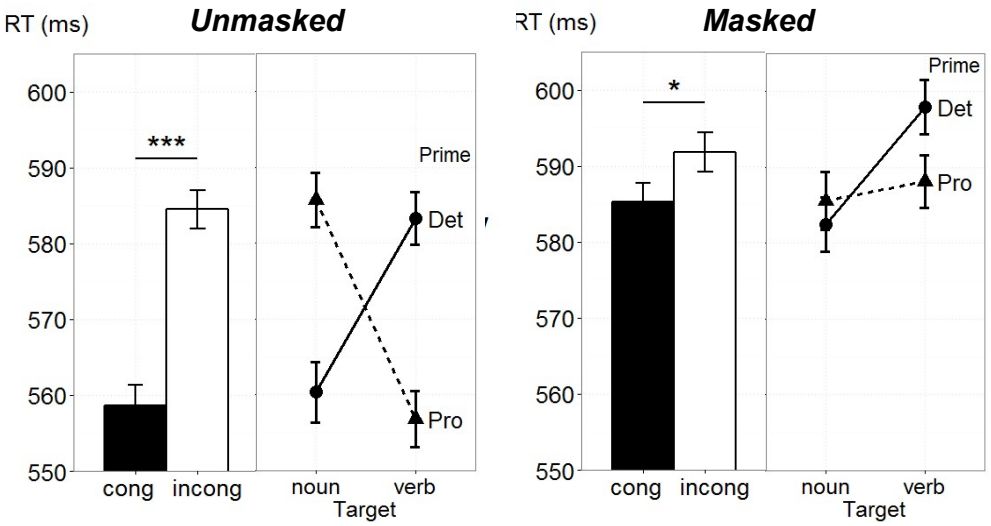
rats =
rat+s
+noun +plural

Can this information be accessed nonconsciously?

Do subjects develop nonconscious expectations of a syntactic category, when presented either with a determiner (expect a noun) or with a pronoun (expect a verb)?



Importantly, different determiners and pronouns were used for nonconscious and for conscious presentation



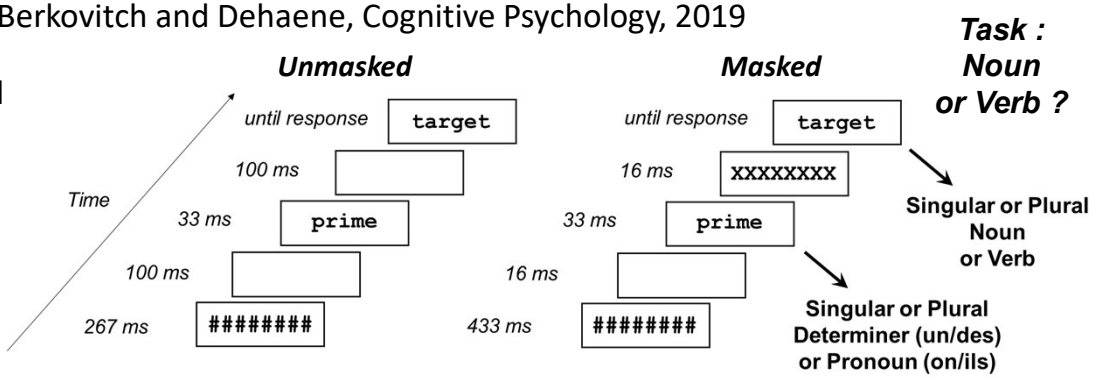
The results show a very strong syntactic congruity effect, for both masked and unmasked primes.

Is the priming effect sensitive to agreement in number?

Berkovitch and Dehaene, Cognitive Psychology, 2019

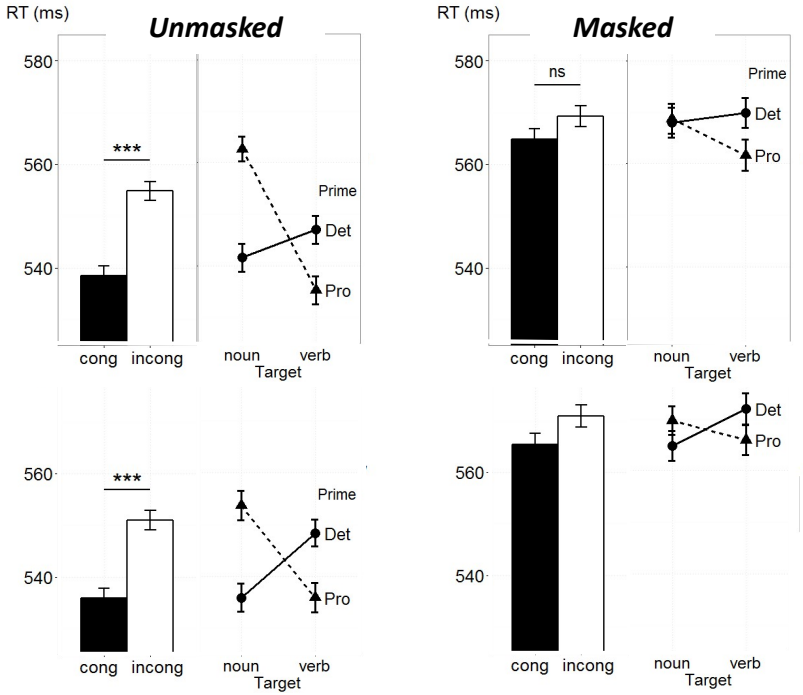
The previous experiment could be explained by a violation of transition probability.

Would Det-Noun and Pronoun-Verb effects occur even with a grammatical violation of number?



Syntactic priming when number is congruent
Ex. ils mangent (they eat) vs ils chiens (they dogs)
un chien (a dog) vs un mange (a eats)

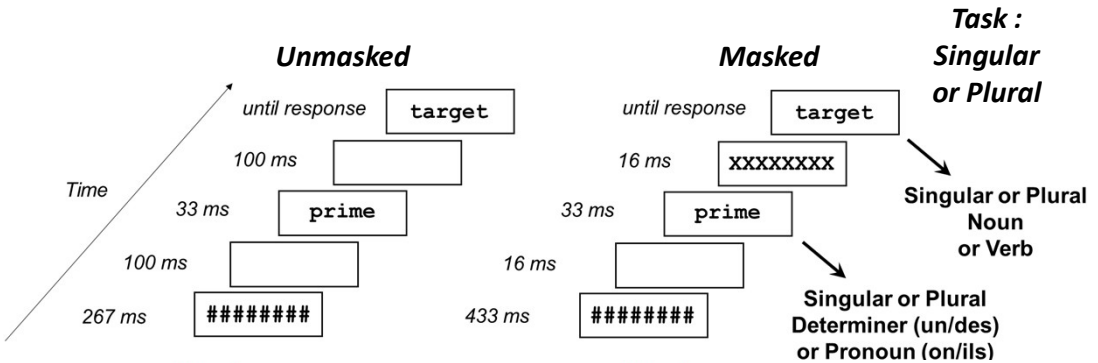
Syntactic priming when number is incongruent
Ex. ils mange (they eats) vs ils chien (they dog)
un chiens (a dogs) vs un mangent (a eat)

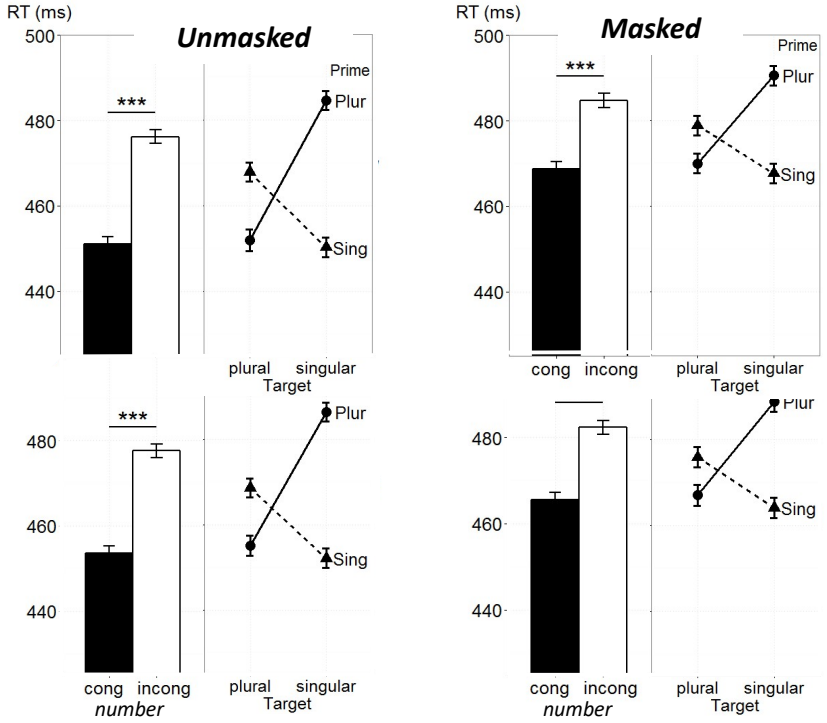


Is the priming effect sensitive to agreement in number?

In the previous experiment, priming occured between syntactic categories, regardless of grammatical number.

Could we reverse the effect and get priming according to grammatical number, regardless of category?





Number priming when syntax is congruent

Ex. ils mangent (they eat) vs ils mange (they eats)
un chien (a dog) vs un chiens (a dogs)

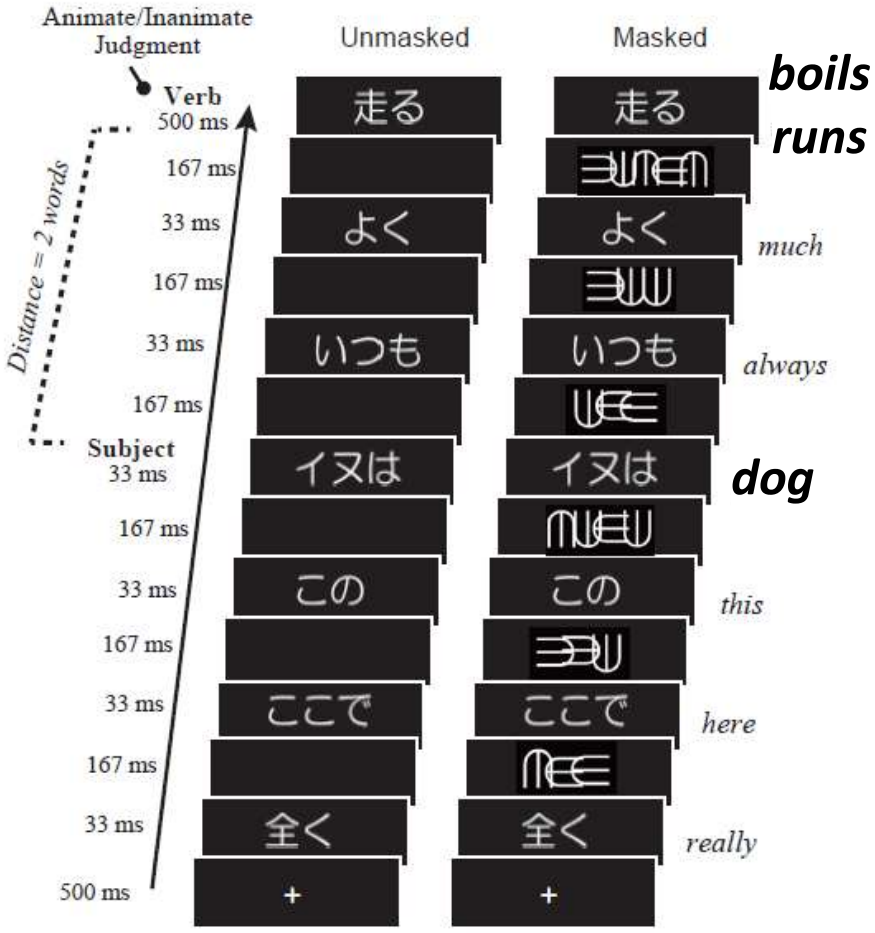
Number priming when syntax is incongruent

Ex. des mangent (the(pl.) eat) vs des mange (the(pl.) eats)
on chien (he dogs) vs on chiens (he dog)

Conclusion: Access to morphosyntactic information and computation of agreement relations between two consecutive words can unfold non-consciously.

Is there a key distance beyond which nonconscious word integration fails ?

Nakamura, K., Makuuchi, M., Oga, T., Mizuochi-Endo, T., Iwabuchi, T., Nakajima, Y., & Dehaene, S. (2018). Neural capacity limits during unconscious semantic processing. *The European Journal of Neuroscience*, *47*(8), 929–937. https://doi.org/10.1111/ejn.13890



An entire sentence was presented consciously or non-consciously. Only the first word (“really”) and the last word (a verb) were always conscious.

The idea was to examine whether the sentence could impact on the processing of the last word.

The last word (a verb) could be congruent or incongruent with its preceding subject noun.

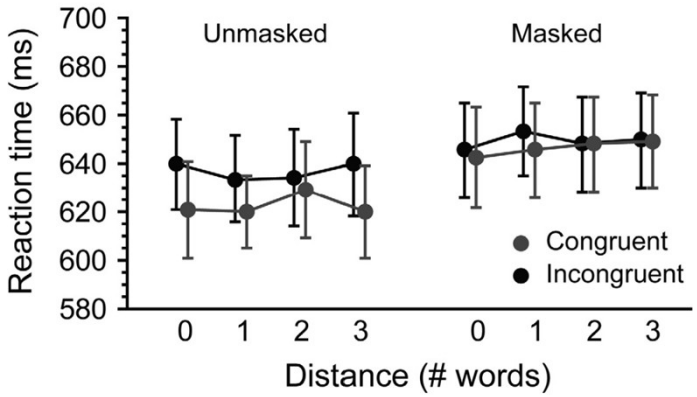
The **distance between the subject and the verb** was varied by inserting additional adverbs.

Participants had to judge whether the last word was inanimate or animate.

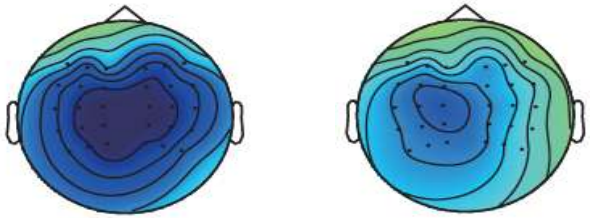
Are there limits to nonconscious word integration?

Nakamura, K., Makuuchi, M., Oga, T., Mizuochi-Endo, T., Iwabuchi, T., Nakajima, Y., & Dehaene, S. (2018). Neural capacity limits during unconscious semantic processing. *The European Journal of Neuroscience*, *47*(8), 929–937. https://doi.org/10.1111/ejn.13890

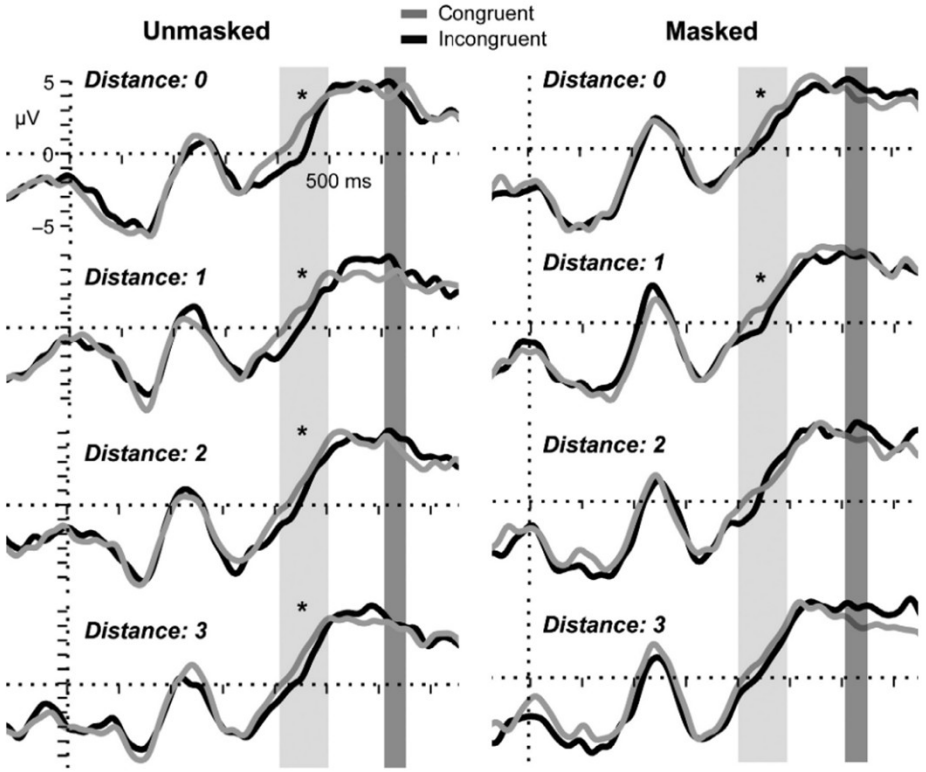
Behaviorally, only the conscious sentence primed the decision, regardless of distance.
For the nonconscious sentence, there was only a trend for priming at distances 0 and 1.



Incongruity generated an N400 wave to the last word



This effect was significant at all distances in the conscious condition, but only at distances 0 and 1 for nonconscious sentences (200 or 400 ms).
Tentative conclusion: consciousness is needed to integrate words together. Nonconscious stimuli decay over time and cannot be integrated.



Probing the limits of nonconscious language processing

Gui, P., Jiang, Y., Zang, D., Qi, Z., Tan, J., Tanigawa, H., Jiang, J., Wen, Y., Xu, L., Zhao, J., Mao, Y., Poo, M., Ding, N., Dehaene, S., Wu, X., & Wang, L. (2020). Assessing the depth of language processing in patients with disorders of consciousness. *Nature Neuroscience*, *23*(6), Article 6.

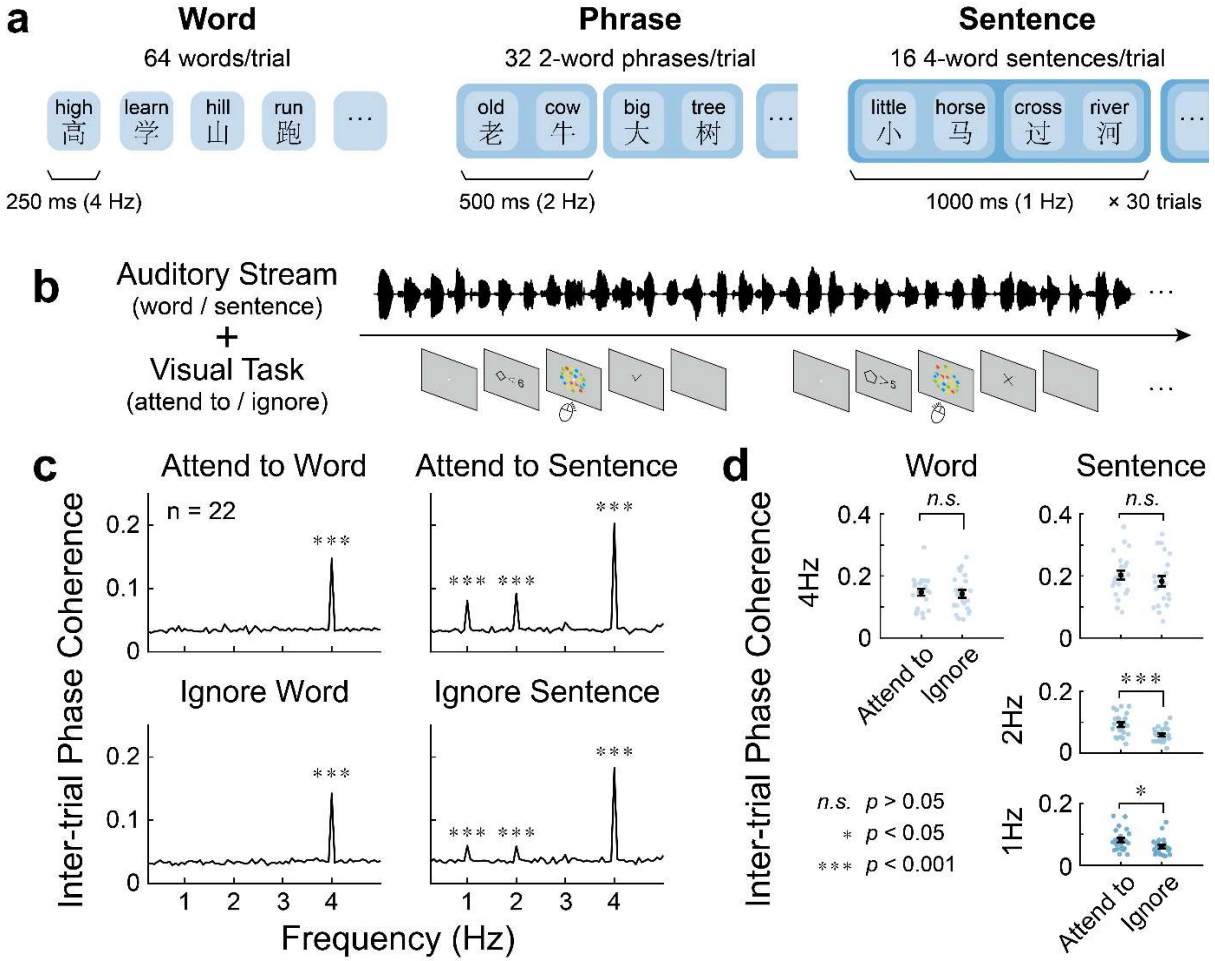
Goal = evaluate the residual processing of speech in patients with disorders of consciousness (coma, vegetative state).

Paradigm similar to Ding et al. 2015 : using EEG frequency tagging to probe the brain's processing of words, phrases and sentences. Measure = inter-trial phase coherence (ITPC).

The first experiment is in normal participants, with a strong distraction task:

- The word, phrase and sentence components induce detectable peaks in the EEG spectrum, at the appropriate frequencies.
- Inattention reduces their amplitude, but they remain detectable.

This experiment is a bit ambiguous with respect to consciousness.



Probing the limits of nonconscious language processing

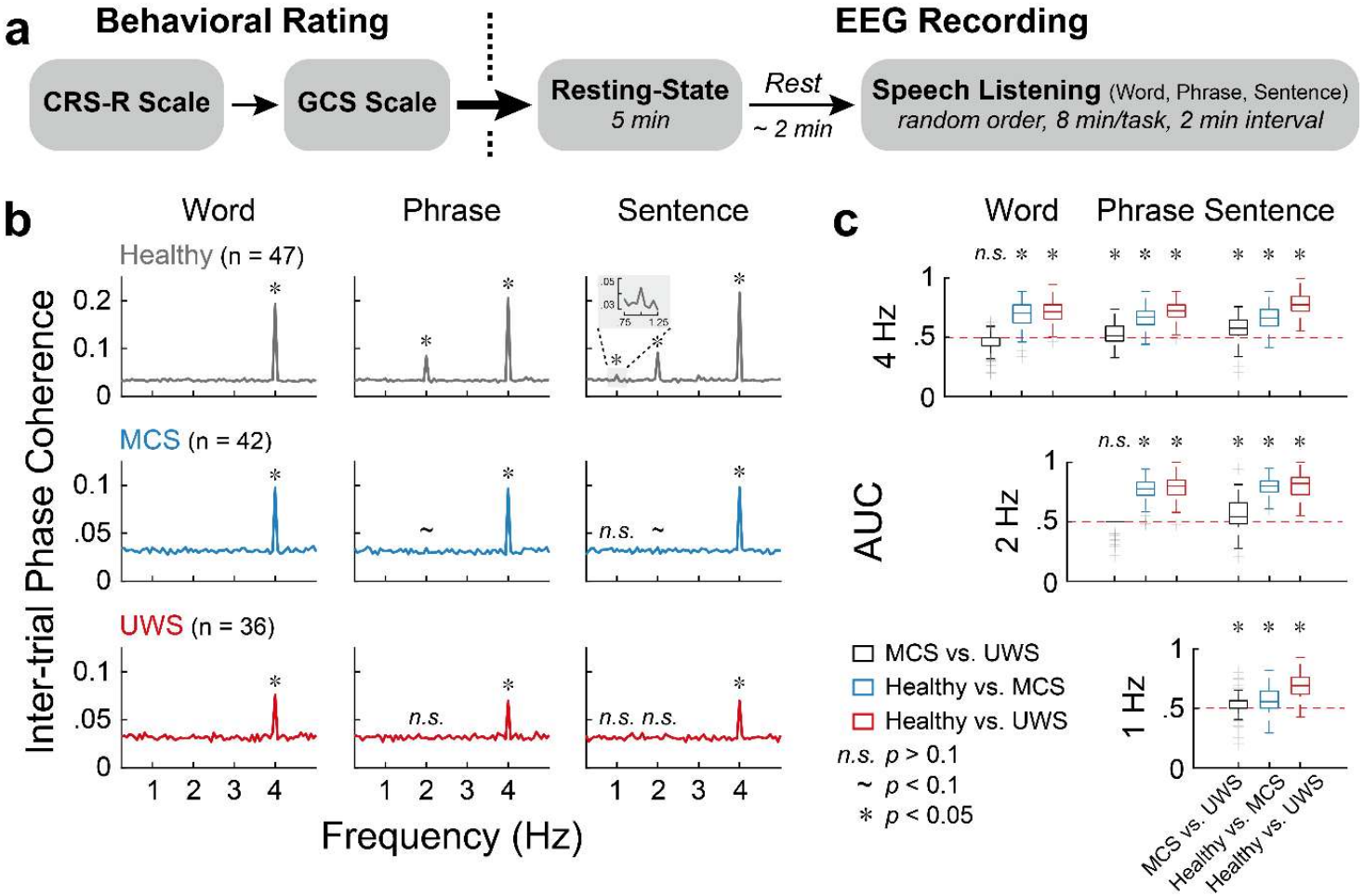
Gui, P., Jiang, Y., Zang, D., Qi, Z., Tan, J., Tanigawa, H., Jiang, J., Wen, Y., Xu, L., Zhao, J., Mao, Y., Poo, M., Ding, N., Dehaene, S., Wu, X., & Wang, L. (2020). Assessing the depth of language processing in patients with disorders of consciousness. *Nature Neuroscience*, *23*(6), Article 6.

The same paradigm was applied to patients with disorders of consciousness (DoC : minimal conscious state = MCS ; unresponsive wakefulness syndrome = UWS)

MCS, but not UWS, still exhibit a small phrase effect.
Both groups lose the sentence effect.
However, individual subjects (11/42 MCS, 4/36 UWS) still showed a residual higher-order ITPC – and they shower a greater recovery.

Conclusions:

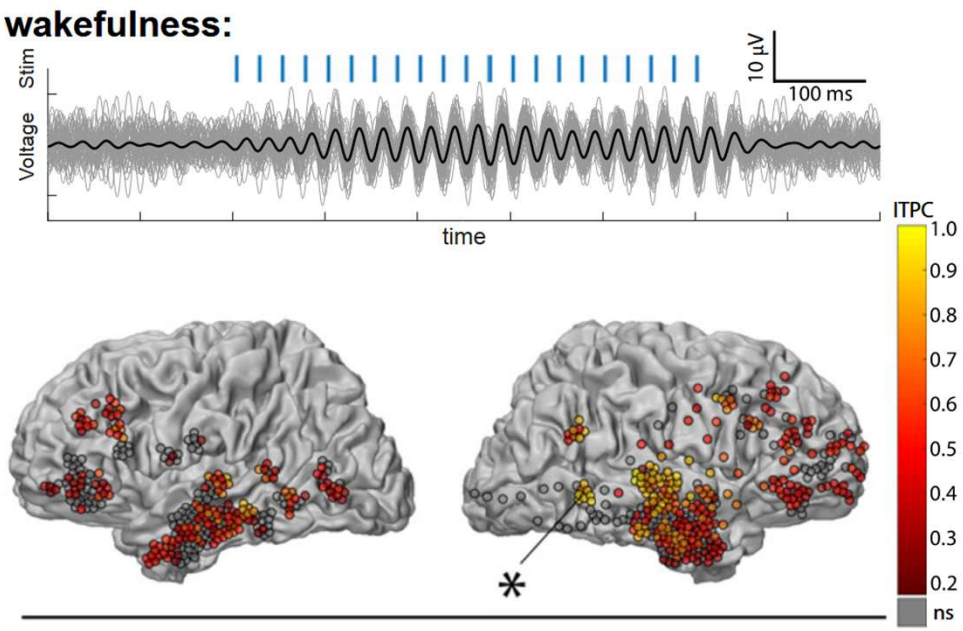
- multi-word integration is severely impaired in DoC patients
- Its preservation can be a useful clinical sign of preserved consciousness and likelihood of recovery.

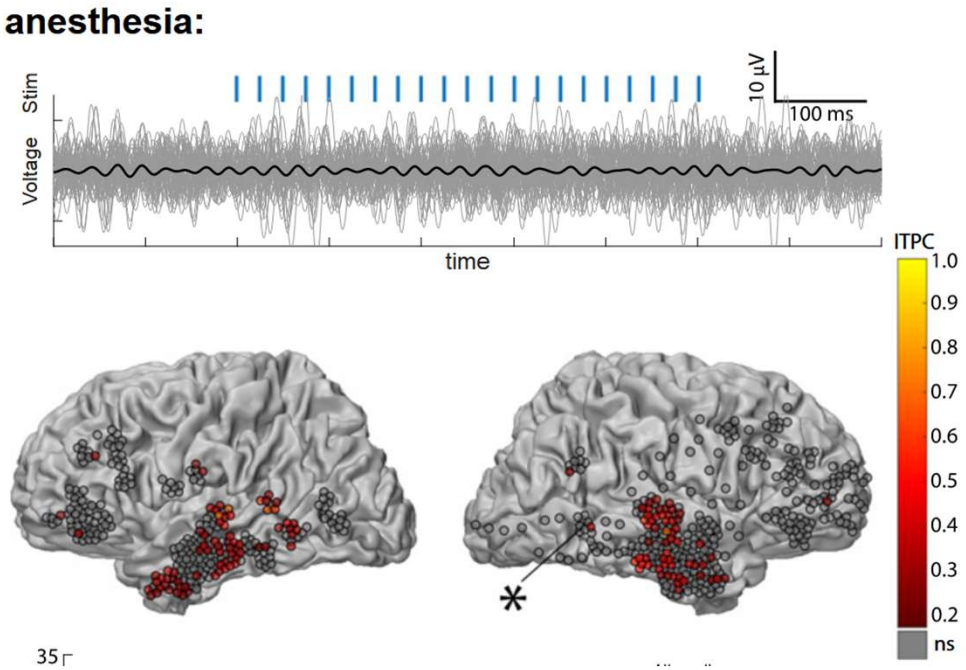


Effect of anesthesia on sound processing

Krom, A. J., Marmelshtein, A., Gelbard-Sagiv, H., Tankus, A., Hayat, H., Hayat, D., Matot, I., Strauss, I., Fahoum, F., Soehle, M., Boström, J., Mormann, F., Fried, I., & Nir, Y. (2020). Anesthesia-induced loss of consciousness disrupts auditory responses beyond primary cortex. Proceedings of the National Academy of Sciences, 117(21), 11770-11780. https://doi.org/10.1073/pnas.1917251117

Intracranial recordings during the presentation of sound sequences, while propofol anesthesia is increased until the patient stops responding.
In iEEG, primary auditory cortex continues to respond, as well as some temporal lobe regions, but all higher order prefrontal and temporal regions seem to disappear.

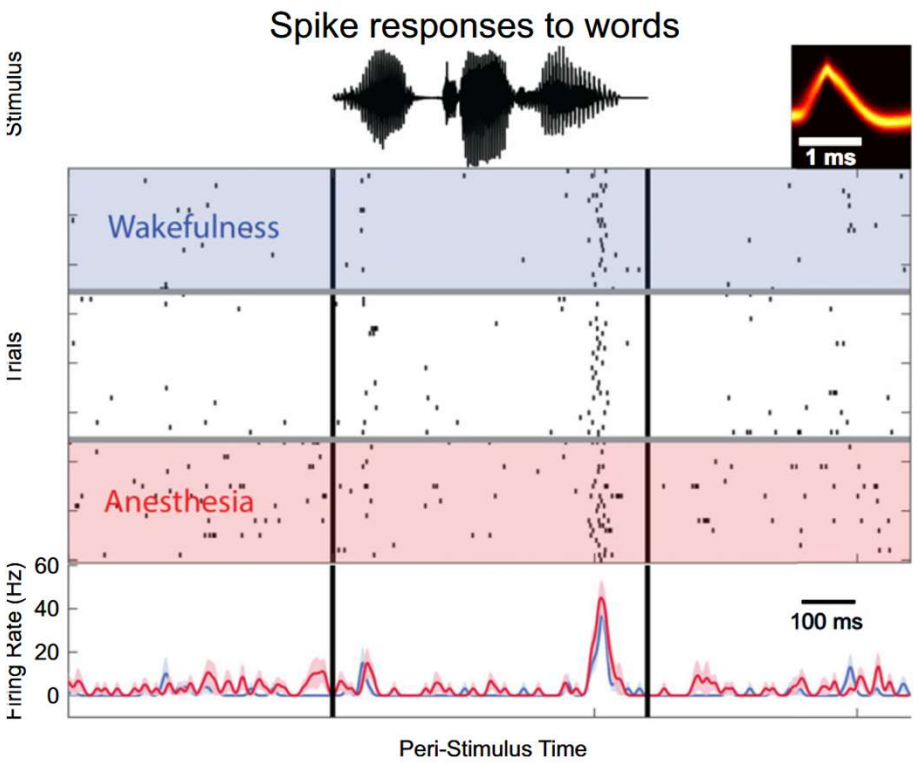




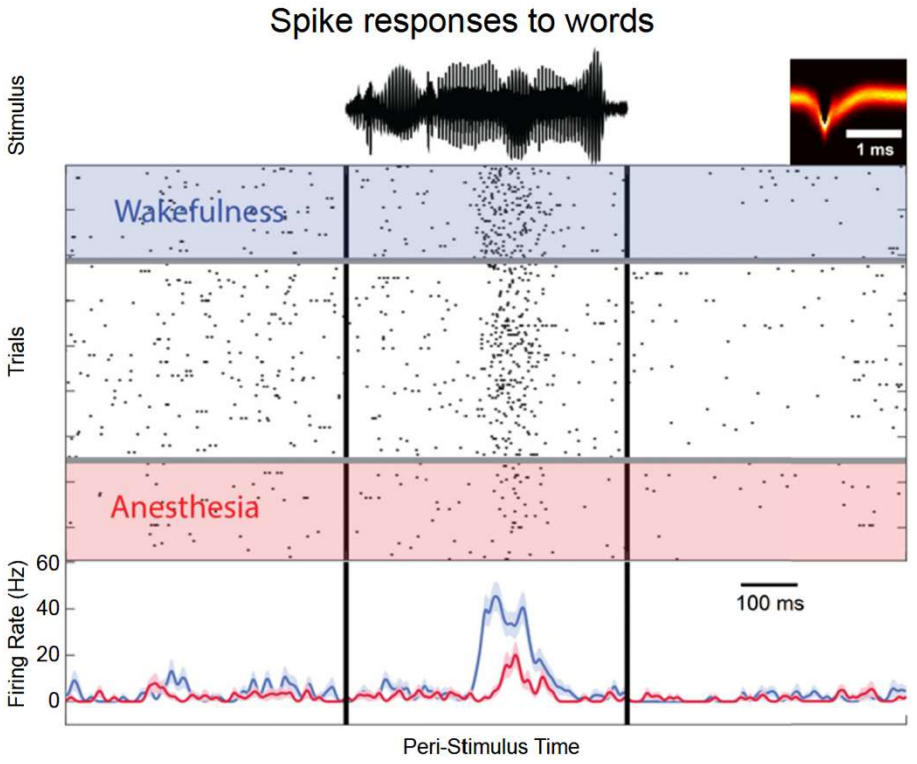
Effect of anesthesia on sound processing

Krom, A. J., Marmelshtein, A., Gelbard-Sagiv, H., Tankus, A., Hayat, H., Hayat, D., Matot, I., Strauss, I., Fahoum, F., Soehle, M., Boström, J., Mormann, F., Fried, I., & Nir, Y. (2020). Anesthesia-induced loss of consciousness disrupts auditory responses beyond primary cortex. Proceedings of the National Academy of Sciences, 117(21), 11770-11780. https://doi.org/10.1073/pnas.1917251117

Single-neuron recordings: Spiking continues, sometimes completing unchanged, in primary auditory cortex.



In higher-order auditory cortex (still in Heschl's gyrus), cell firing becomes much smaller and more variable.

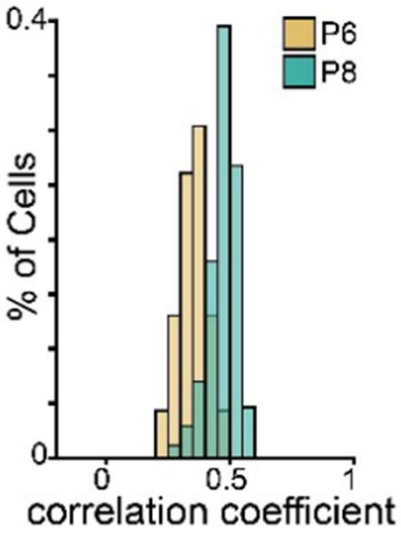


Semantic processing in the hippocampus under anesthesia

Katlowitz, K. A., Shah, S., Franch, M. C., Adkinson, J., Belanger, J. L., Mathura, R. K., Meszéna, D., Mickiewicz, E. A., McGinley, M., Muñoz, W., Banks, G. P., Cash, S. S., Hsu, C.-W., Paulk, A. C., Provenza, N. R., Watrous, A., Williams, Z., Heilbronner, S. R., Kim, R., … Sheth, S. A. (2025). Learning and language in the unconscious human hippocampus (p. 2025.04.09.648012). bioRxiv. https://doi.org/10.1101/2025.04.09.648012

Recordings from the exposed hippocampus in the middle of surgery for anterior temporal resection (before its ultimate removal). Neuropixel probes record from hundreds of cells.

Experiment 2 : the patient, still anesthesized, listens to short stories. Words are encoded by 6 main dimensions of Word2Vec.



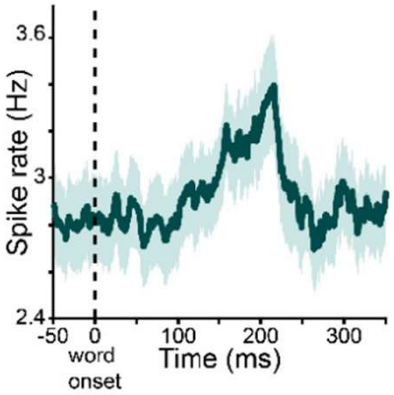
Results:

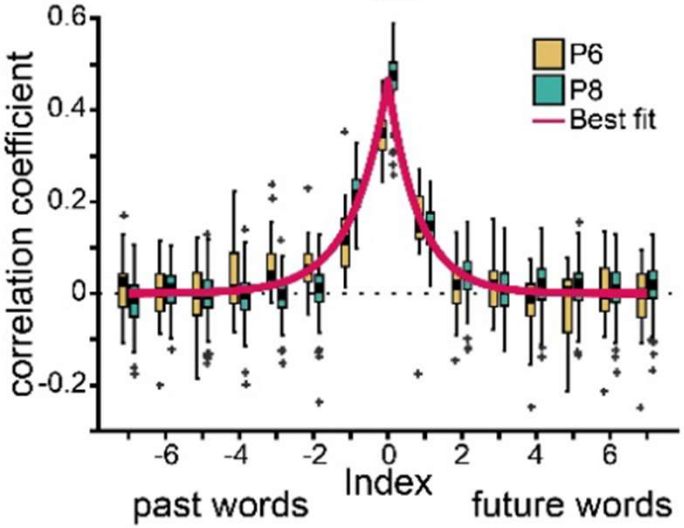
- Very clear activation to words
- The activation of many cells can be predicted from the semantic dimensions of the words
- The features of the previous word and even the next word contribute to predict firing.

The authors interpret this finding as predictive coding, but this is uncertain because words inside a sentence are highly intercorrelated.

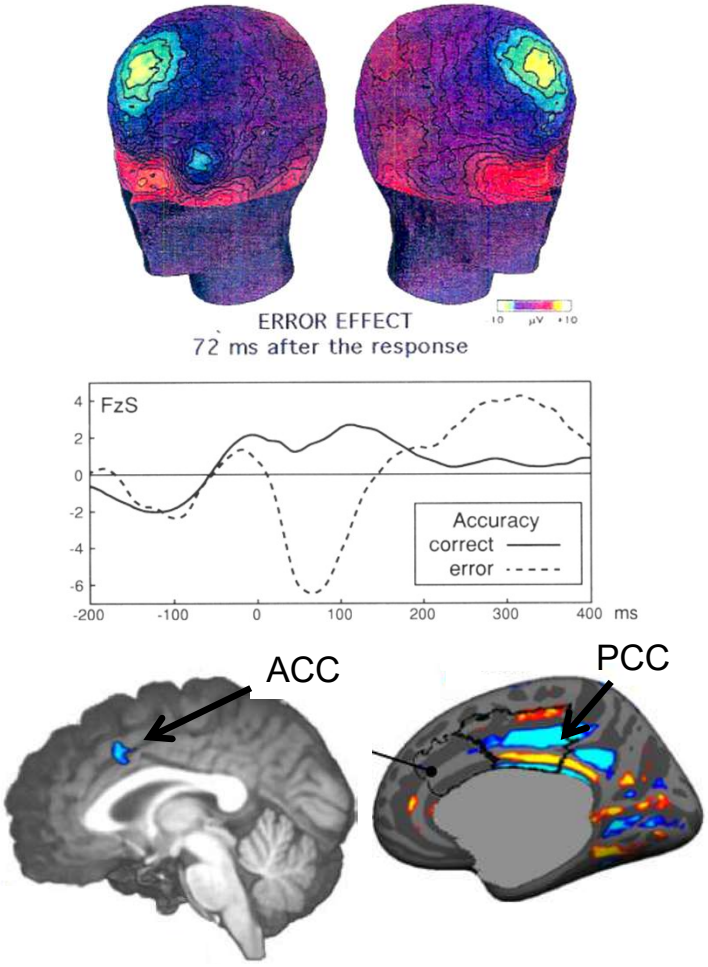
Conclusions :

- Single words can be nonconsciously processed along the auditory pathway, up to the semantic level, and into the hippocampus.
- Some aspects of phrasal integration may occur non-consciously, but perhaps not sentence-level integration.





What are the limits of nonconscious processing ?
Is consciousness required for some computations ?



Debener 2005, Agam 2011

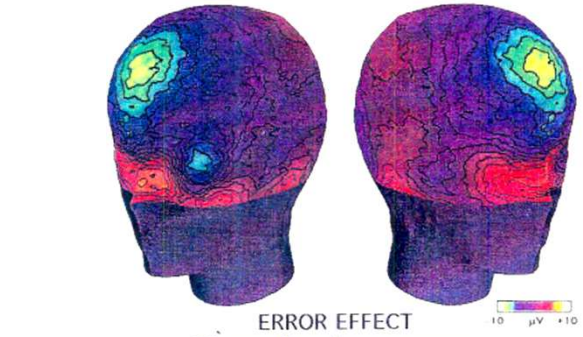
Consciousness acts as **a slow bottleneck**, arising late after the stimulus (~300 ms) and only capable of processing one mental object at a time.

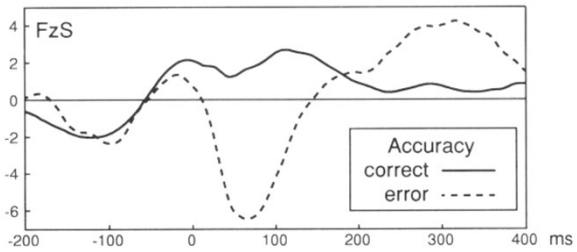
What might its functions be?

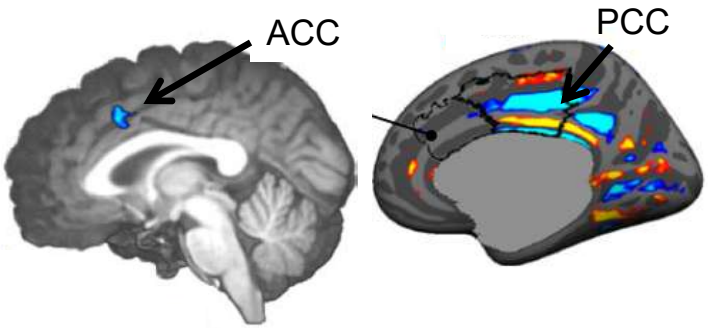
- To **integrate** multiple pieces of information.
- To **reflect upon** the information: subliminal information is evanescent, while conscious information is stable
- To **route the information** to other processing stages, allowing us to perform chains of operations (**a human Turing machine**)
- To **monitor** our behavior and diagnose our errors

Example: the **Error-Related Negativity** (ERN) which arises ~80 ms after an erroneous response (prior to any feedback)**.**

Do meta-cognition and error detection depend on conscious access?







Debener 2005, Agam 2011

Meta-cognition (the ability to entertain thoughts about your own thoughts) seems to require consciousness, almost by definition (e.g. Rosenthal's HOT theory).

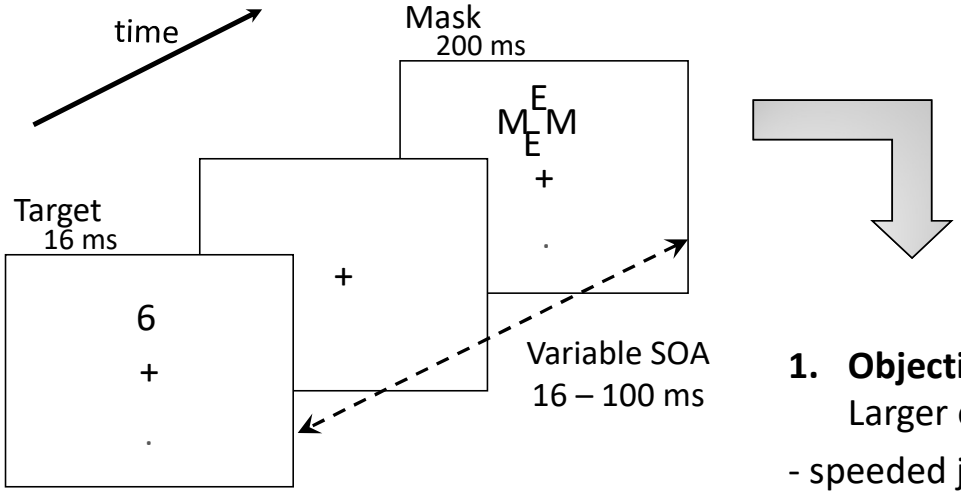
Yet according to Bayesian or signal detection theories, whenever we have evidence, we can also have (some) confidence. **Could confidence be computed non-consciously**?

• Hypothesis 1 : Error detection belongs to prefrontal operations of **executive attention**, and may involve a **comparison** between the consciously intended response and the actual response.
In this case it would require conscious processing of the stimulus.

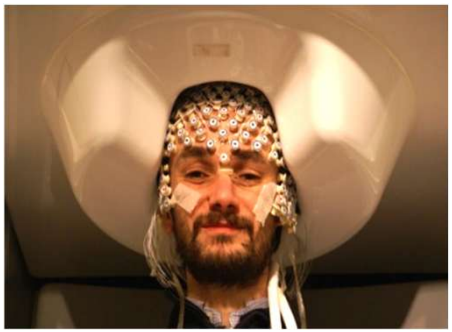
• Hypothesis 2: Alternatively, it could reflect an automatic and non-conscious computation of **response conflict.**

A paradigm to study non-conscious meta-cognition and error detection

Charles, L., Van Opstal, F., Marti, S., & Dehaene, S. (2013). Distinct brain mechanisms for conscious versus subliminal error detection. *Neuroimage*, *73*, 80-94. https://doi.org/10.1016/j.neuroimage.2013.01.054



Event-related potentials + Magnetoencephalography



Lucie Charles

1. **Objective Task :**
   Larger or Smaller than 5

- speeded jugement in experiment 1
- self-paced in experiment 2

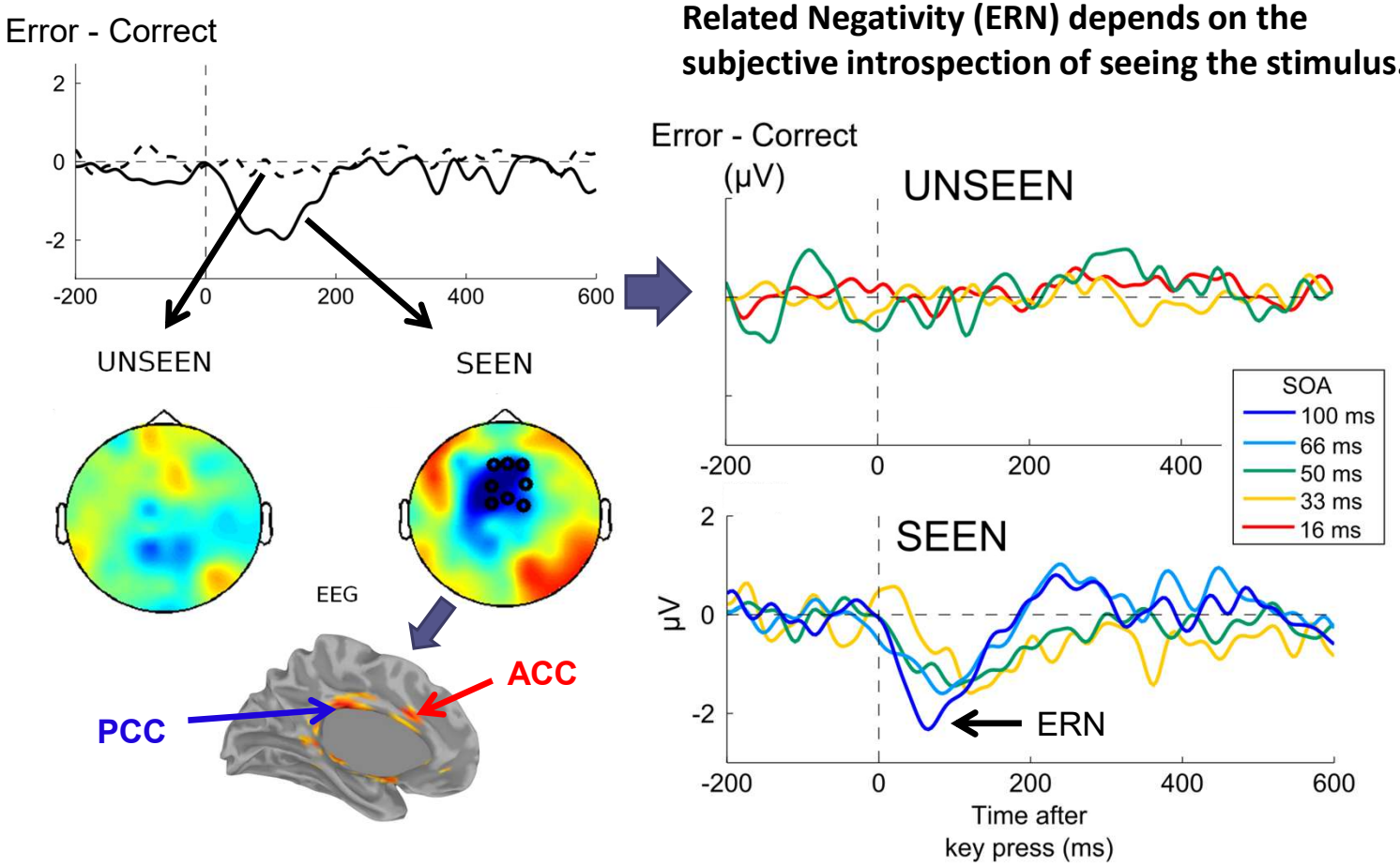
2. **Subjective Visibility Rating :**
   « Seen » or « Not seen »

3. **Subjective Error Detection :**
   « Error » or « Correct »

Error Detection is only deployed on conscious trials

Charles et al., *NeuroImage* 2013

For the same objective stimulus, the Error-Related Negativity (ERN) depends on the subjective introspection of seeing the stimulus.



Above-chance confidence judgments without consciousness

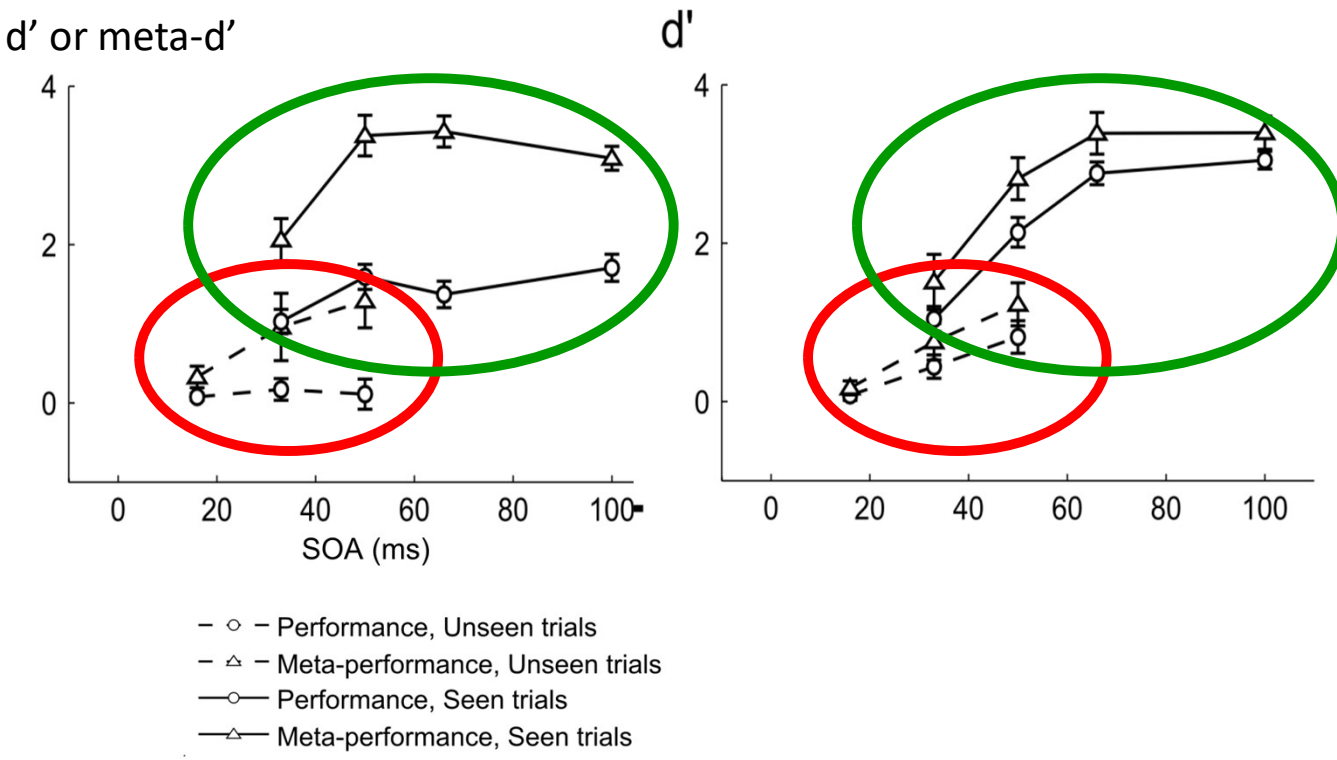
Type 1 signal detection analysis

	target <5	target >5
response left	Hit	Miss
response right	False Alarm	Correct Rejection

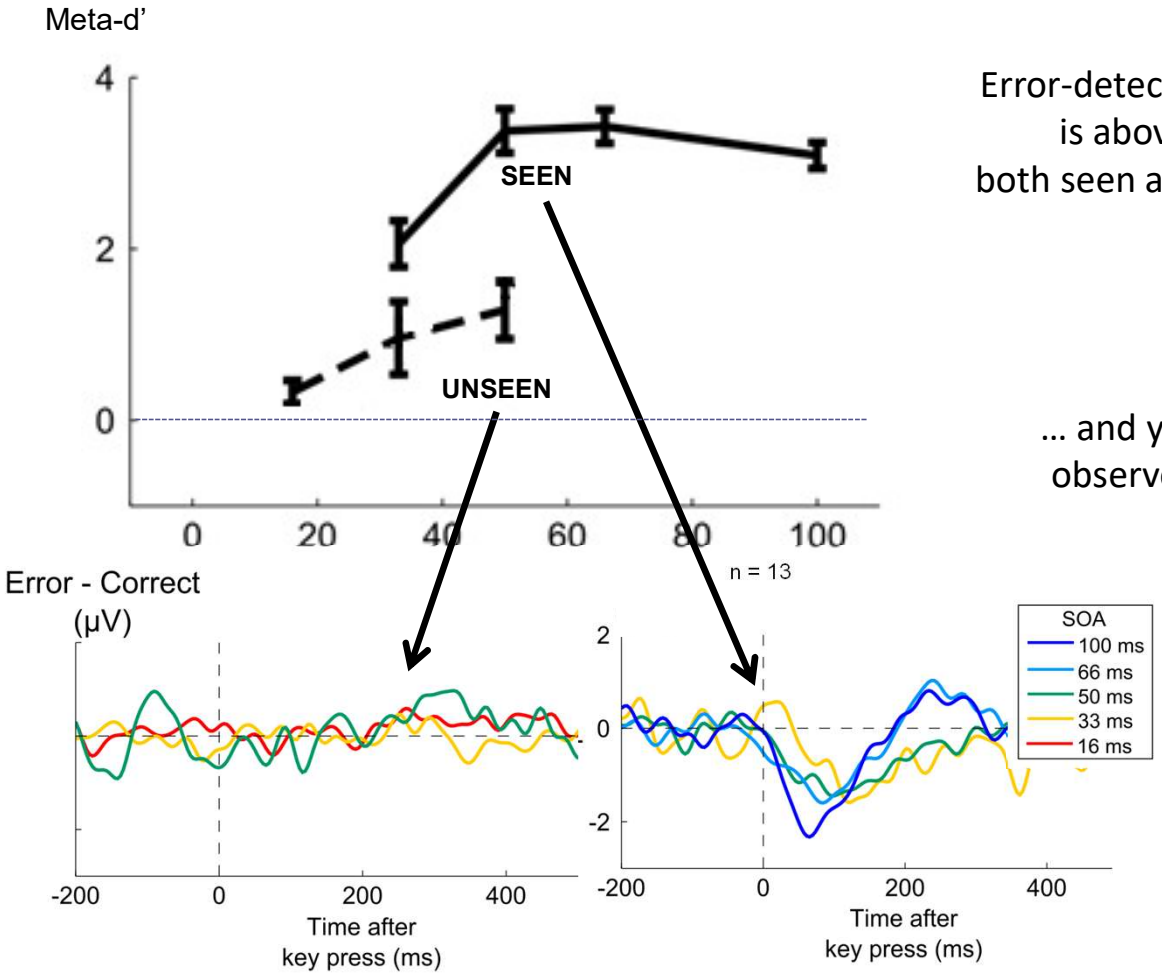
Type 2 and meta-d'

	Error	Correct
Perceived Error	Hit	Miss
Perceived Correct	False Alarm	Correct Rejection

- **SEEN condition**
  - Performance and Meta-performance are above chance level
- **UNSEEN Condition**
  - Performance is at chance
  - Meta-Performance **is above chance**
- Converging results in experiment 2
  - Both Performance and Meta-Performance are now above chance



Summary of the key dissociation



Error-detection performance is above chance on both seen and unseen trials…

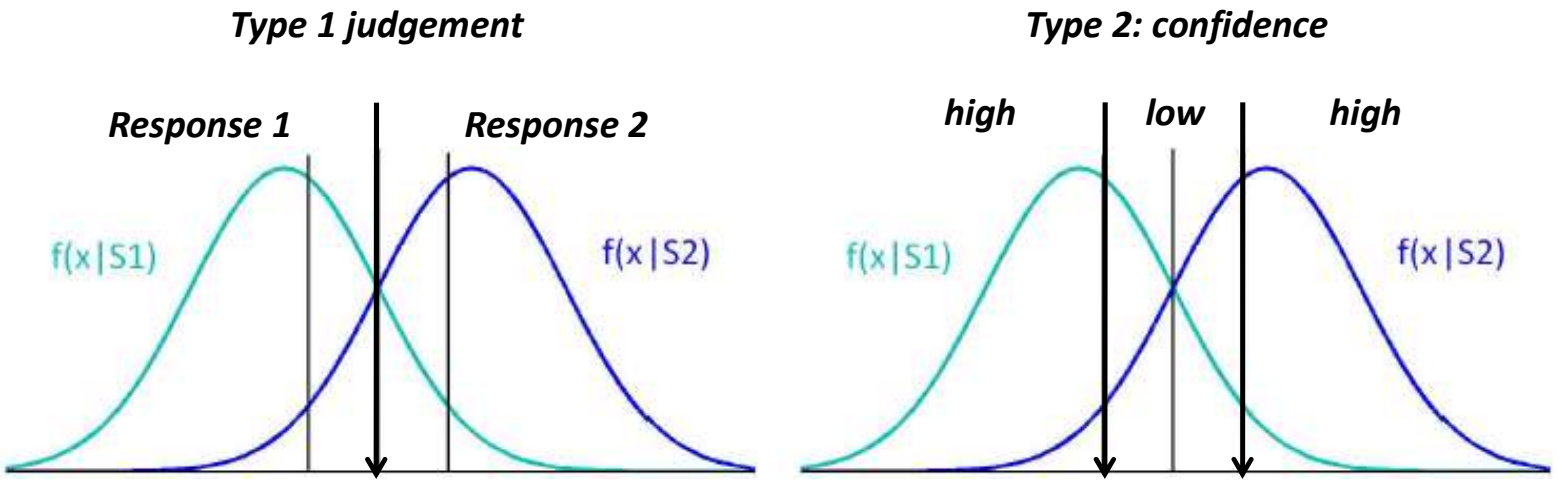
… and yet the ERN is only observed on seen trials.

→ Two dissociable processes of performance monitoring?

A signal-detection mechanism for nonconscious meta-cognition

Our subliminal behavioral data are compatible with type-II signal detection theory :

Any decision (even non-conscious) leads to the possibility of higher-than chance confidence, wagering, or error/correct judgment.

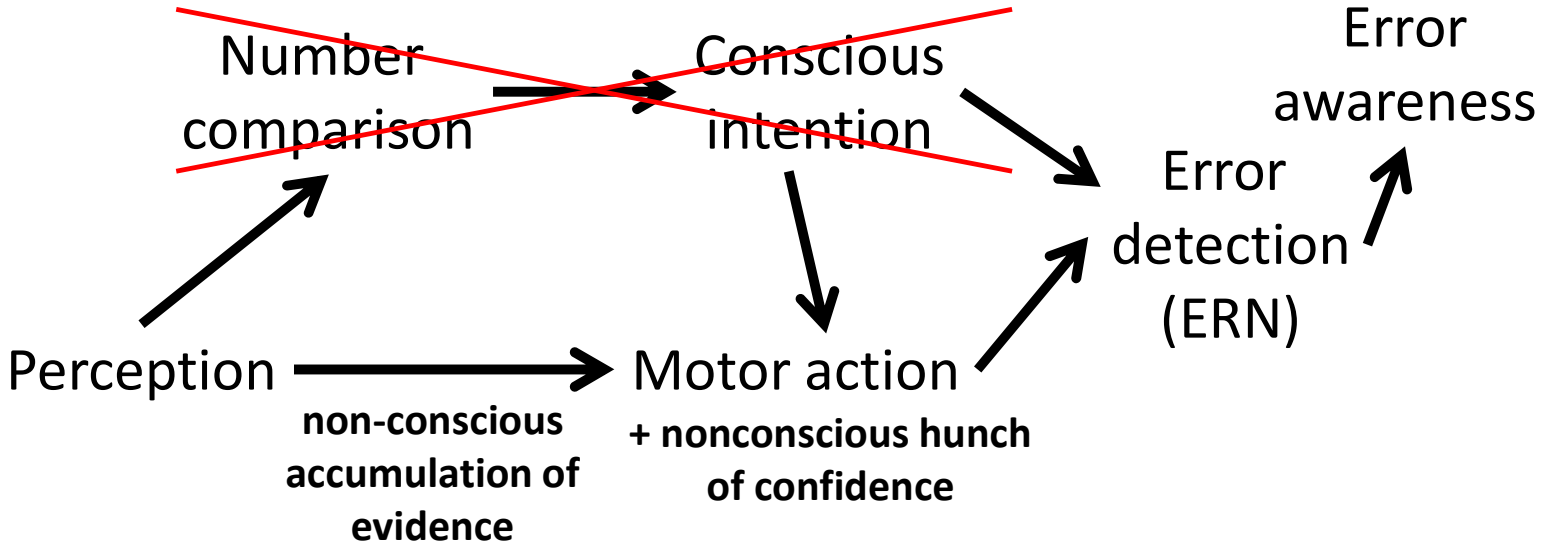


According to the Bayesian perspective, each brain area may encode the uncertainty associated with its computation.

Error negativity as a **comparison** of **conscious** **intention** and **actual action**

Charles, L., King, J.-R., & Dehaene, S. (2014). Decoding the dynamics of action, intention, and error detection for conscious and subliminal stimuli. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(4), 1158-1170. https://doi.org/10.1523/JNEUROSCI.2465-13.2014

A dual-route model of decision making (Del Cul, Dehaene et al., 2009)

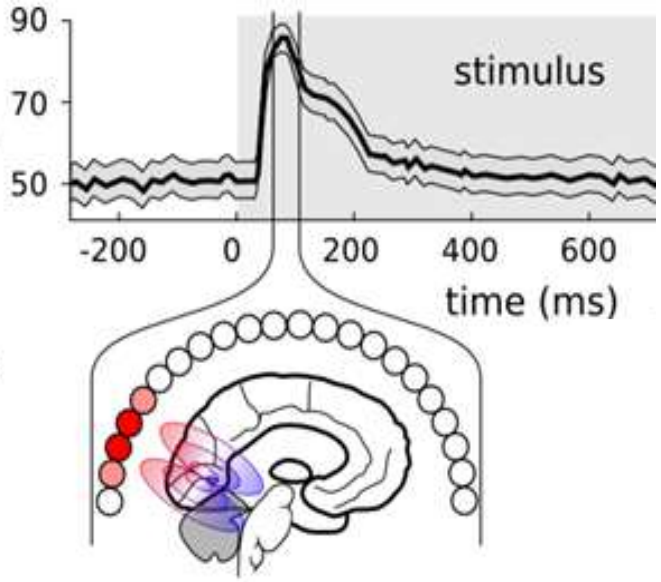


- When the digit is seen, there is both an action and an intention → ERN
- When the digit is NOT seen, there is an action but no intention → no ERN

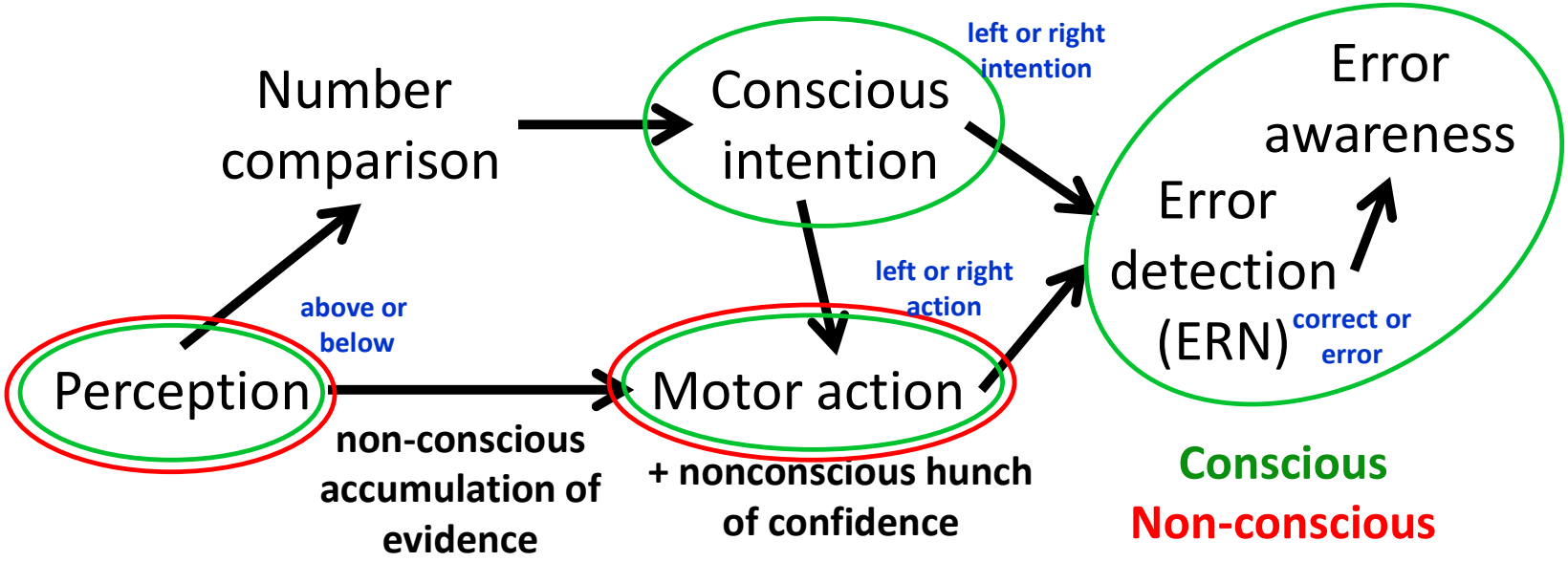
Multivariate decoding of MEG signals can reveal the time course of mental representations

King, J.-R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: the temporal generalization method. Trends in Cognitive Sciences, 18(4),

At each time point, we train a Support Vector Machine (SVM) to decode one aspect of the stimulus from the topography across sensors.



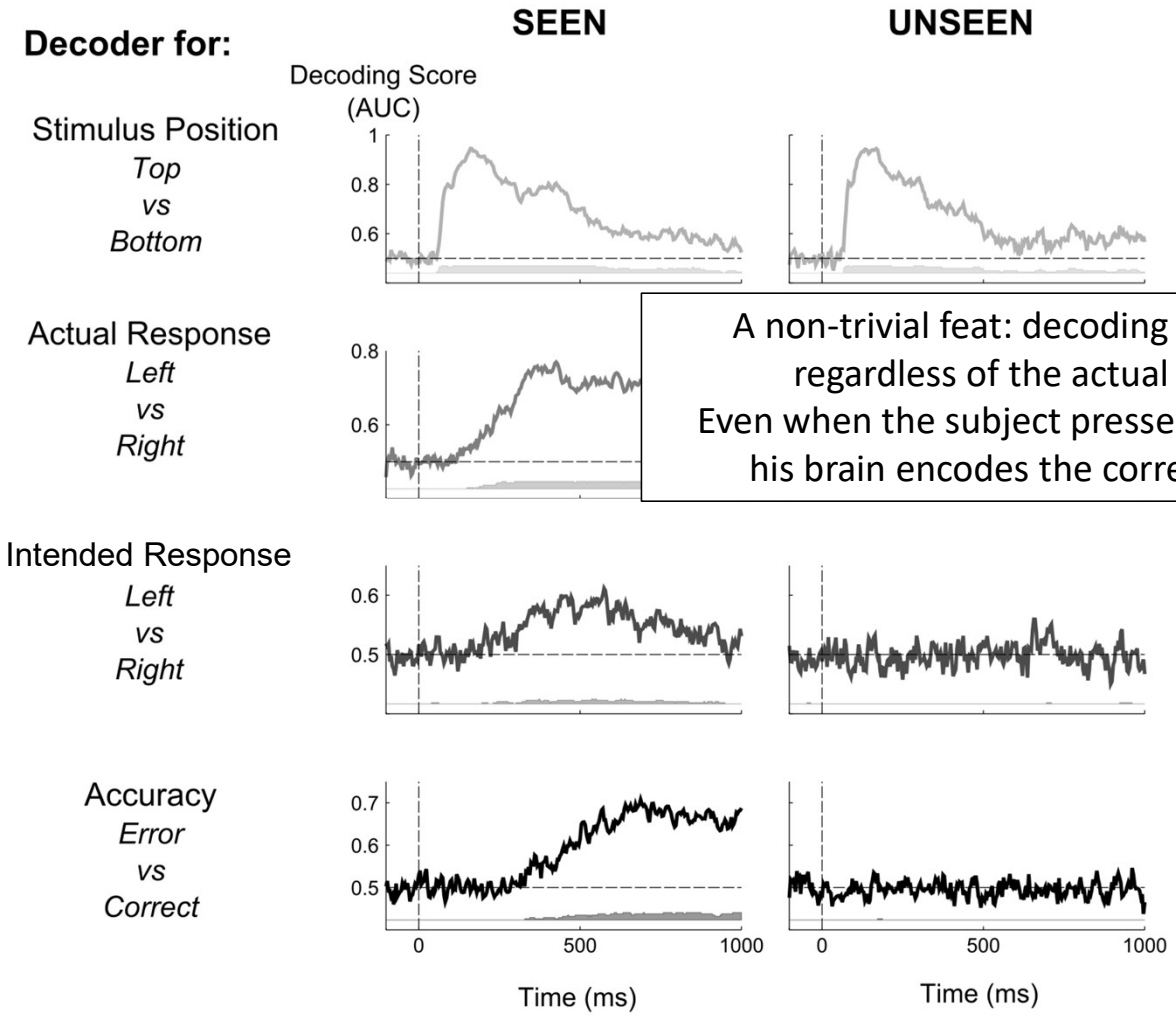
Predictions for multivariate decoding



- We can create a **decoder** that, for each time slice, uses the distribution of activation over 304 MEG and 256 EEG sensors to identify the current **mental content**.
- We should be able to decode the contents of **perception** and **motor action** on both **conscious** and **non-conscious** **trials**
- We should only be able to decode **intentions** and **errors** on **conscious** **trials**.

Decoding the time course of conscious and non-conscious contents

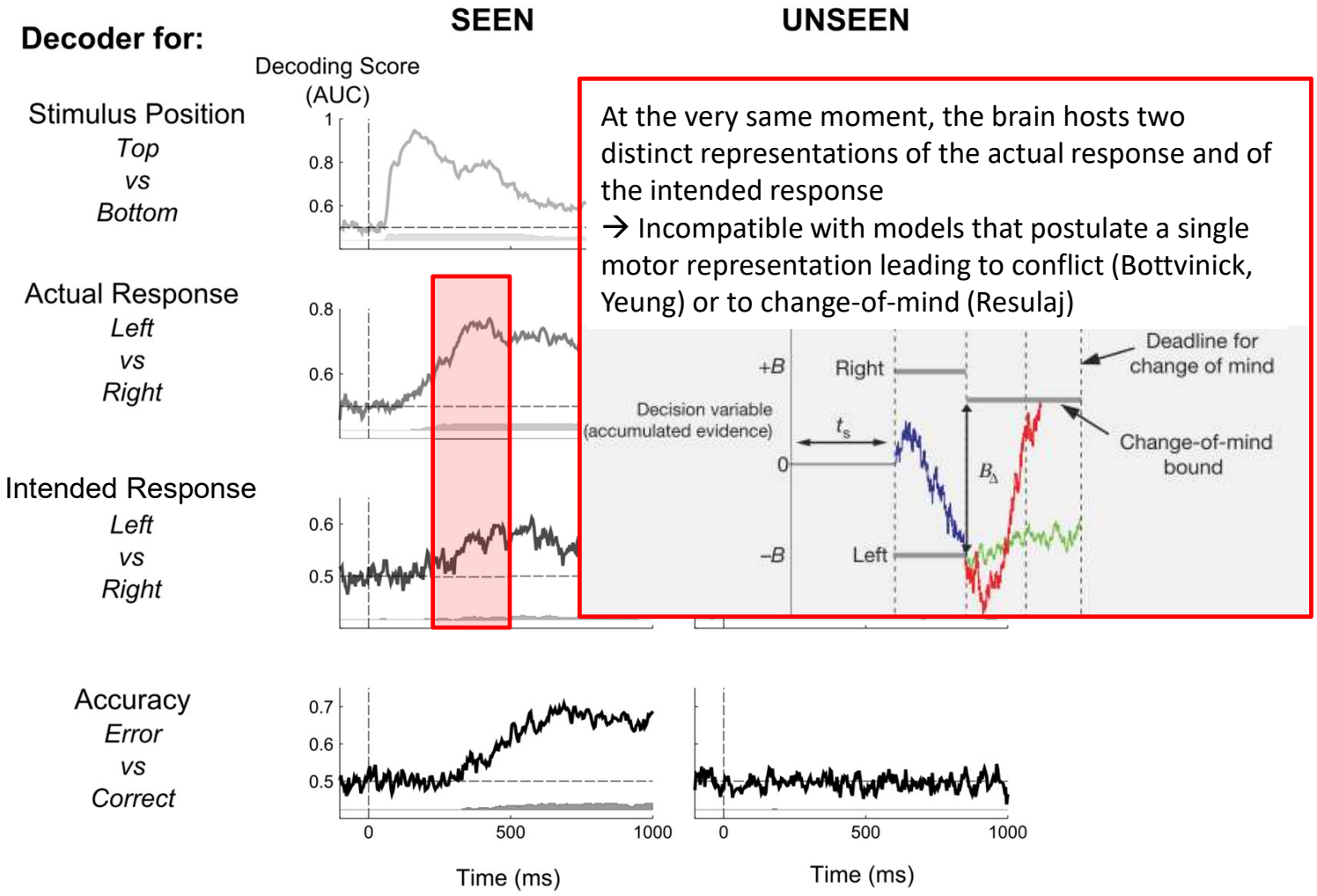
Charles et al., J Neuroscience 2014



A non-trivial feat: decoding the intention, regardless of the actual response!
Even when the subject presses the wrong key, his brain encodes the correct response.

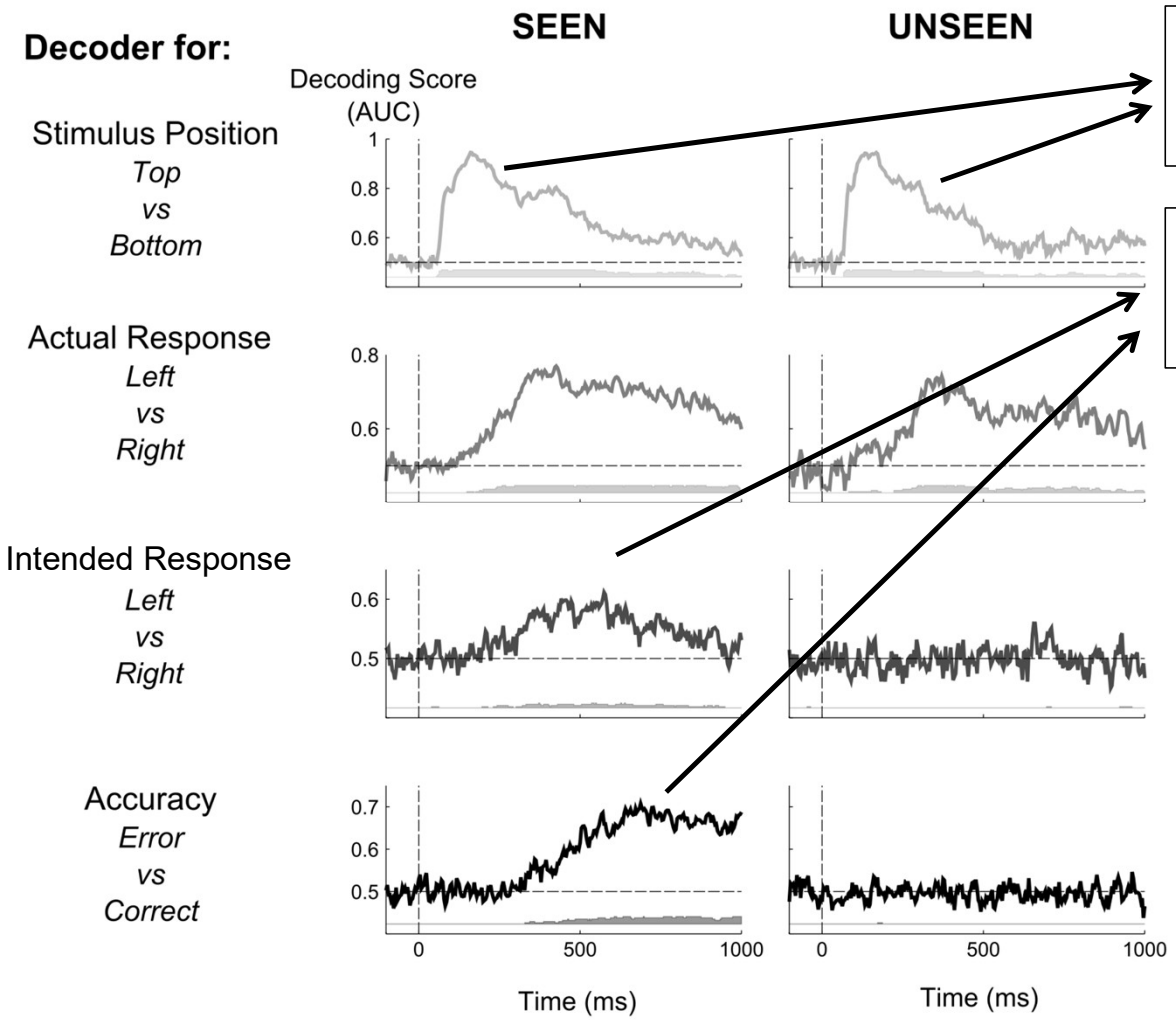
Decoding the time course of conscious and non-conscious contents

Charles et al., J Neuroscience 2014



Decoding the time course of conscious and non-conscious contents

Charles et al., J Neuroscience 2014



amplification of sensory information

uniquely conscious computations

Conclusion :

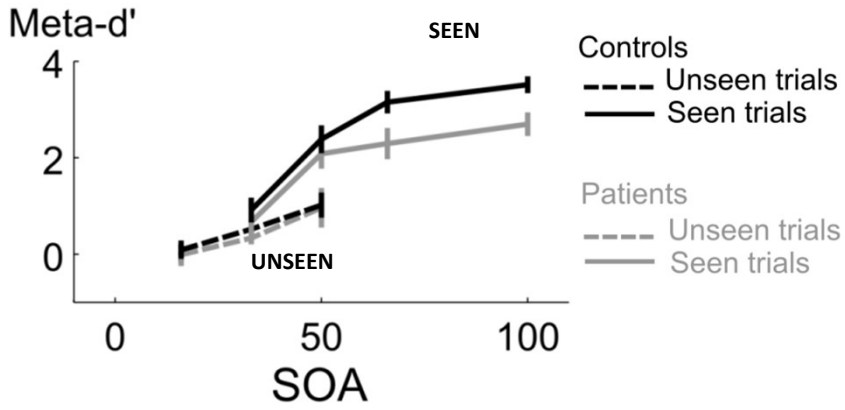
The probabilistic computations underlying decision and confidence can partially unfold nonconsciously.

But… the precise trial-by-trial monitoring of our performance requires conscious processing.

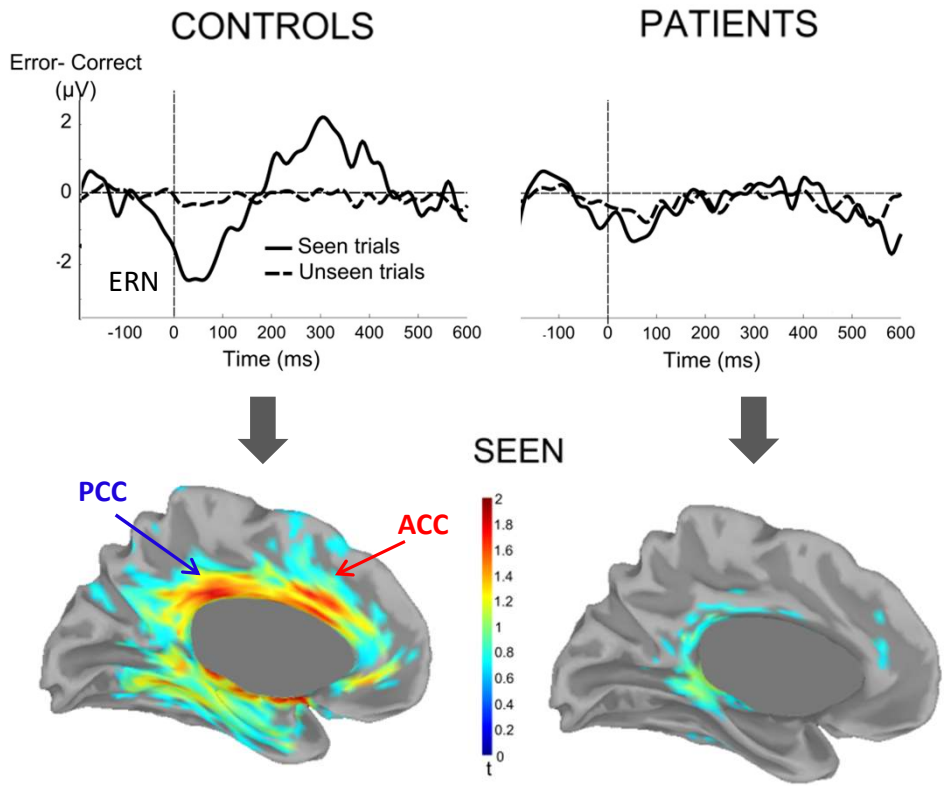
Impaired conscious, but preserved non-conscious metacognition in schizophrenia

Charles, L., Gaillard, R., Amado, I., Krebs, M.-O., Bendjemaa, N., & Dehaene, S. (2017). Conscious and unconscious performance monitoring : Evidence from patients with schizophrenia. *NeuroImage*, *144*(Pt A), 153-163. https://doi.org/10.1016/j.neuroimage.2016.09.056

Behavioral data



MEEG data



Two computationally distinct systems :

Non-conscious metacognition is preserved: patients are identical to controls in having a hunch of when they made an error, even when they did not see the digit.

Conscious error detection is impaired :

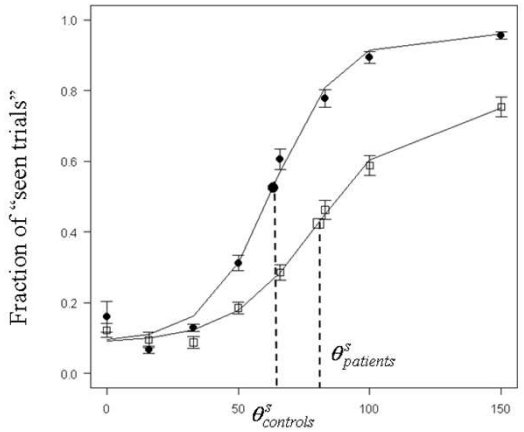
Behaviorally, patients show a reduced sensitivity to their errors;

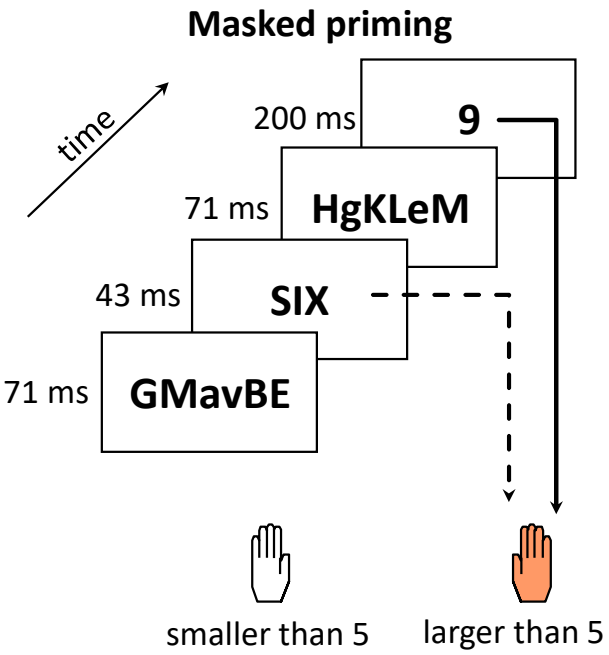
and the cingulate response (ERN) is drastically reduced.

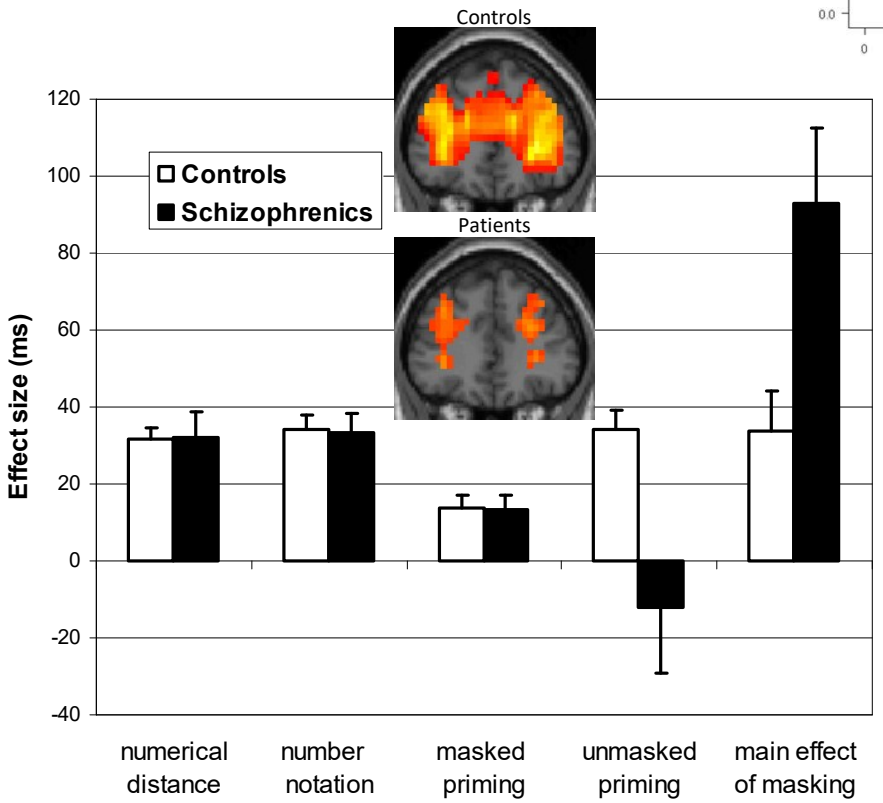
Schizophrenia as an impairment of conscious access, with preserved subliminal processing

Dehaene et al., *PNAS* 2003; Del Cul et al., *Arch. Gen. Psychiatry* 2006; **Review in Berkovitch, Gaillard & Dehaene, TICS 2017**

Patients with schizophrenia exhibit an elevated threshold for conscious visibility

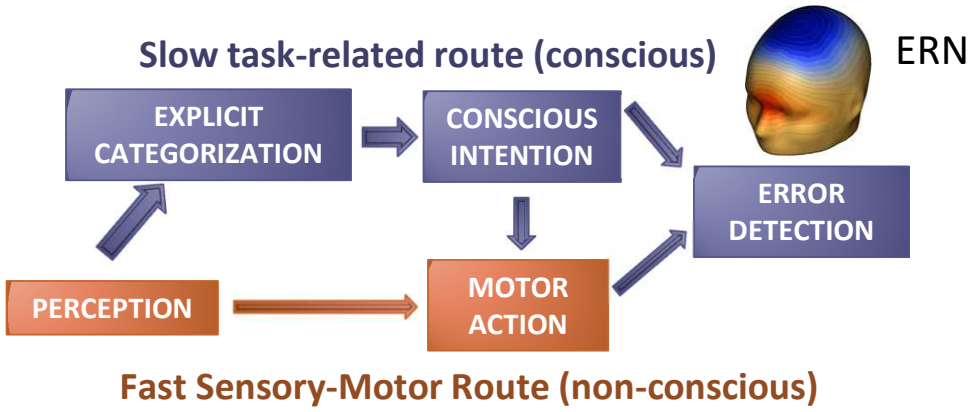






During masked versus unmasked priming, patients with schizophrenia show normal subliminal processing, but impaired conscious control associated with reduced PFC/ACC activation.

Conclusions: multiple systems for error monitoring

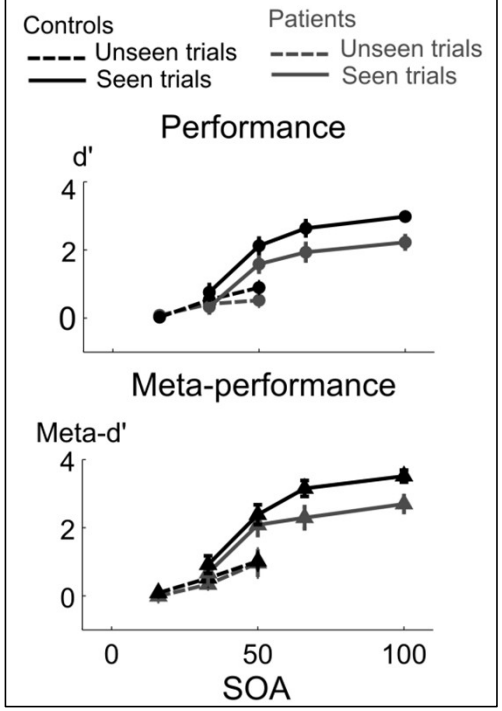


ERN

1. A **statistical assessment of confidence** : every processor may compute over probability distributions
   - → Can be completely nonconscious
   - → Dissociable from the ERN and the anterior cingulate

2. A **single-trial process of response evaluation**, based on the comparison of intended and actual responses.
   - → Associated with the Error-Related Negativity (ERN)
   - → Which requires consciousness of the intention, and often leads to consciousness of the error

In schizophrenia, there is **impaired** conscious perception, explicit error detection, ERN and anterior cingulate… but **preserved** implicit sense of confidence:



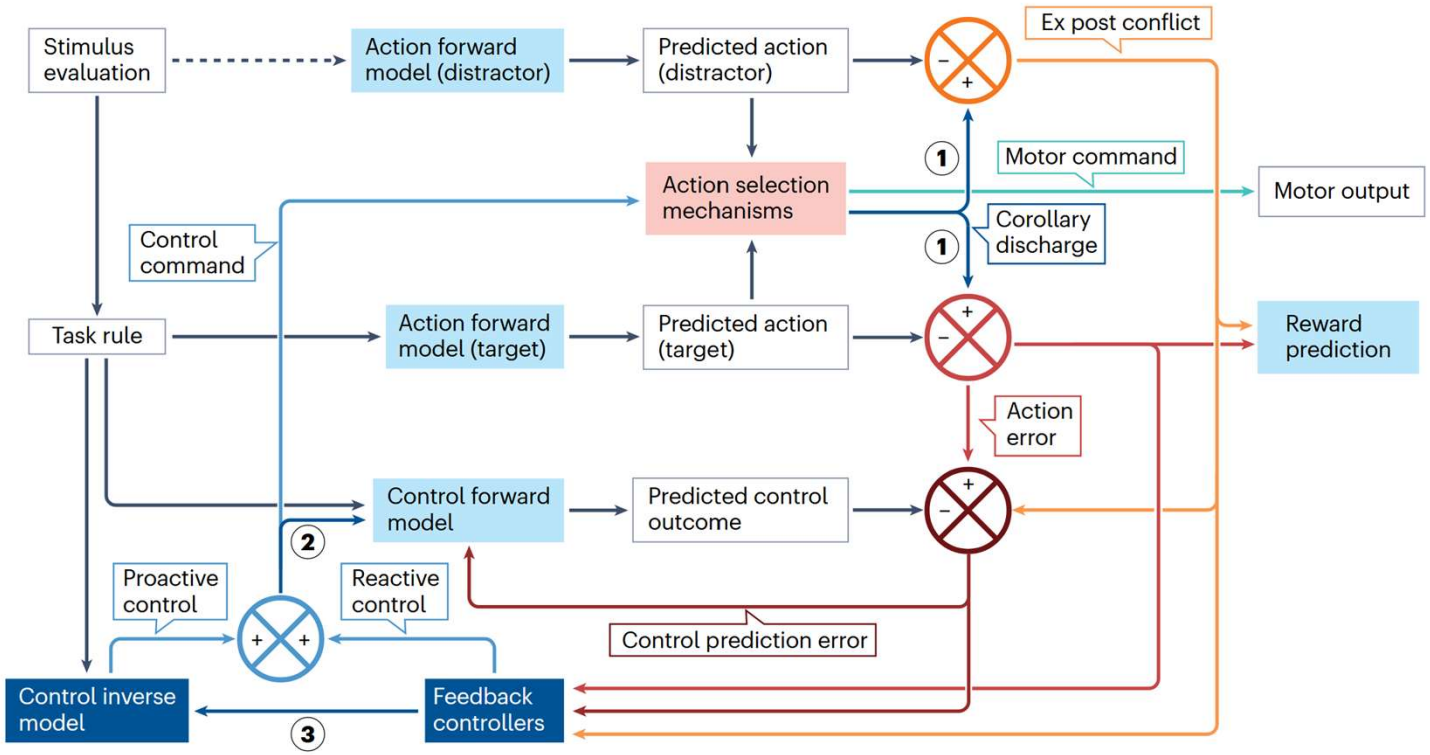


Fig. 5 | Conceptual framework for action error computation. The available actions under a given task rule and stimulus are predicted by action forward models (light blue). This includes both the correct response (target) and the incorrect response (distractor). The action selection process (red box) then chooses between one of the possible actions. Action selection is modulated by the control command (blue line), which is composed of proactive and reactive components (blue). The feedback controllers use performance-monitoring goal-compatible action. The orange cross computes ex post conflict signals that are the result of comparing the selected action, conveyed as corollary discharge 1, with the predicted goal-incompatible action. The dark red cross computes control prediction error by comparing the predicted control outcome and the actual control outcome (error, conflict); it can also recruit feedback control. Action errors and ex-post conflict are used to predict the occurrence of reward. The control forward model predicts whether the current control

Fu, Z., Sajad, A., Errington, S. P., Schall, J. D., & Rutishauser, U. (2023). Neurophysiological mechanisms of error monitoring in human and non-human primates. *Nature Reviews Neuroscience*, *24*(3), 153-172. https://doi.org/10.1038/s41583-022-00670-w

General conclusions : the depth of non-conscious processing

Below the level of our awareness, **virtually all brain regions** are constantly performing complex computations :

- Invariant visual recognition
- In parallel across multiple items
- **Language processing** : the human brain can process several words nonconsciously, and integrate them into a simple phrase
  - but within limits: full sentence-level integration may require consciousness.
- **Metacognition** : our brain can nonconsciously compute a hunch of the probability of being correct
  - again within limits: single-trial error detection and self-reflection probably require consciousness.

Next week : neural correlates of crossing the consciousness threshold.