

Cours 2013:

Le bébé statisticien

Stanislas Dehaene

Chaire de Psychologie Cognitive Expérimentale

Cours n°2

Le sens des probabilités chez l'enfant



Corrège, *La sainte famille* (© Musée des Beaux-Arts d'Orléans)

L'hypothèse du « bébé statisticien »

L'enfant serait capable, très tôt:

de se représenter des distributions de probabilité
de mettre à jour ces probabilités, en appliquant
la règle de Bayes

et ce, à de multiples niveaux hiérarchiques
de représenter mentalement de très vastes
espaces d'hypothèses, hiérarchisées, chacune
avec sa probabilité a priori et sa
vraisemblance

d'évaluer simultanément la plausibilité de
plusieurs hypothèses:

- non seulement pour des hypothèses
« statiques »

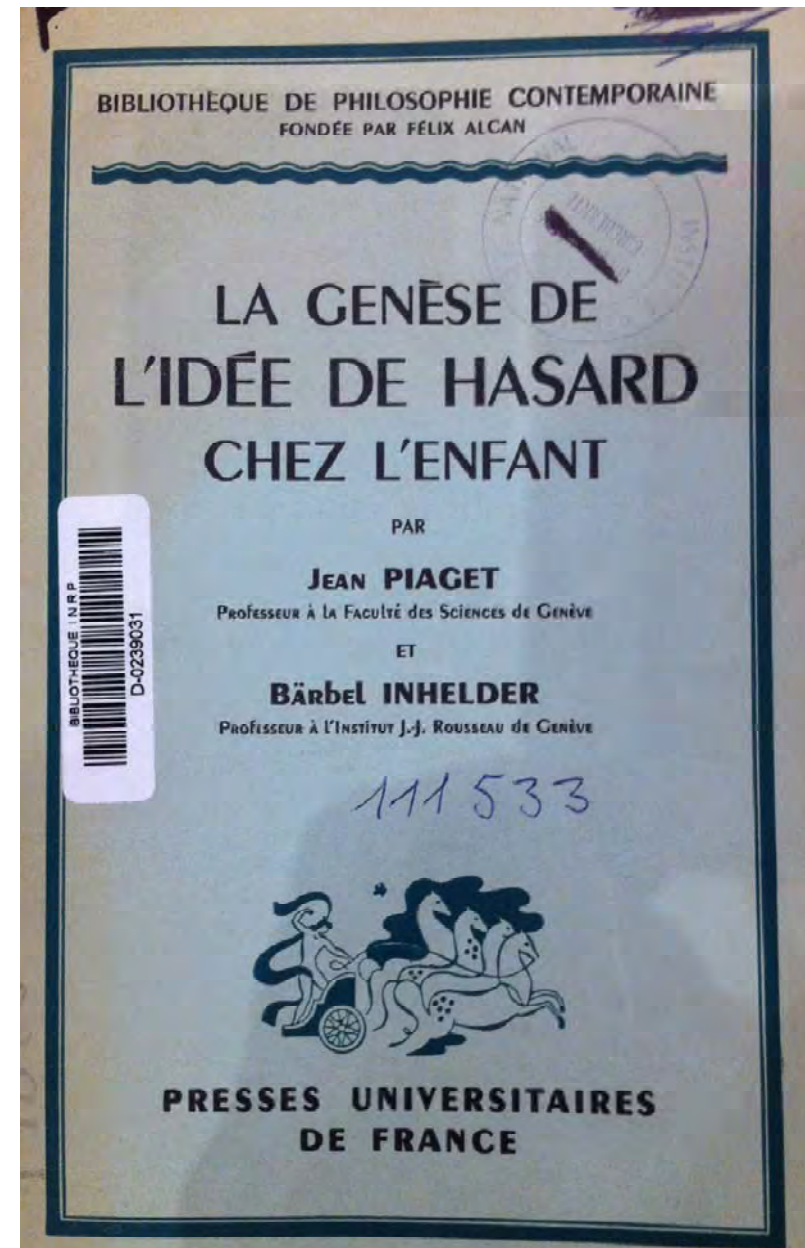
- mais également pour des hypothèses « dynamiques » : des programmes récur­sifs en nombre potentiellement infini
- de les utiliser pour générer des prédictions et les comparer aux données reçues du monde extérieur (erreur de prédiction, signal de surprise)

Piaget: La genèse de l'idée de hasard chez l'enfant (1951)

Piaget pose avec acuité la question de l'existence d'un « sens des probabilités »:
« Un mathématicien (...) nous suggérerait un jour l'étude de ce problème psychologique: n'existerait-il pas, chez l'homme normal, une « intuition de la probabilité » aussi fondamentale et d'usage aussi fréquent que, par exemple, l'intuition du nombre entier? ».

Sur la base des travaux de Levy-Bruhl, Piaget conclut que l'absence de la notion de hasard fait partie des « caractères essentiels de la mentalité primitive ».

Il pose alors une hypothèse similaire dans le domaine du développement: « ne se pourrait-il pas que la découverte de l'idée de hasard, c'est-à-dire la compréhension même de l'irréversibilité, dût attendre le niveau de l'achèvement des opérations réversibles? ».
Le hasard serait compris tardivement, comme l'antithèse de l'effet des opérations réversibles.



Piaget: La genèse de l'idée de hasard chez l'enfant (1951)

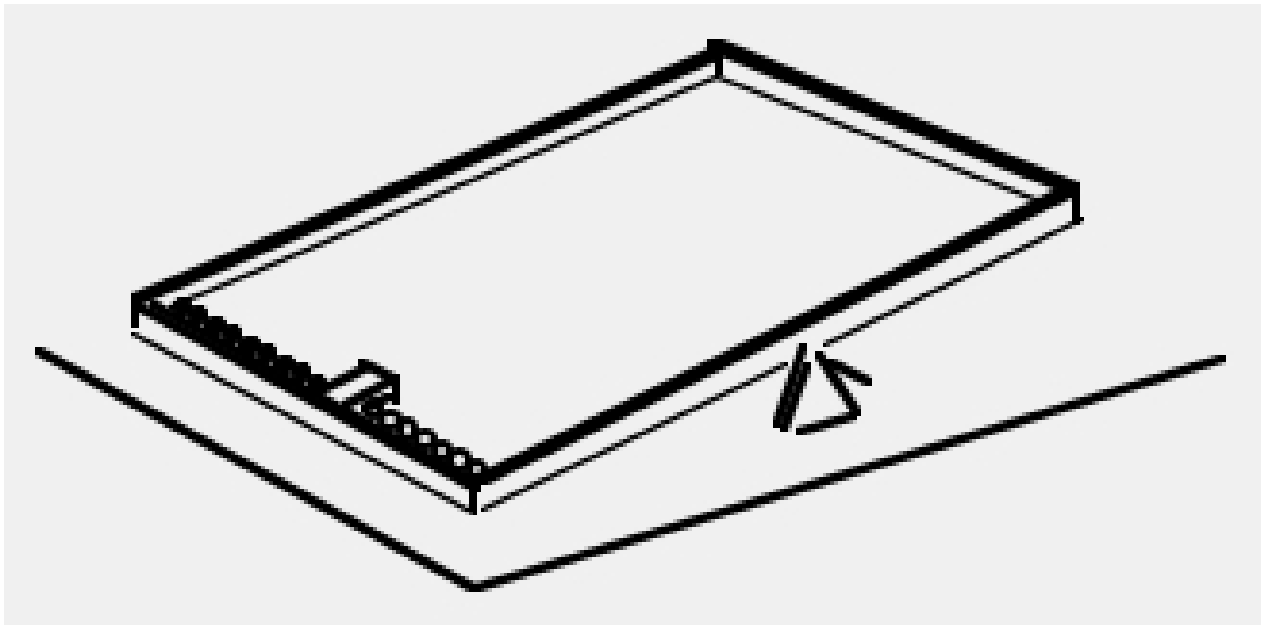
Un exemple d'expérience: l'intuition de l'irréversibilité du hasard

Huit perles rouges et huit perles blanches sont alignées, séparées par une petite cloison.

« Lors de chaque mouvement de bascule, les perles se dirigeront vers le petit côté opposé, puis reviendront se placer en position initiale, mais **avec une série de permutations possibles.** »

Piaget demande à l'enfant de prédire les effets de la bascule, et de dessiner les trajectoires des billes.

Il trouve, à quatre-cinq ans, une intuition de la réversibilité. Même après un premier mélange, les billes vont revenir à leur place, « elles savent où elles roulent, puisqu'elles roulent toutes seules ».



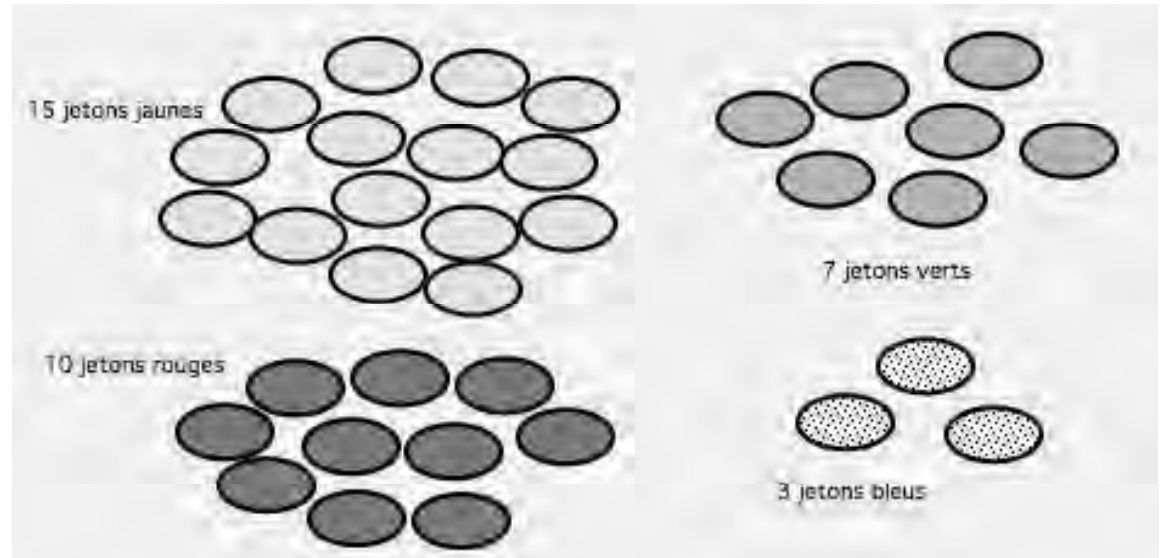
Ce ne serait que vers 7-11 ans qu'apparaît l'idée du hasard, « mais les opérations élémentaires d'ordre (ordonner une suite ABC...) sont loin de suffire à la constitution d'un schème complet de permutations proprement dites ». Cette maîtrise apparaîtrait vers 12 ans.

Piaget: La genèse de l'idée de hasard chez l'enfant (1951)

L'intuition des tirages aléatoires:

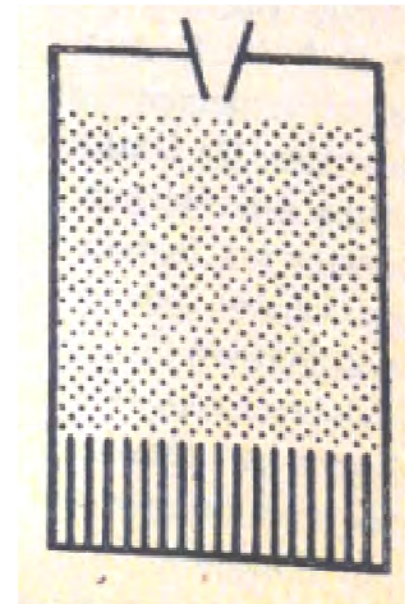
L'enfant tire des jetons d'un sac opaque. Il a sous les yeux une copie du contenu du sac. On lui demande de prédire le résultat de ses tirages.

Leu (5:8): 6 rouges, 4 bleus, 1 blanc: « si tu prends une puce dans le sac, n'importe laquelle, de quelle couleur sera-t-elle plutôt? » « la blanche » « pourquoi » « Parce qu'il n'y a qu'une blanche »...



L'intuition des distributions:

Les enfants savent-ils prédire la distribution qui résulte d'une multitude d'évènements aléatoires?



Piaget: La genèse de l'idée de hasard chez l'enfant (1951)

Piaget tire de ses études une série de conclusions... discutables. Selon lui:

1. Le concept de hasard se construit.

2. Il existerait une « étroite corrélation » entre « la formation des notions de hasard et de probabilité » et « celle des diverses opérations dont nous connaissons aujourd'hui la genèse psychologique »

- avant 8 ans: « pas d'opérations proprement dites , c'est-à-dire de composition réversible »: l'enfant se repose sur des « raisonnements prélogiques ». A ce moment, « l'enfant ne distingue pas le possible et le nécessaire », « il n'y a dès lors ni hasard ni probabilité ».

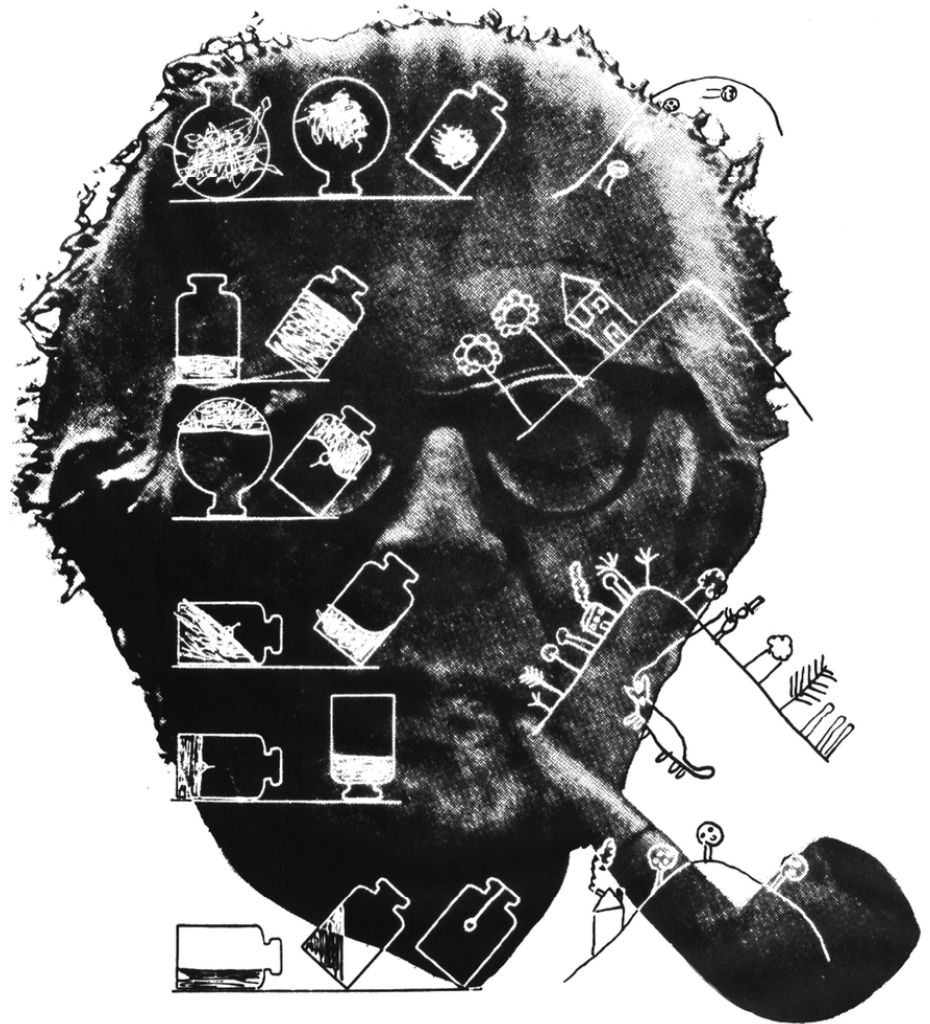
- de 7-8 à 11-12 ans, « construction de groupements opératoires d'ordre logique et des relations numériques, mais sur un plan essentiellement concret ». La découverte de la nécessité déductive ou opératoire permet de concevoir, par antithèse, la notion de transformation fortuite.

-Ce ne serait que vers 12 ans que se met en place une représentation du hasard comme « l'indéterminé », en distinguant l'indétermination logico-arithmétique (ne pas savoir quelles opérations ont été appliquées) et l'indétermination spatio-temporelle ou physique (séries causales indépendantes, paramètres physiques non-spécifiés).

Piaget: La genèse de l'idée de hasard chez l'enfant (1951)

Quelles sont les erreurs de Piaget?

- son approche ne fait guère appel à des situations simples et classiques en théorie des probabilités, où l'on peut faire des prédictions mathématiques précises.
- Elle s'appuie sur l'introspection verbale des enfants ainsi que sur leurs dessins – plutôt que sur les connaissances implicites que l'on peut déceler dans leur regard, leur surprise, leur choix spontané...
- Elle est entièrement déterminée par un cadre théorique *a priori* fondé sur la hiérarchie des opérations logico-mathématiques. On peut se demander à quel point ce cadre a influencé sur le choix des expériences et même des observations à rapporter (qui sont toujours décrites sous forme de listes d'observations ou de dialogues anecdotiques, sans statistiques).



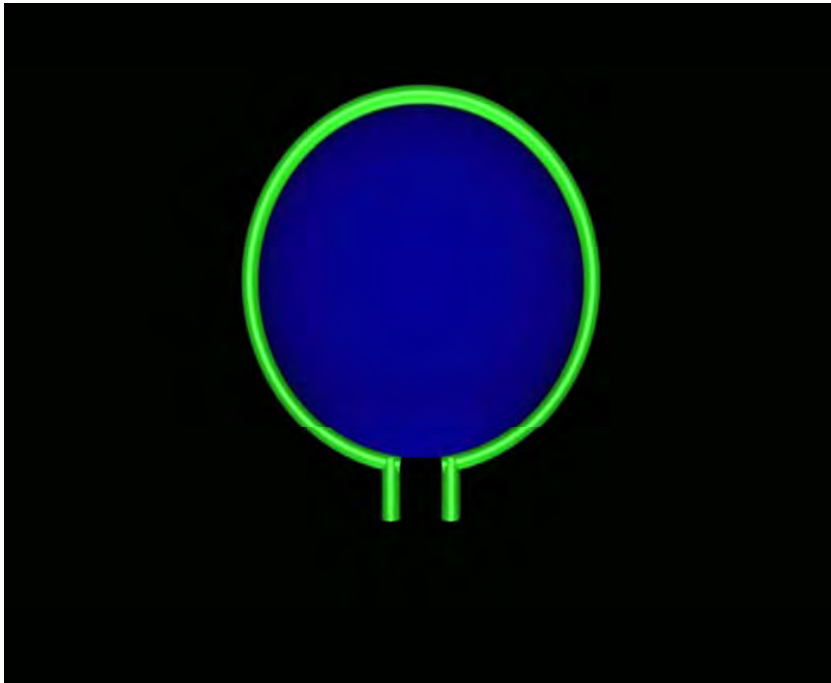
Jean Piaget (1896–1980)

<http://neuroportraits.eu/portrait/jean-piaget>

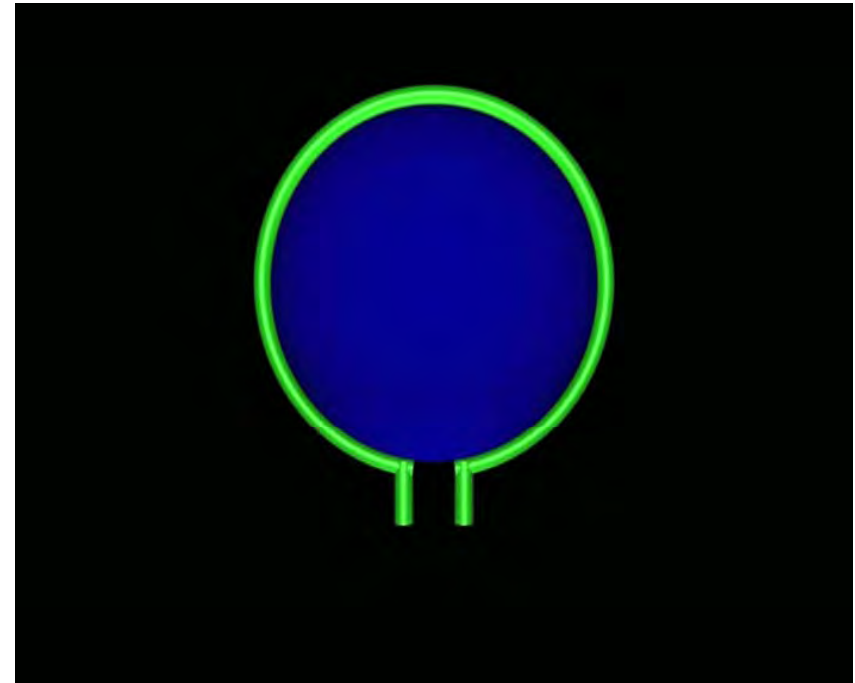
Le bébé statisticien: les études de Teglas et Bonatti

Teglas, E., Girotto, V., Gonzalez, M., & Bonatti, L. L. (2007). Intuitions of probabilities shape expectations about the future at 12 months and beyond. *Proc Natl Acad Sci U S A*, 104(48), 19156-19159.

Les bébés de 12 mois peuvent-ils anticiper la probabilité d'un événement futur?



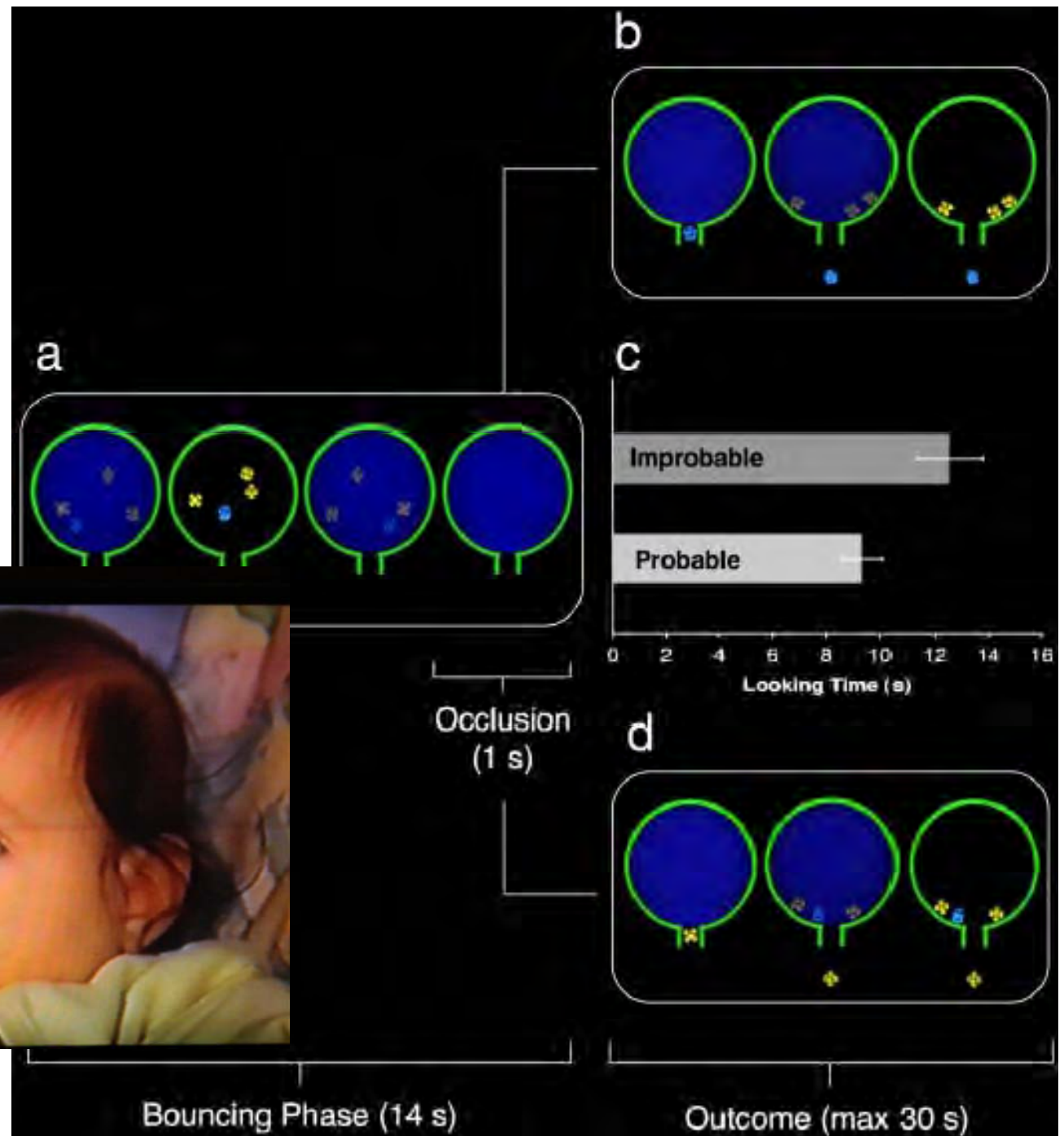
Événement improbable



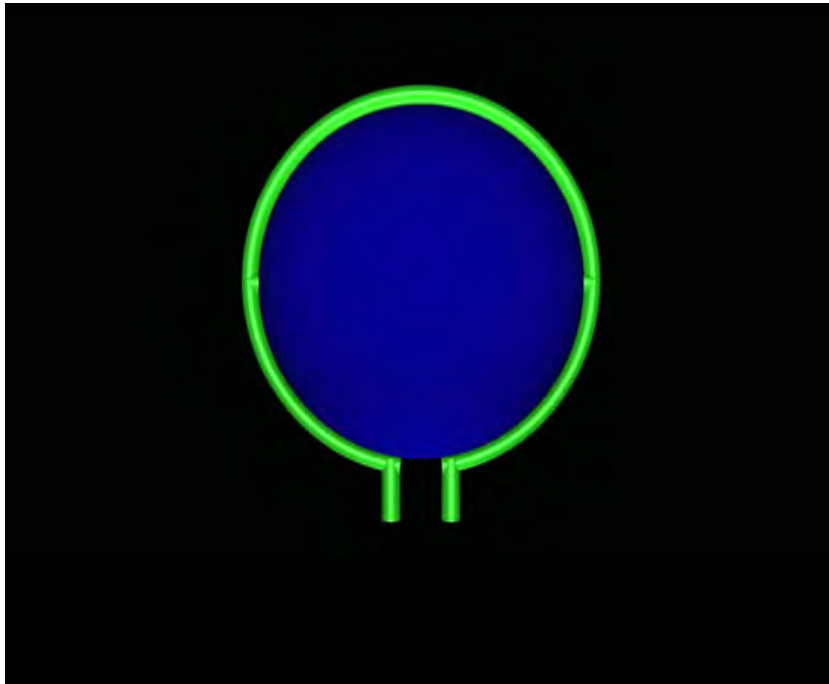
Événement probable

Résultats:

Les bébés de 12 mois regardent plus longtemps l'événement improbable que l'événement probable.



Ces résultats peuvent-ils s'expliquer par un simple biais perceptif?



Un contrôle expérimental astucieux:

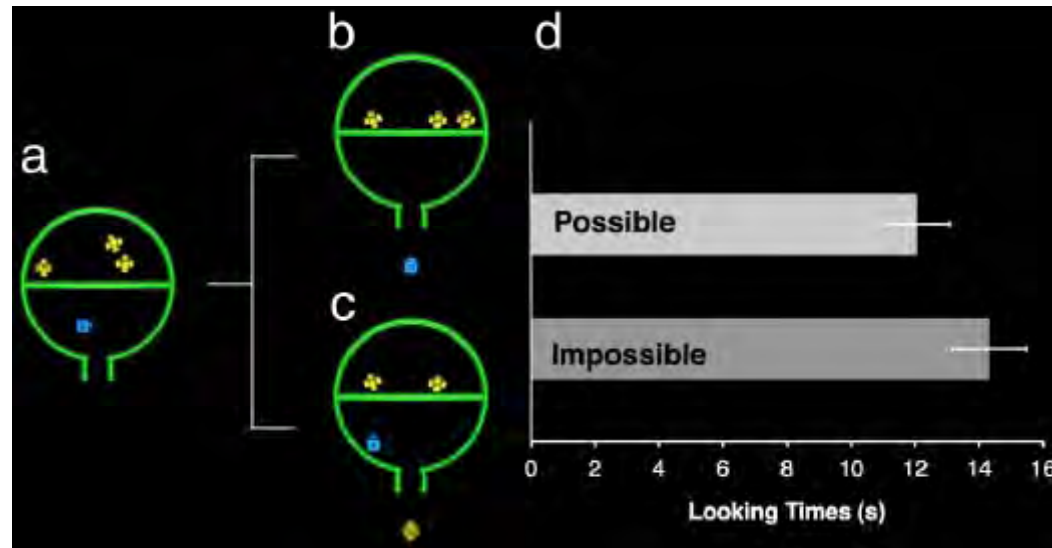
- le nombre d'objets est exactement le même qu'auparavant.

- une cloison rend impossible la sortie des objets les plus nombreux

- ainsi les résultats de l'expérience sont inversés: la seule issue possible est la sortie de l'objet unique

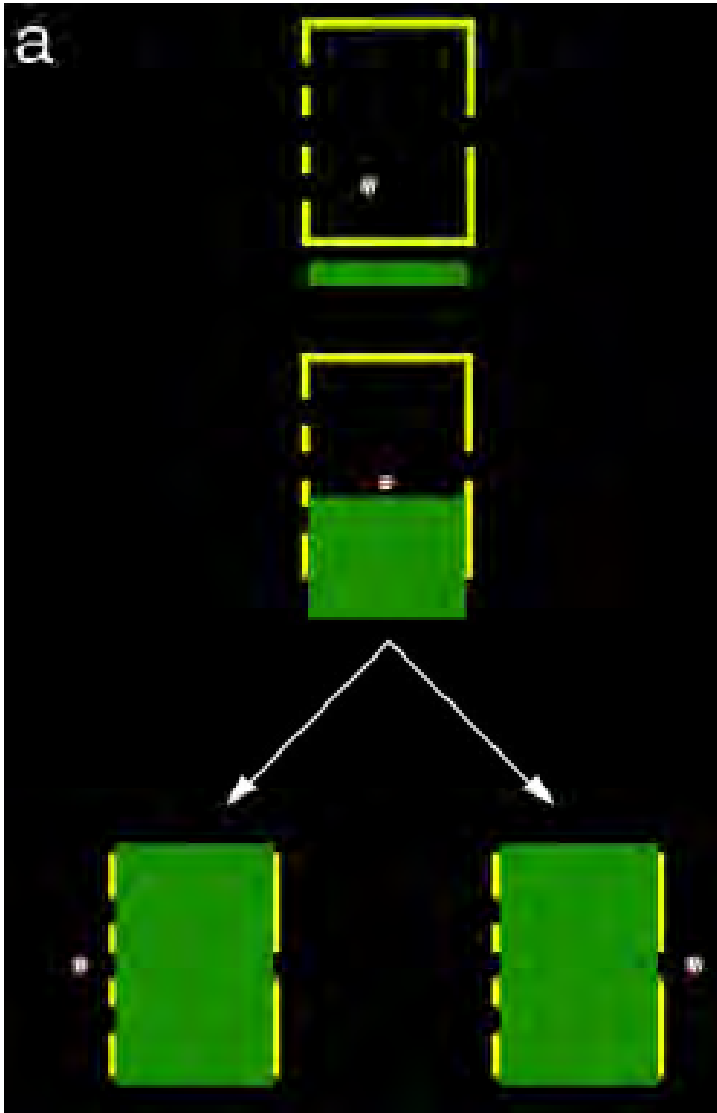
Résultats:

- la durée du regard s'inverse
- l'enfant regarde plus longtemps l'événement impossible, qui correspond à présent à la sortie d'un des objets les plus nombreux.



Quand la probabilité et la fréquence entrent en conflit

Teglas, E., Girotto, V., Gonzalez, M., & Bonatti, L. L. (2007). Intuitions of probabilities shape expectations about the future at 12 months and beyond. *Proc Natl Acad Sci U S A*, 104(48), 19156-19159.



Dans les expériences précédentes, la fréquence ne joue aucun rôle, car les films sont vus très peu de fois, et aussi fréquemment dans la condition probable que dans la condition improbable (2 fois).

Chez l'enfant de 3 et 5 ans, il est possible de prolonger l'expérience et donc d'étudier la relation entre la probabilité *a priori* et la fréquence.

Chez deux groupes d'enfants:

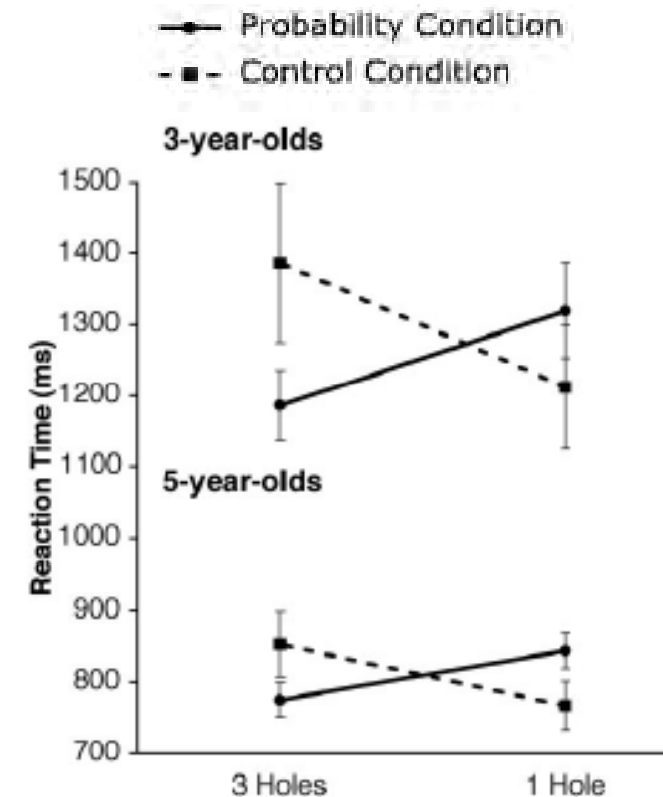
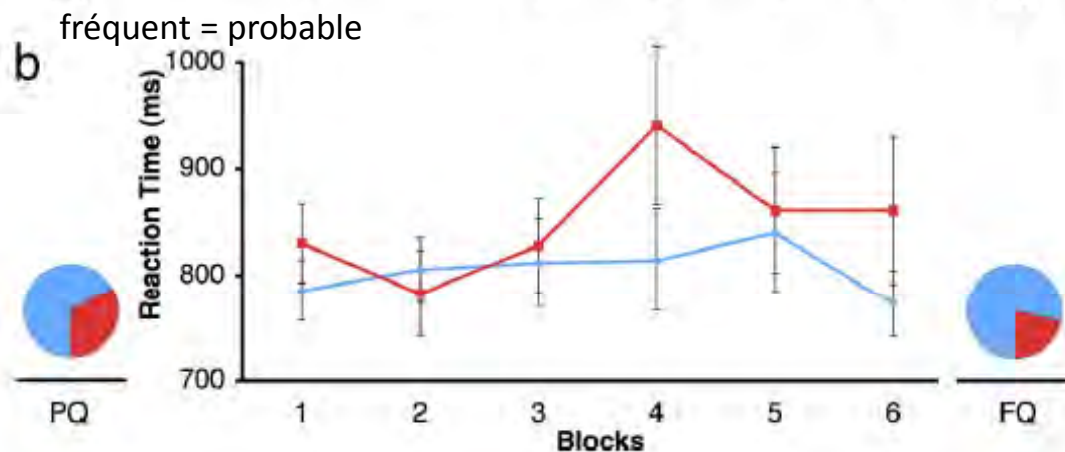
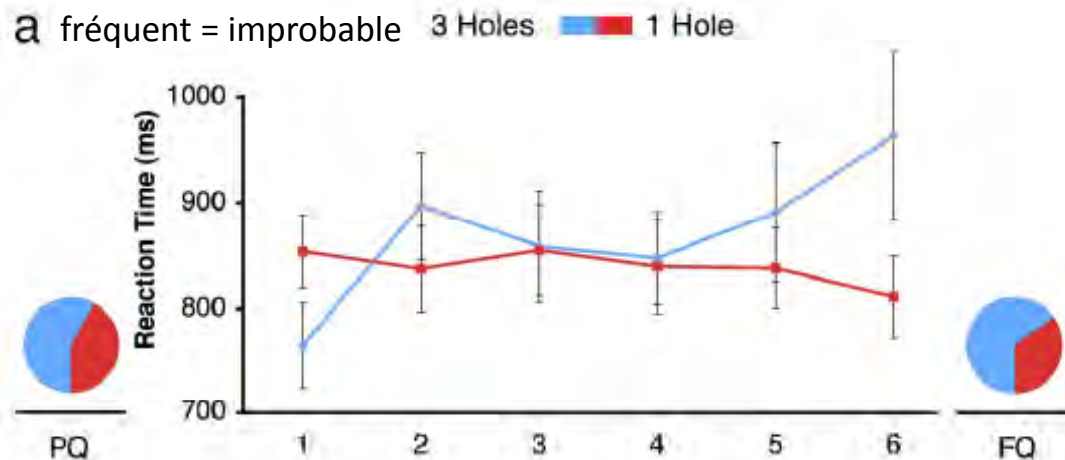
- l'objet sort plus fréquemment du côté à un trou (improbable)
- ou bien du côté à trois trous (probable)

Mesure expérimentale = temps de réaction pour détecter la sortie de l'objet.

Quand la probabilité et la fréquence entrent en conflit

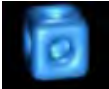
Résultats: dans le premier bloc, seule compte la probabilité.
Les réponses sont plus lentes quand l'objet sort du côté à un seul trou.

Contrôle: ce résultat s'inverse quand on explique à l'enfant que c'est une balle spéciale qui a envie de sortir de ce côté.



Seulement chez l'enfant de 5 ans, les résultats s'inversent progressivement en fonction de la fréquence des essais.
Et cependant, par oral, ils continuent de répondre en fonction de la probabilité!
Conclusion: le sens des probabilités est chevillé au corps des jeunes enfants.

L'interprétation de Luca Bonatti: un calcul **logique** et pas seulement **statistique**

A: Event representation	The Fact level	
B: Representation of possible outcomes	The Logical Level	 
C: Numerical representation	The Number Level	 
D: Probabilities	The Probability Level	Yellow > Blue
E: Predictions	The Future Level	Expect Yellow

L'enfant n'a pas eu la possibilité de voir suffisamment d'essais pour calculer la **fréquence** de chacune des alternatives. Son comportement reflète donc une inférence logique sur la **plausibilité** des observations, calculable en un seul essai.
 (point de vue « Bayésien » opposée à l'interprétation « fréquentiste » des probabilités)

L'interprétation de Luca Bonatti: un calcul **logique** et pas seulement statistique

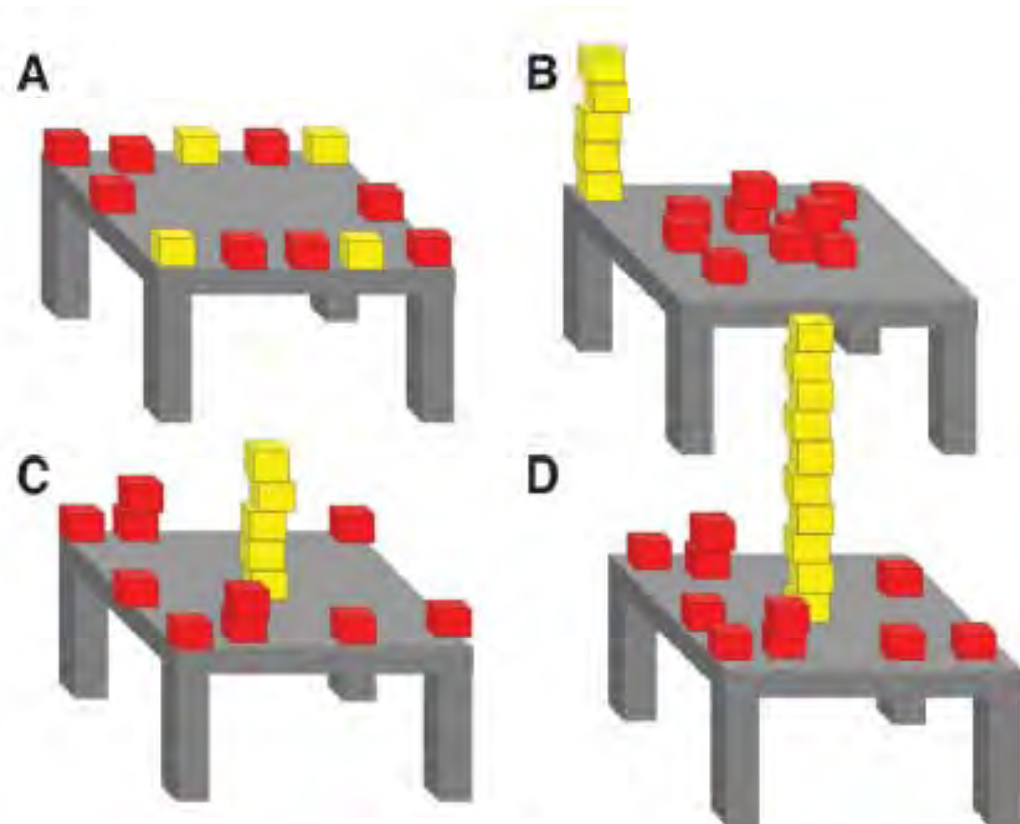
Notre connaissance du monde nous permet d'effectuer des calculs **logiques, prédictifs**, fondés sur le jugement de **plausibilité**, dans toutes sortes de situations **nouvelles**.

Si la table est secouée, quelle est la probabilité que le premier cube à tomber soit jaune?

Nous sommes tous dotés d'un certain "bon sens" ou « sens commun ».

Ce "bon sens" est-il bayésien?

Est-il disponible chez le bébé de quelques mois?



Les enfants sont-ils capables de combiner plusieurs indices?

Teglas, E., Vul, E., Girotto, V., Gonzalez, M., Tenenbaum, J. B., & Bonatti, L. L. (2011). Pure reasoning in 12-month-old infants as probabilistic inference. *Science*, 332(6033), 1054-1059.



Astuce nouvelle:
Le film est masqué alors que l'un des deux groupes d'objets est proche de la sortie.

- Les enfants tiennent-ils compte de la proximité?
- Comment pondèrent-ils proximité et probabilité?

Temps d'occlusion

Court



Long

Indice physique:
proximité de la sortie

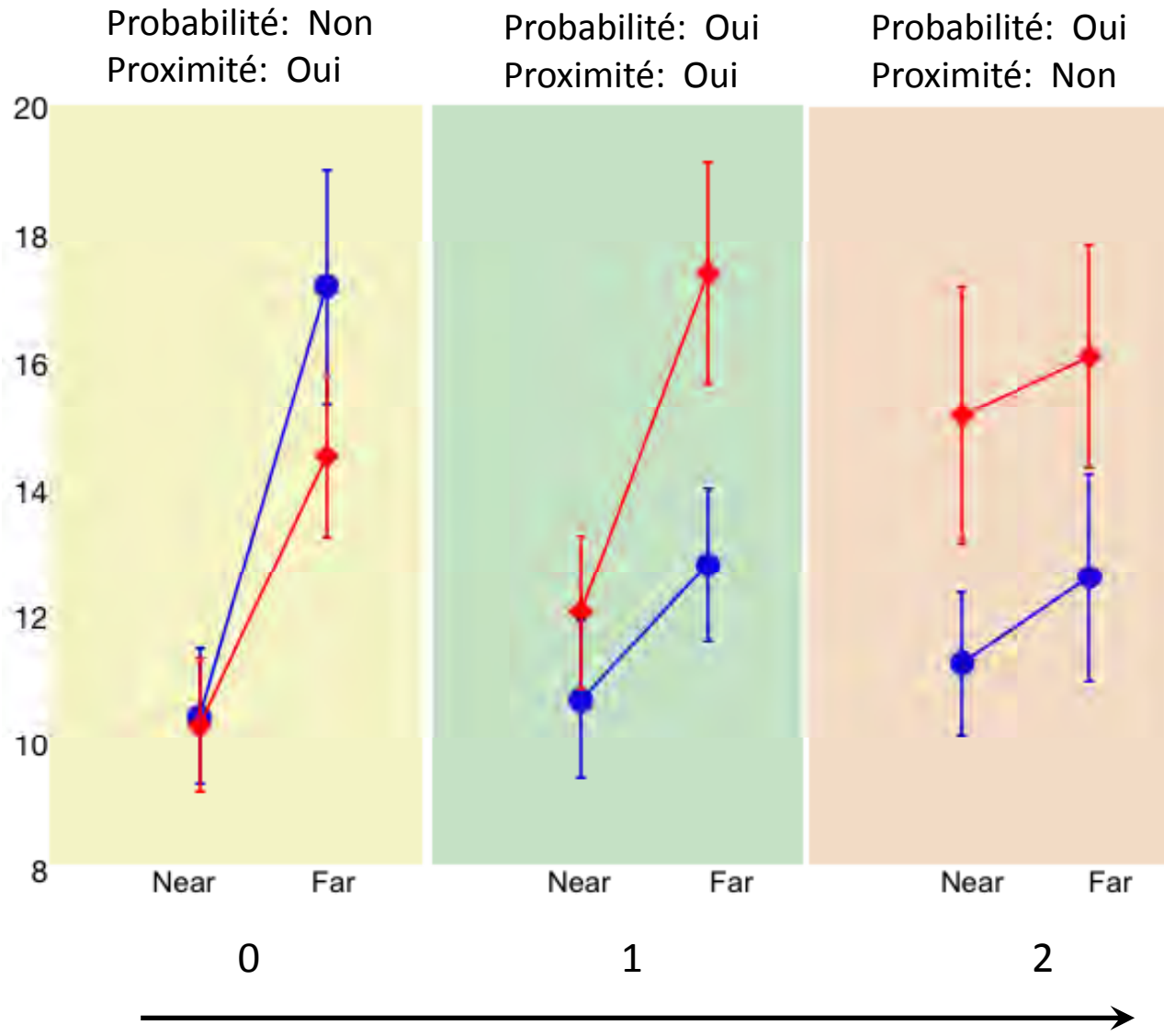


Indices logiques:
Nombres et probabilité



Résultats:

Les enfants pondèrent proximité et probabilité, en fonction du temps écoulé.



objet improbable

1-class object

objet probable

3-class object

temps d'occlusion
(secondes)

Un modèle bayésien du raisonnement de l'enfant

Le modèle stipule que l'enfant génère des prédictions probabilistes fondées sur une simulation mentale.

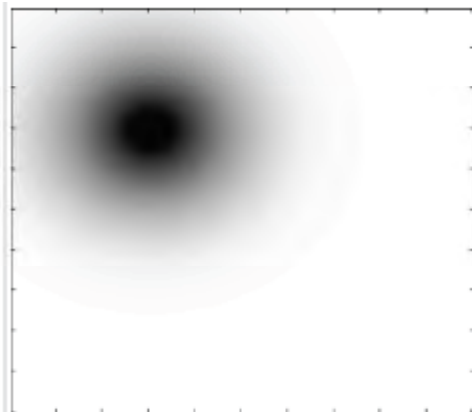
Celle-ci suit des principes physiques (Spelke):

- Principes de **solidité**

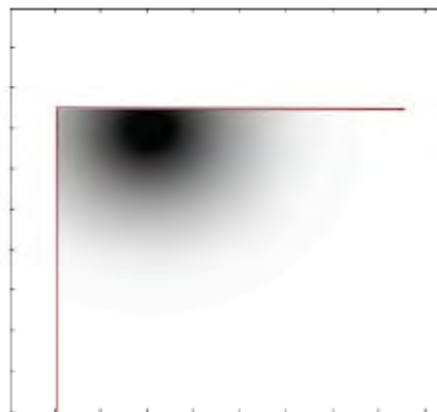
- Principe de **continuité spatio-temporelle**:

$P(S_t | S_{t-1})$.

- Mais aussi une règle probabiliste: simulation par un mouvement Brownien.



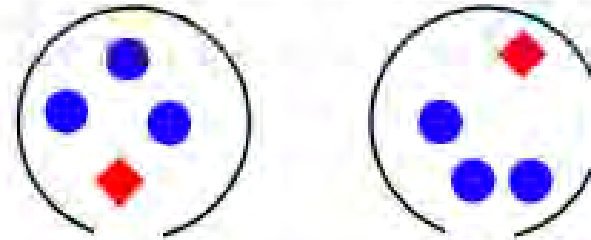
Unconstrained Gaussian



Constrained Gaussian

-et un mécanisme de raisonnement bayésien:

Pre-occlusion configurations



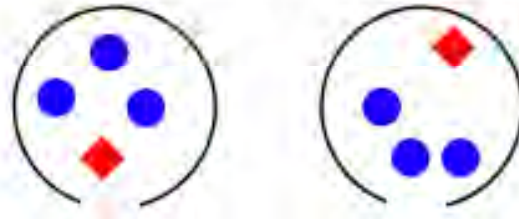
Formally, the probability of a final outcome D_F given the observed data $D_0, \dots, D_{F-1}$ is approximated as a sum of the scores of K hypothetical trajectories (sequences of states $S_0, \dots, S_F$),

$$P(D_F | D_0, \dots, D_{F-1}) \propto \sum_{k=1}^K P(D_F | S_F^k) \times \prod_{t=1}^F P(D_{t-1} | S_{t-1}^k) P(S_t^k | S_{t-1}^k) \quad (1)$$

where the score is a product over time steps t of how well the k th hypothesis fits the observed data $P(D_t | S_t^k)$ and how probable it is under the prior on object dynamics $P(S_t^k | S_{t-1}^k)$. Intuitively, an observed outcome is expected insofar as many predicted future trajectories are consistent with it or unexpected if it is consistent with few predicted trajectories.

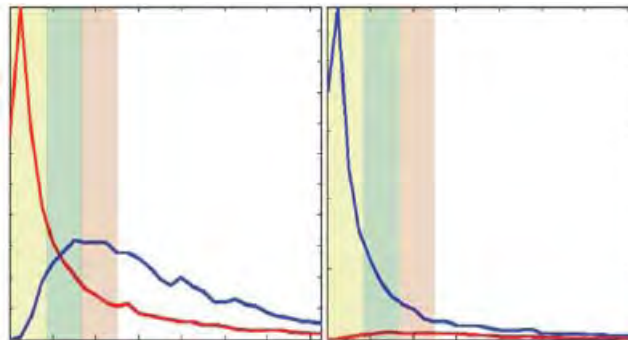
3-class object
1-class object

Pre-occlusion configurations

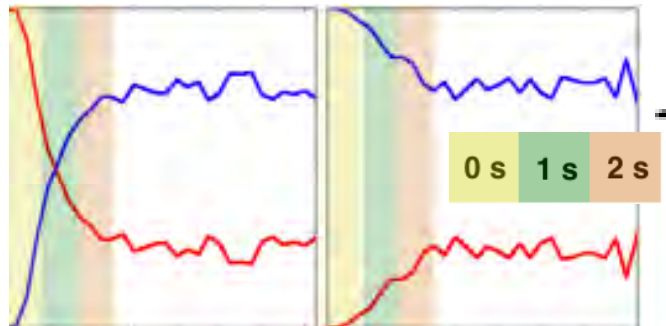


Résultats des simulations

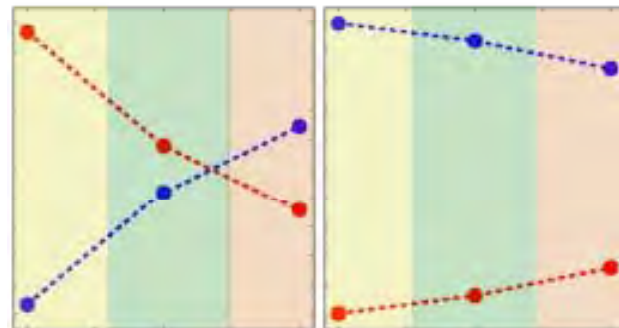
Probabilité qu'un objet rouge ou bleu soit le premier à sortir au temps t



Probabilité conditionnelle: $P(\text{objet} | \text{objet sort au temps } t)$

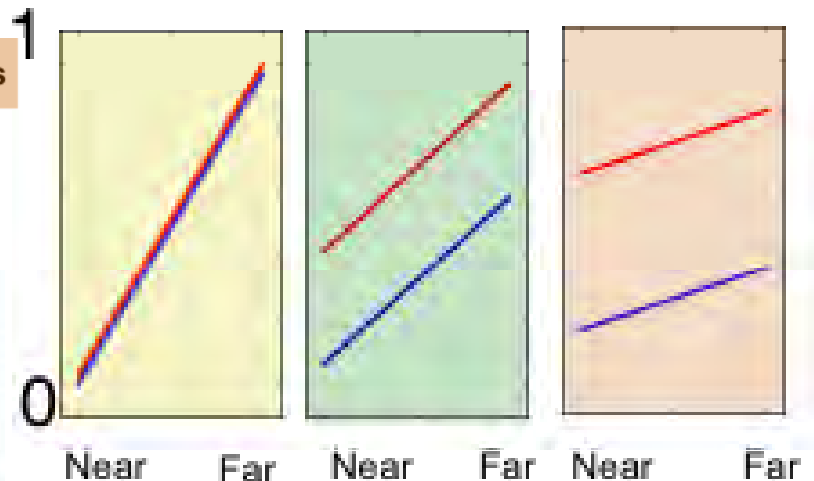


Probabilité conditionnelle aux temps utilisés dans l'expérience (zoom)

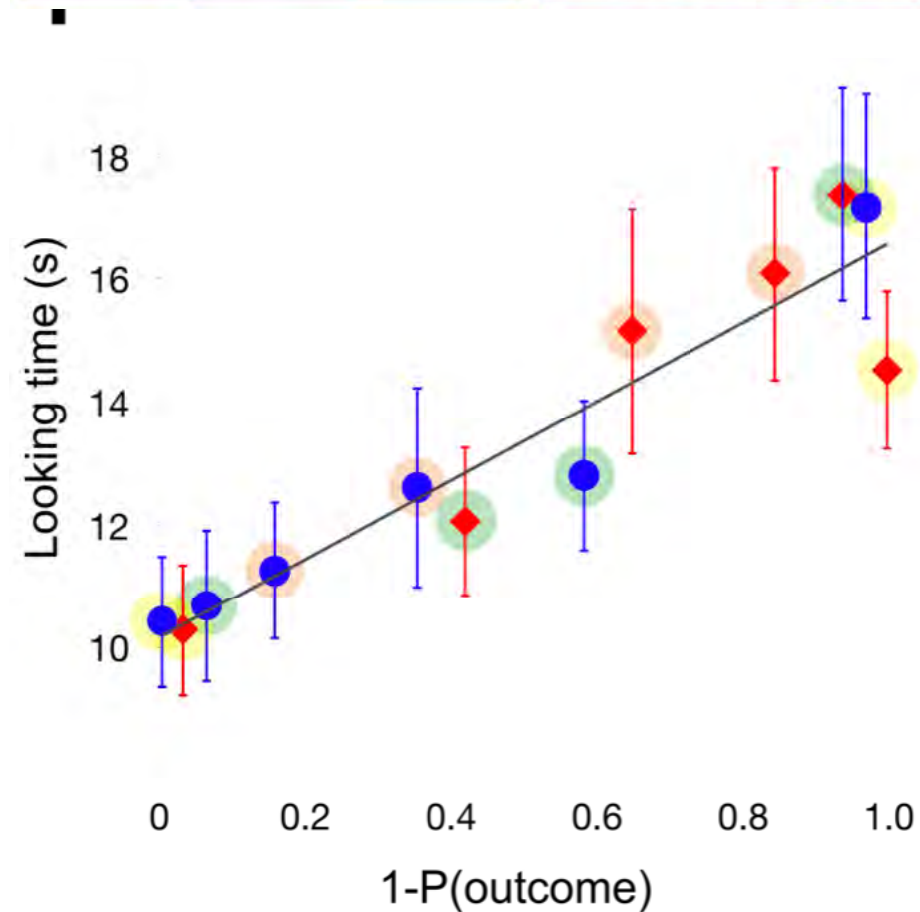


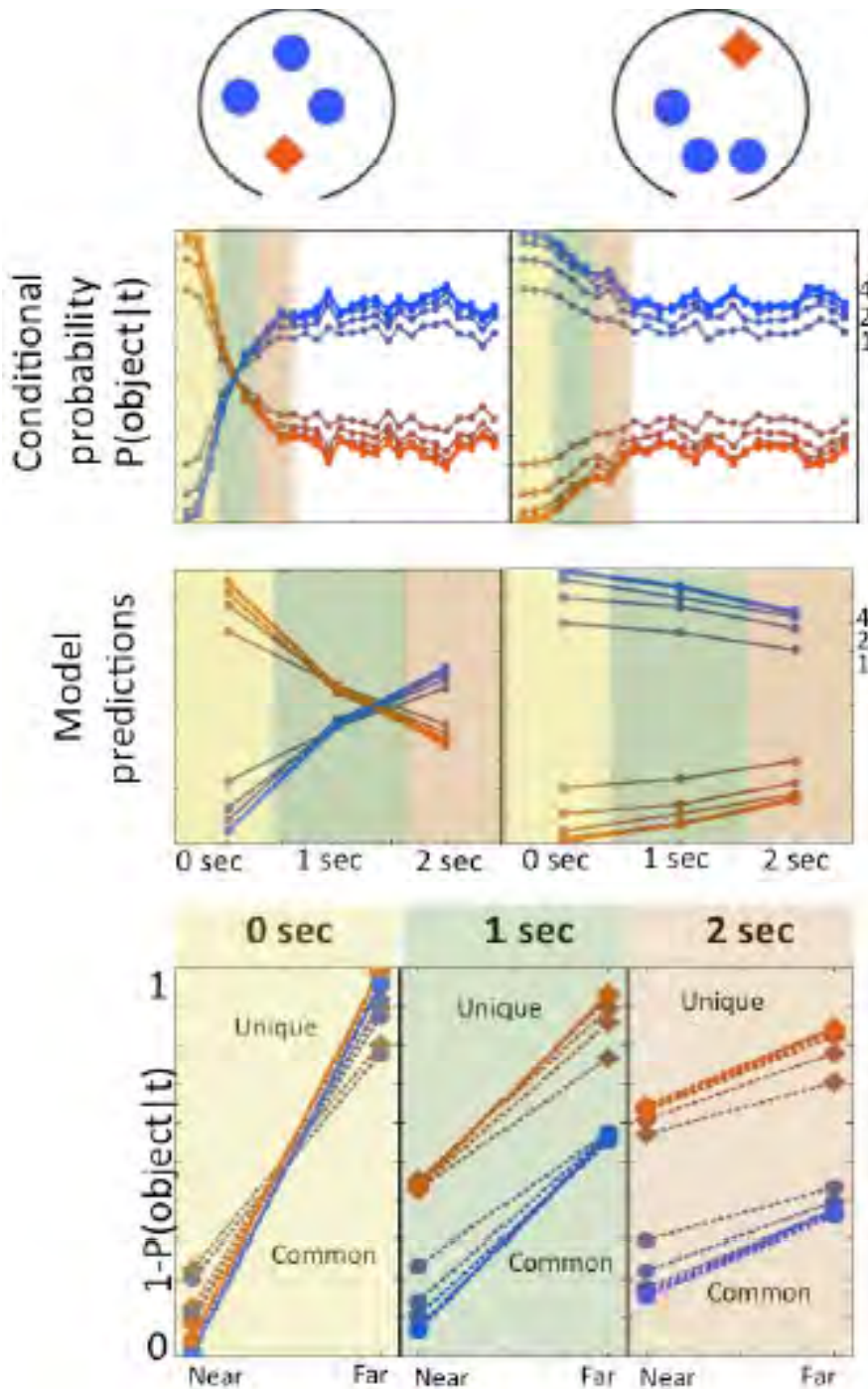
Le modèle génère des courbes de probabilité que ce soit l'objet rouge ou le bleu qui sorte en premier, en fonction du temps.

On en déduit le degré de surprise du bébé = $1 - P(\text{objet} | t)$



Le modèle statistique rend compte, *quantitativement*, de la durée du regard de l'enfant.





Indépendance du nombre de simulations mentales

Pour générer leurs courbes théoriques, Teglas et al. simulent sur ordinateur des milliers de réalisation du mouvement brownien.

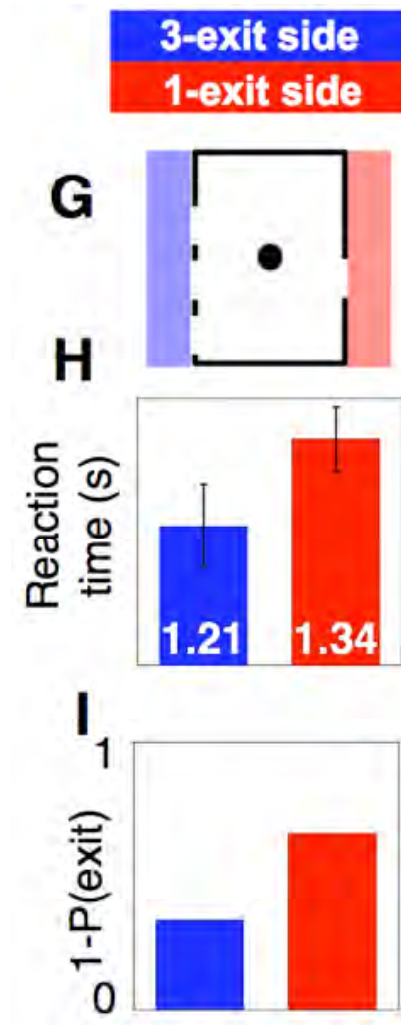
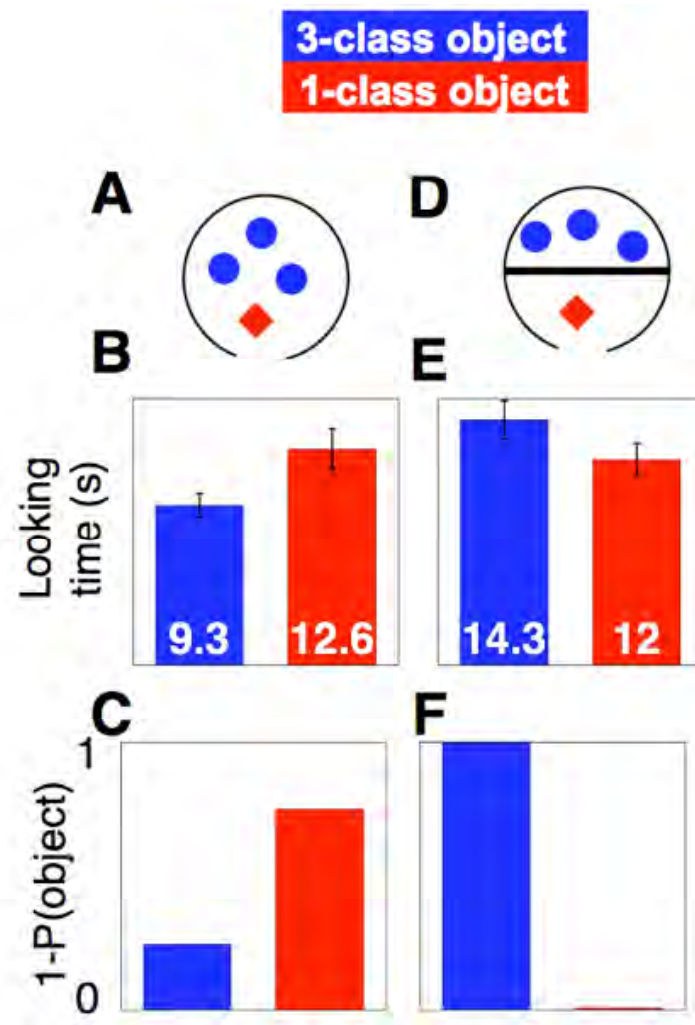
Est-il réaliste de penser qu'au cours d'un essai, un bébé effectue des milliers de simulations mentales?

Peut-être pas... mais les prédictions sont pratiquement inchangées si l'on suppose que l'enfant effectue 4, 2 ou seulement 1 simulation mentale par essai.

Le modèle rend compte de nombreuses autres expériences, sans hypothèses supplémentaires

Simulation par Teglas et al. (*Science*, 2011) des expériences de :

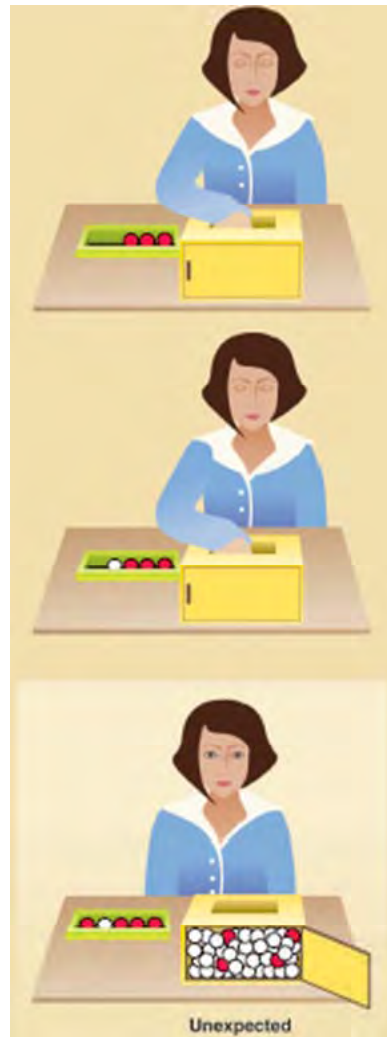
Teglas, E., Girotto, V., Gonzalez, M., & Bonatti, L. L. (2007). Intuitions of probabilities shape expectations about the future at 12 months and beyond. *Proc Natl Acad Sci U S A*, 104(48), 19156-19159.



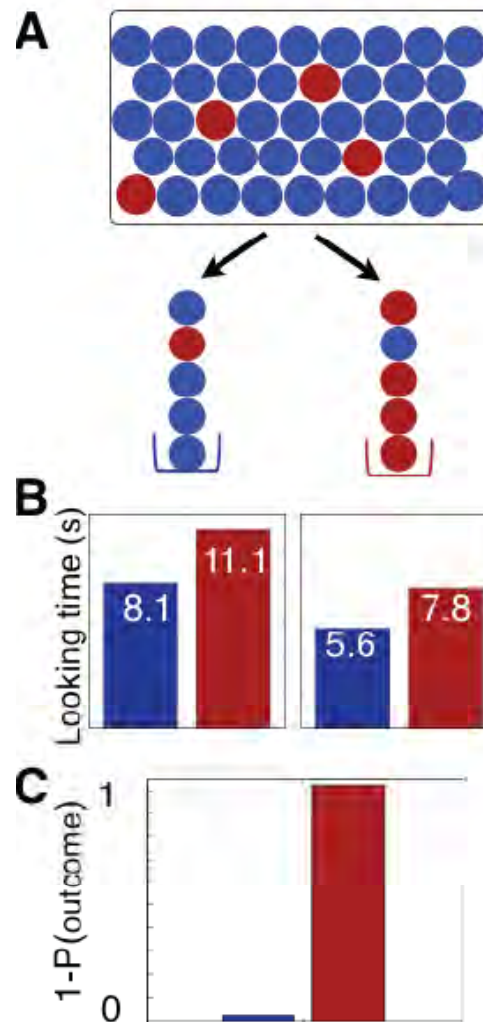
Le tirage d'une urne, considéré comme la simulation d'un modèle interne

Simulation par Teglas et al. (*Science*, 2011) des expériences de :

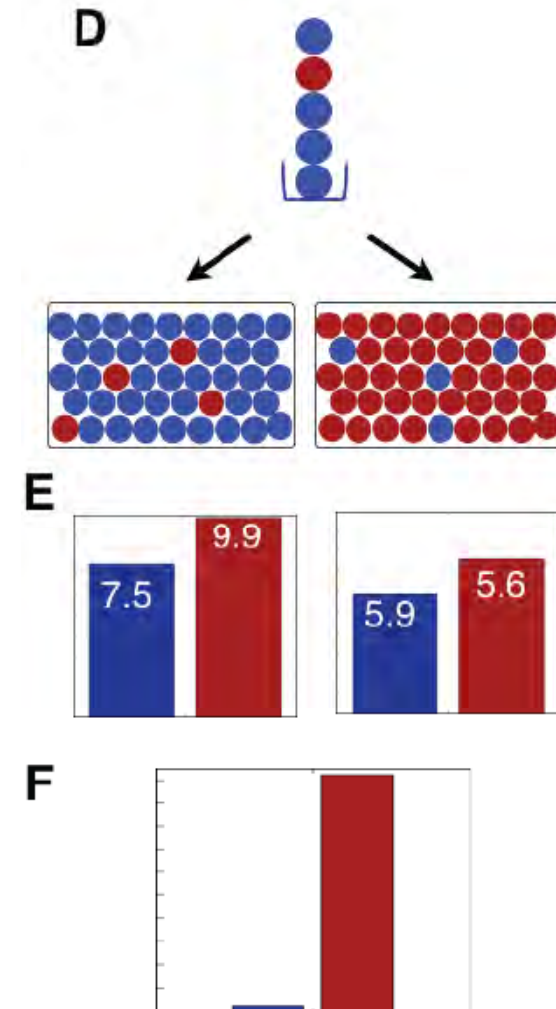
Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proc Natl Acad Sci U S A*, 105(13), 5012-5015.



From Population to Sample



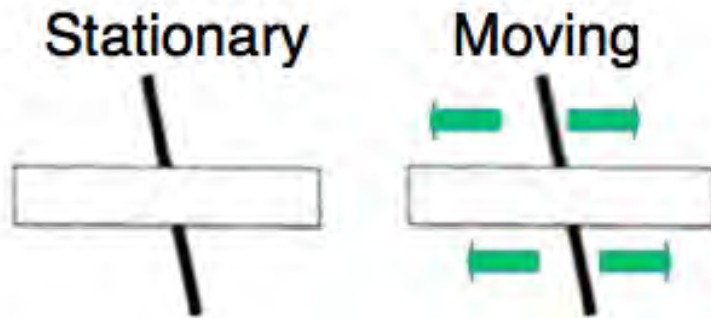
From Sample to Population



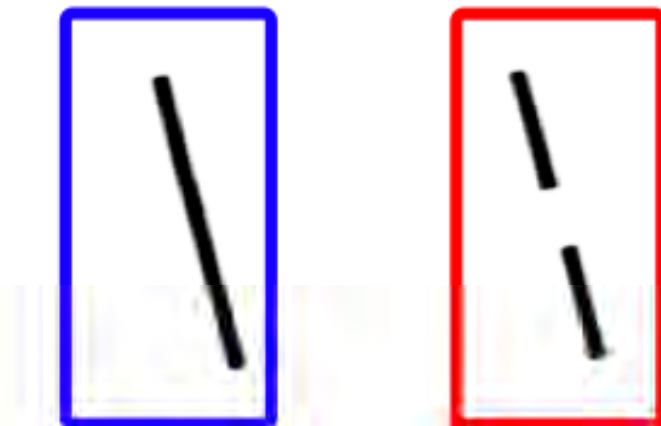
Les expériences portant sur le concept d'objet

Simulation par Teglas et al. (*Science*, 2011) de :
Kellman, P. J., & Spelke, E. S. (1983). Perception of Partly
Occluded Objects in Infancy. *Cognitive Psychology*, 15, 483-524.

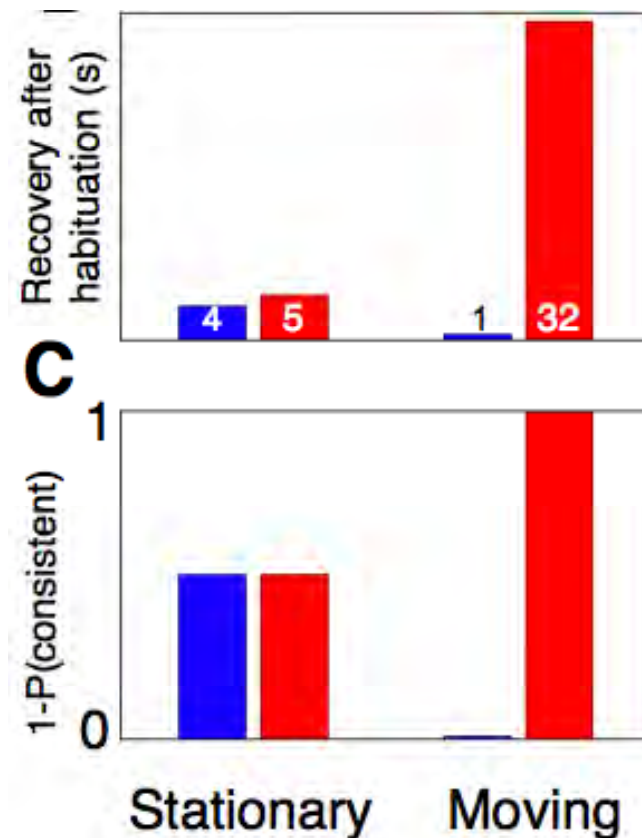
Un objet est présenté à des enfants de 3 mois,
derrière un écran, statique ou en mouvement.



L'écran s'en va et révèle l'une de deux
interprétations possibles:



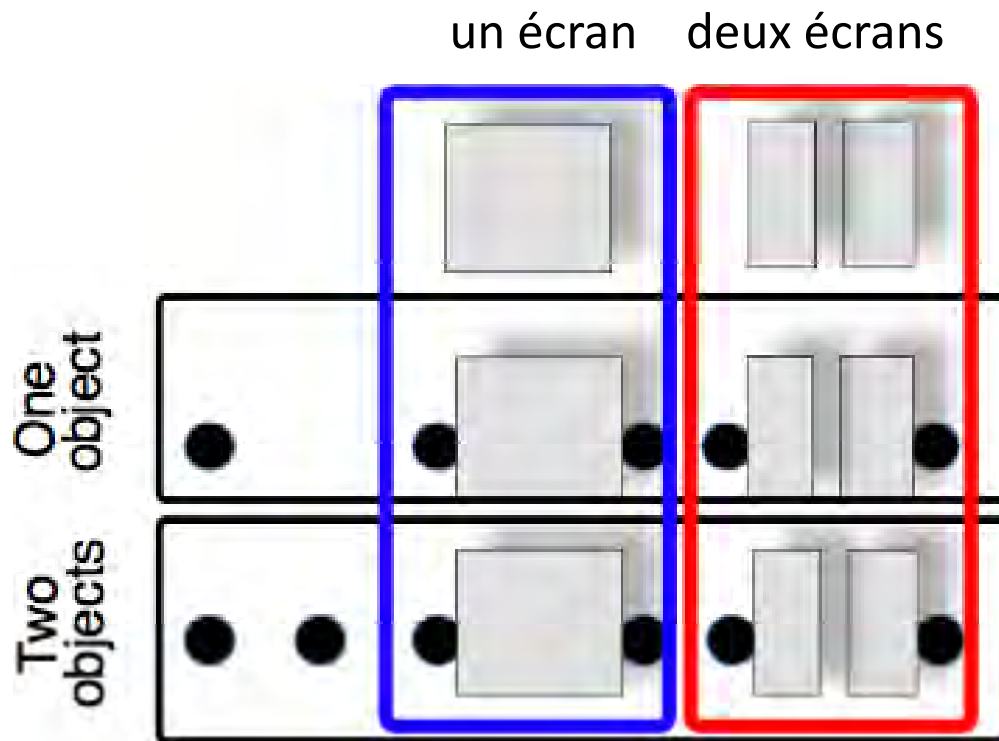
Le modèle considère les deux
hypothèses équiprobables. Mais, dès que
l'objet bouge, la probabilité a posteriori
de l'hypothèse 2 diminue car ce serait
une coïncidence remarquable que les
deux mouvements soient identiques.



Les inférences sur la numérosité, l'espace, et l'occlusion

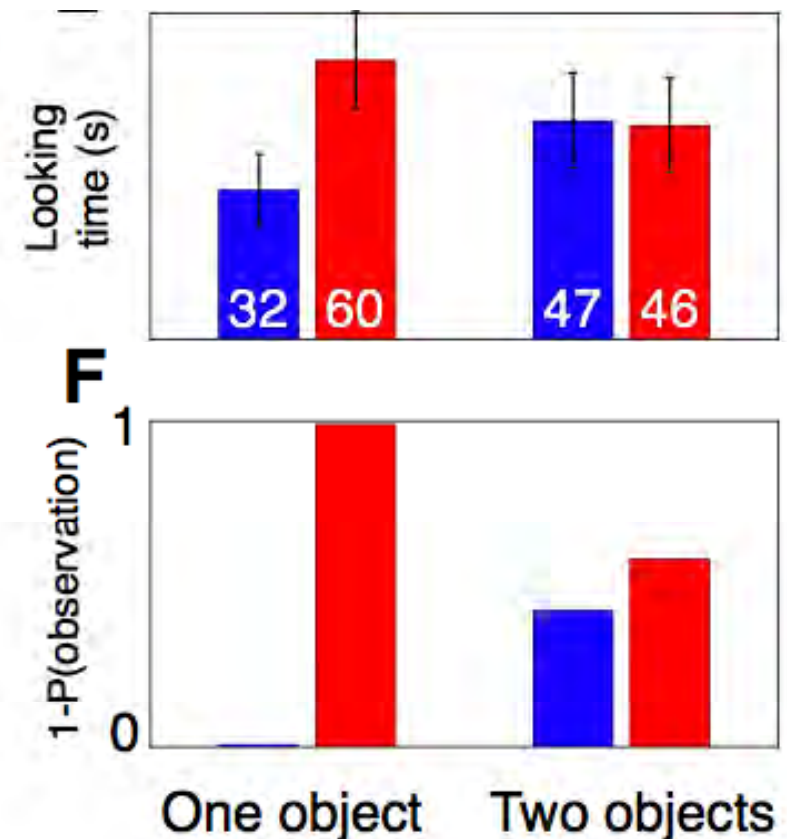
Simulation par Teglas et al. (*Science*, 2011) de :
Aguiar, A., & Baillargeon, R. (1999). 2.5-month-old infants' reasoning about when objects should and should not be occluded. *Cognitive Psychology*, 39(2), 116-157.

Des enfants de 2.5 mois voient un ou deux objets se cacher derrière un ou deux écran. Puis un objet sort tantôt à droite et tantôt à gauche.



Le modèle échantillonne toutes les trajectoires compatibles avec la sortie d'un objet à droite et à gauche.

Quand il y a deux écrans, il est très peu probable que l'objet passe très vite d'un côté à l'autre, sans être vu.

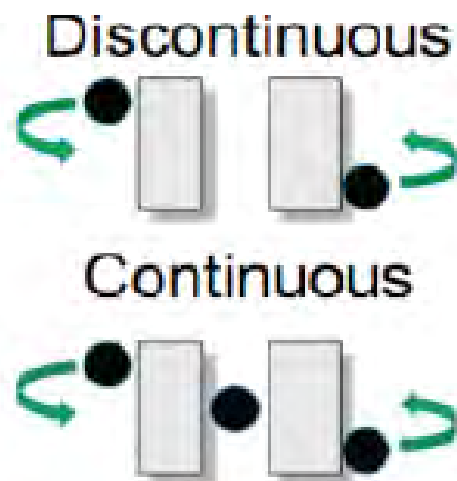


L'inférence inverse: du mouvement au nombre d'objets

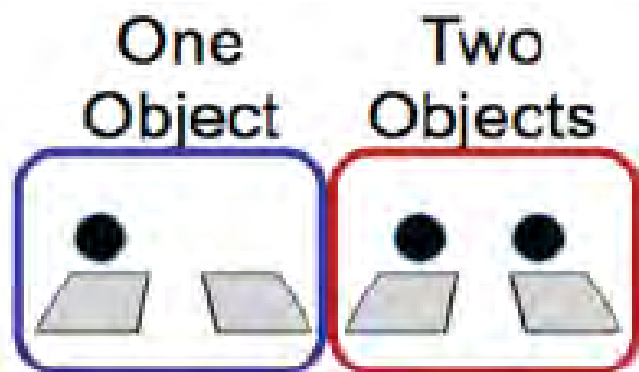
Simulation par Teglas et al. (*Science*, 2011) de :

Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology*, 30, 111-153.

Des enfants de dix mois voient deux écrans. Un objet en sort de droite et de gauche, tantôt en passant par le milieu, tantôt pas.

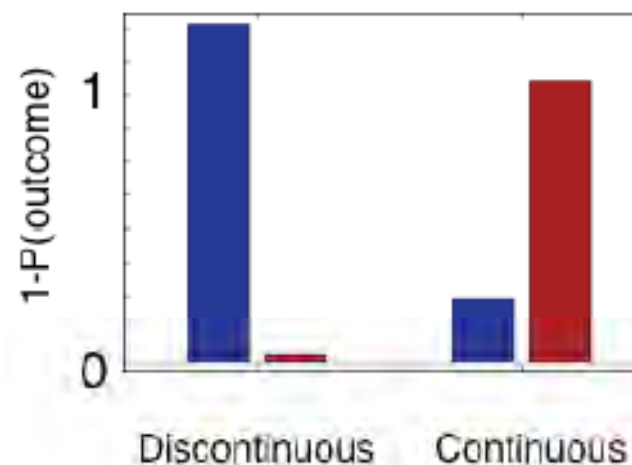
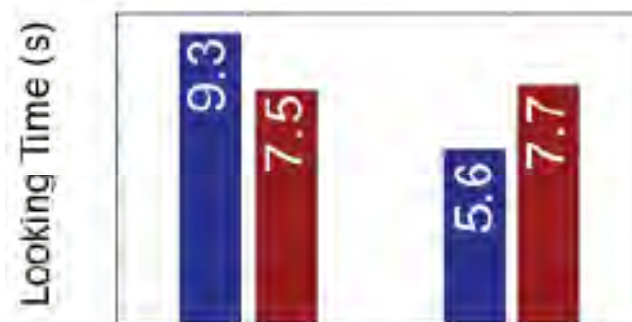


Les écrans s'abaissent et révèlent l'une de deux interprétations possibles:



Le modèle considère les deux hypothèses équiprobables. Mais

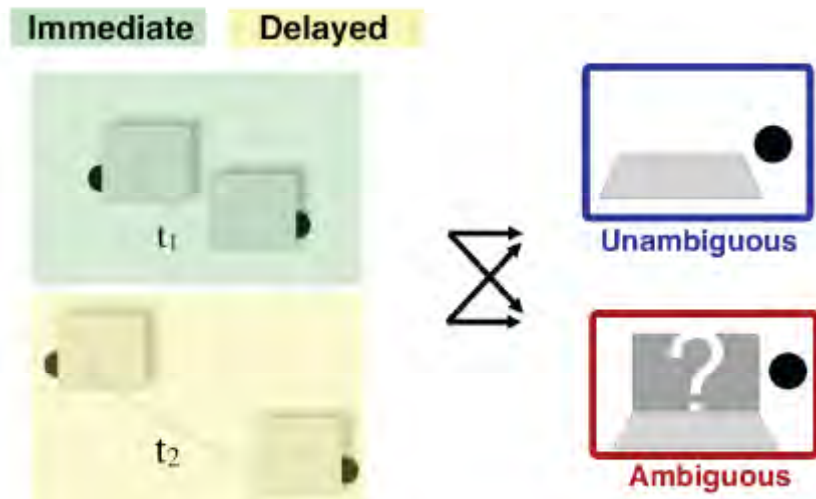
- dans la condition « discontinue », l'hypothèse 2 est plus probable.
- dans la condition « continue », l'hypothèse 1 est plus probable car la 2 prédit l'observation fréquente de 2 objets.



La prise en compte de délais temporels dans la simulation mentale de l'enfant

Simulation par Teglas et al. (*Science*, 2011) de :
Wilcox, T., & Schweinle, A. (2003). Infants' use of speed information to individuate objects in occlusion events.
Infant Behavior & Development, 26(2), 253-282.

Un objet sort alternativement à gauche et à droite d'un écran, avec un délai variable.



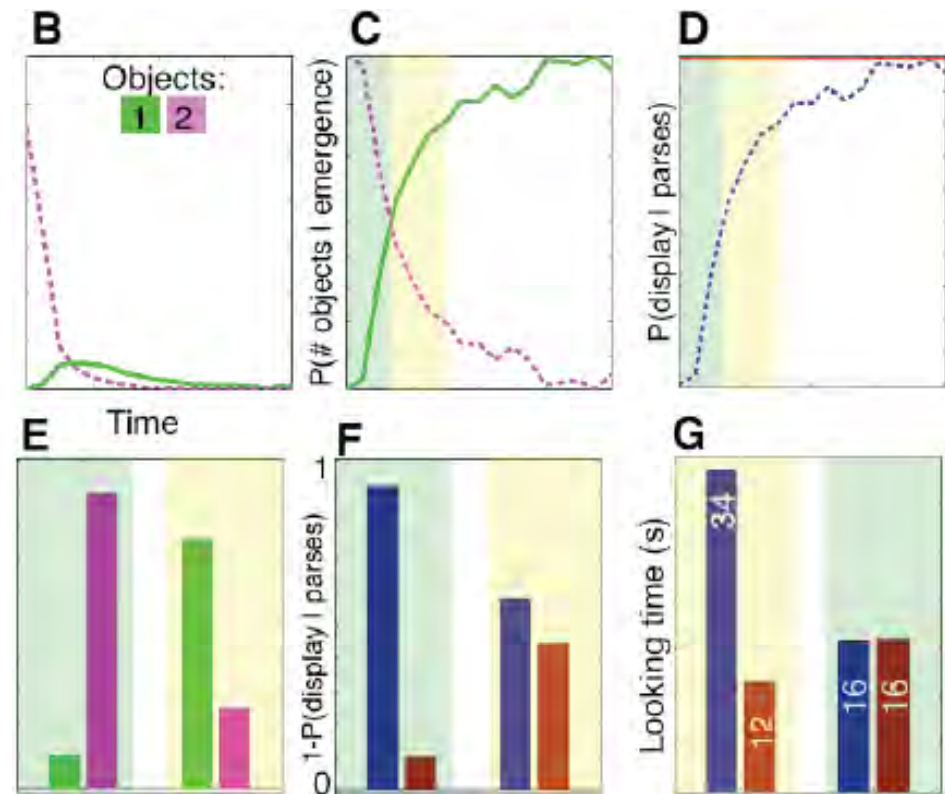
L'écran s'abaisse, et révèle:

- soit une situation non-ambigüe: une seule balle.
- soit une situation ambigüe: une balle, mais également un deuxième écran qui pourrait cacher une seconde balle.

Le modèle prédit que si le délai est court, il est plus probable qu'il y ait deux balles, et s'il est long, qu'il y ait une seule balle.

La situation ambigüe est compatible avec les deux possibilités. Seule la situation non-ambigüe immédiate est improbable.

C'est celle que les enfants regardent longtemps!



Conclusions des expériences de Teglas, Bonatti et coll.

1. Les enfants développent des **modèles du monde extérieur** qui combinent
 - des **principes physiques** (solidité, continuité spatio-temporelle)
 - et des **raisonnements statistiques**

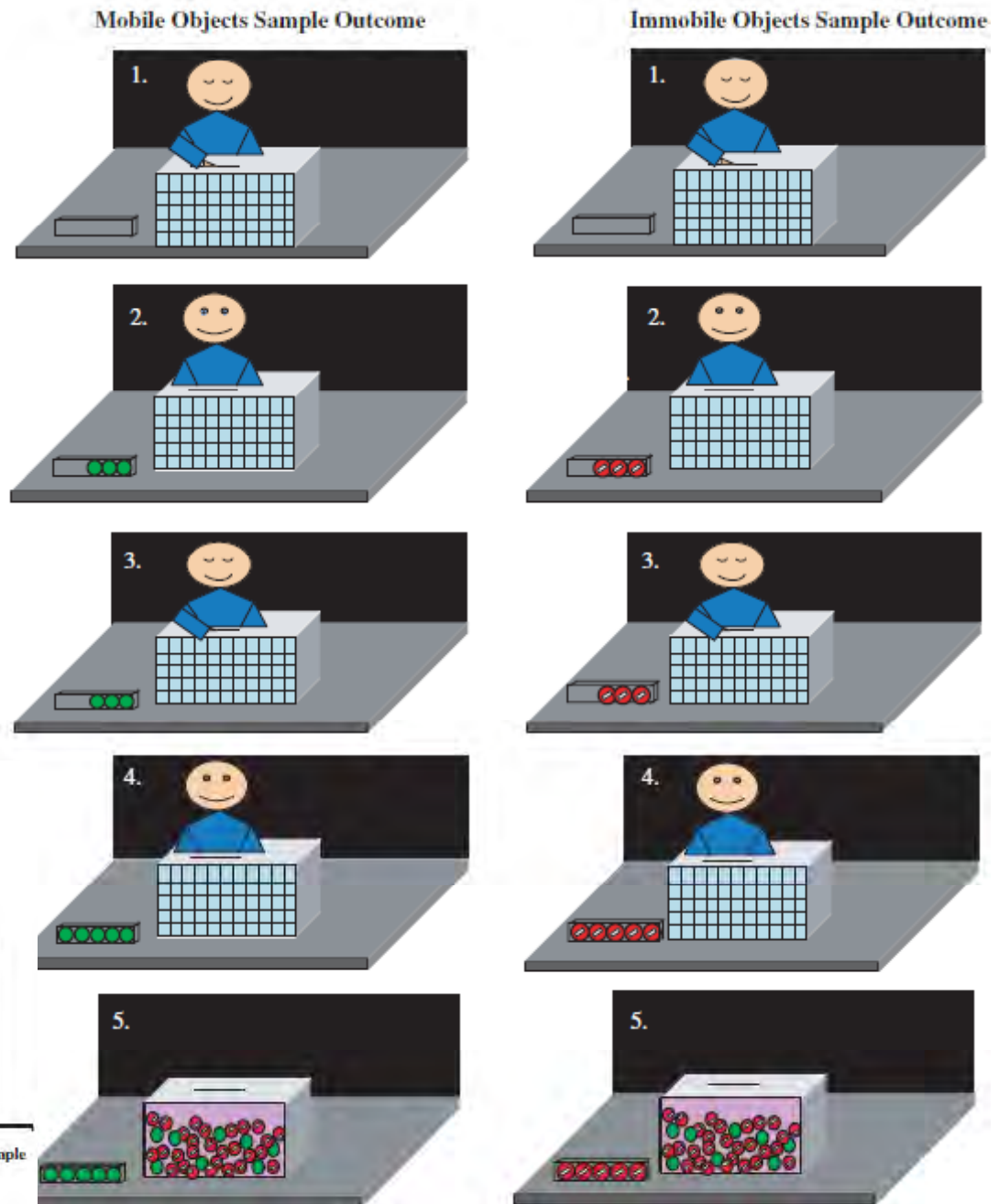
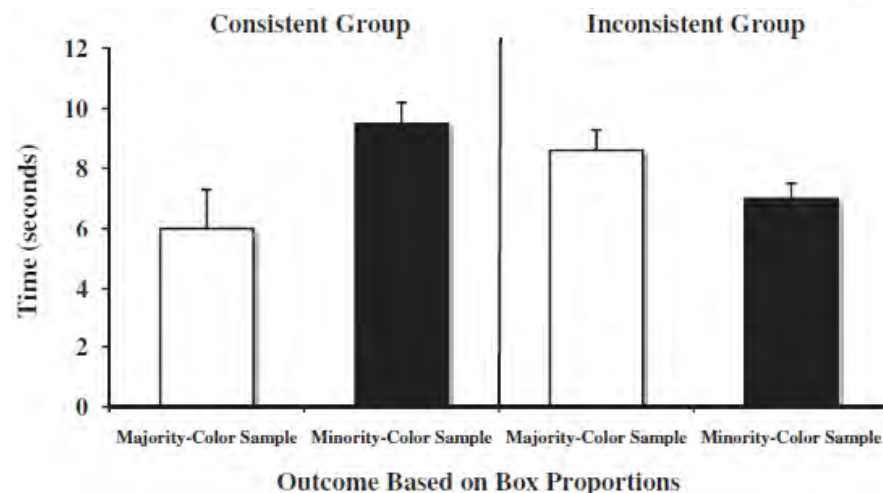
Bien des expériences que nous pensions porter uniquement sur le « sens physique » des enfants portent, en fait, également sur leur sens des probabilités.

Un autre exemple de combinaison de la physique et des probabilités

Denison, S., & Xu, F. (2010). Integrating physical constraints in statistical inference by 11-month-old infants. *Cogn Sci*, 34(5), 885-908.

Denison et Xu (2010) répliquent, chez le bébé de 11 mois, leur expérience d'échantillonnage, avec une variante cruciale: les enfants apprennent que les balles rouges sont fixées avec du velcro, on ne peut pas les retirer.

Les résultats changent radicalement: les enfants sont surpris de voir tirer des objets fixés, quelle que soit leur représentation dans l'urne.



Conclusions des expériences de Teglas, Bonatti et coll.

1. Les enfants développent des modèles du monde extérieur qui combinent

- des **principes physiques** (solidité, continuité spatio-temporelle)
- et des **raisonnements statistiques**

Bien des expériences que nous pensions porter uniquement sur le « sens physique » des enfants portent, en fait, également sur leur sens des probabilités

2. Le regard de l'enfant **reflète, en continu, son appréciation** du monde extérieur.

- La durée du regard peut être interprétée comme $1 - p$ (ou peut-être serait plus approprié d'utiliser $\log((1-p)/p)$?)

3. Le cerveau **projette en permanence des interprétations** sur le monde extérieur

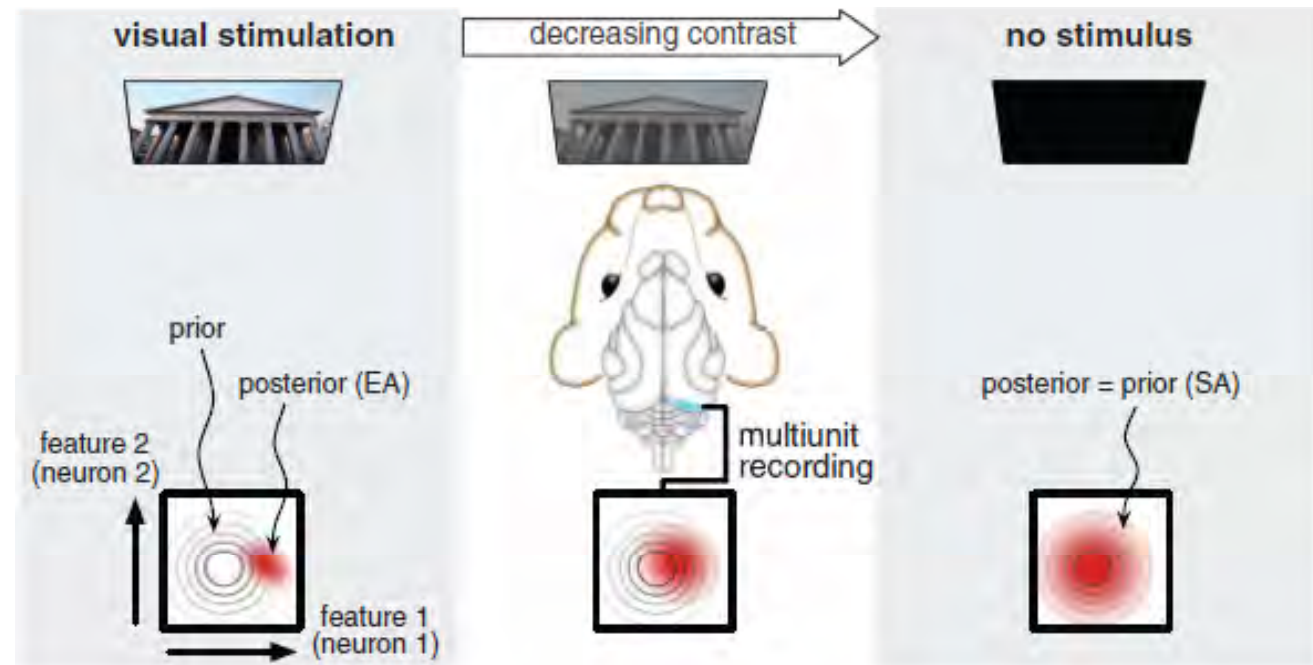
- L'activité « spontanée » du cerveau pourrait-elle refléter cette production permanente de modèles, cet échantillonnage constant des modèles plausibles?

L'activité corticale spontanée considérée comme un modèle interne de l'environnement

Berkes, P., Orban, G., Lengyel, M., & Fiser, J. (2011). Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science*, 331(6013), 83-87.

La perception repose sur un modèle interne de l'environnement.

La mise en relation du modèle interne avec les entrées sensorielles peut être modélisée comme une **inférence** bayésienne (cf cours 2012):



Le modèle fournit l' *a priori* qui est combiné avec les données et conduit à une distribution de probabilité *a posteriori* des états du monde extérieur.

Prédiction du modèle: en l'absence d'entrées sensorielles, le système nerveux doit générer (en permanence?) des *a priori* sur les entrées qu'il pourrait recevoir.

→ modèle possible de l'activité spontanée omniprésente dans le système nerveux?

L'activité corticale spontanée considérée comme un modèle interne de l'environnement

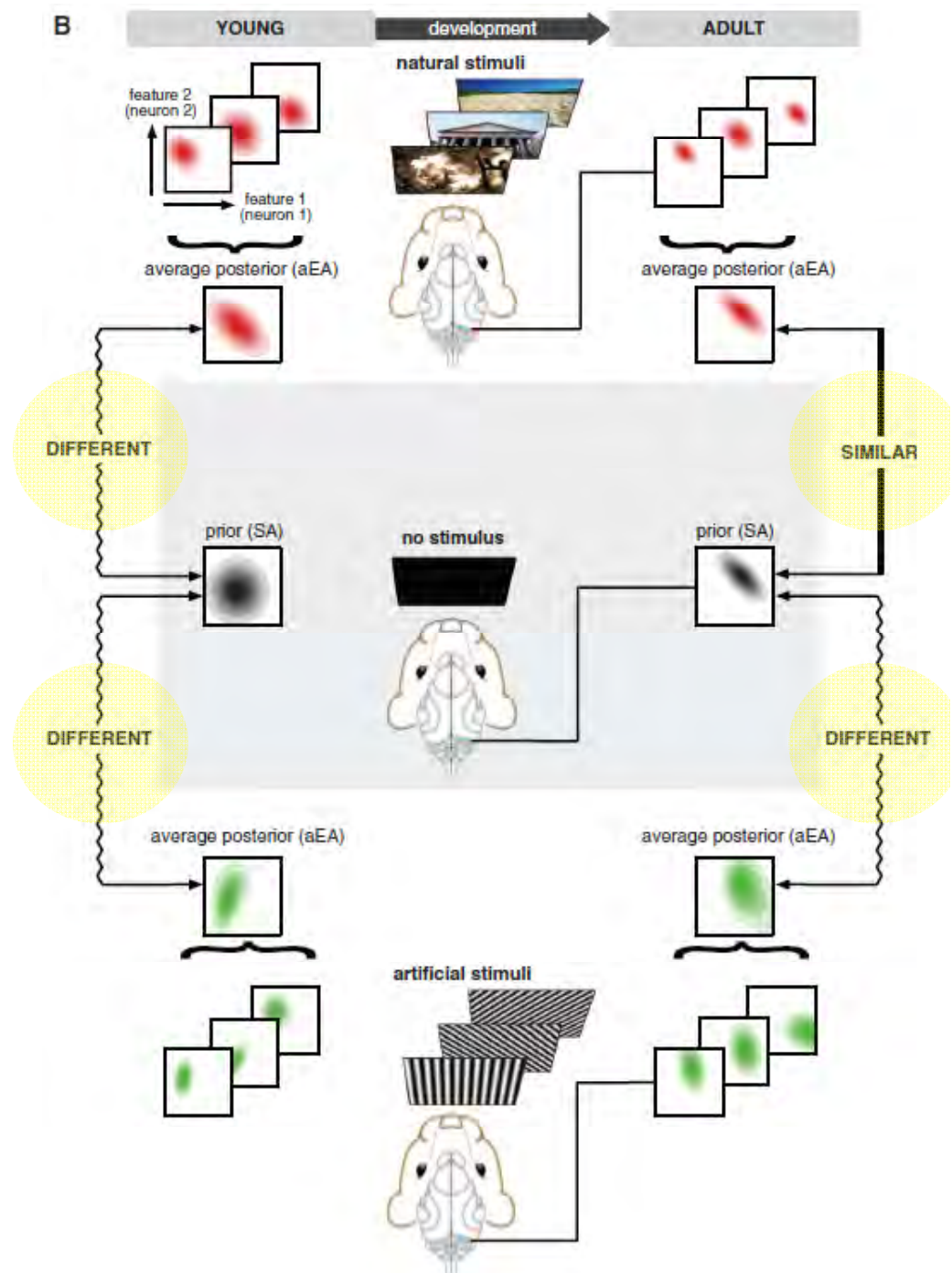
Berkes, P., Orban, G., Lengyel, M., & Fiser, J. (2011). Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science*, 331(6013), 83-87.

Selon ce modèle, la justesse de la perception dépend de l'adéquation de ce modèle interne avec les propriétés statistiques du monde extérieur.

Au cours du développement, l'activité spontanée devrait s'ajuster progressivement aux fréquences réelles des traits du monde extérieur.

Dans l'expérience proposée, deux variables sont croisées:

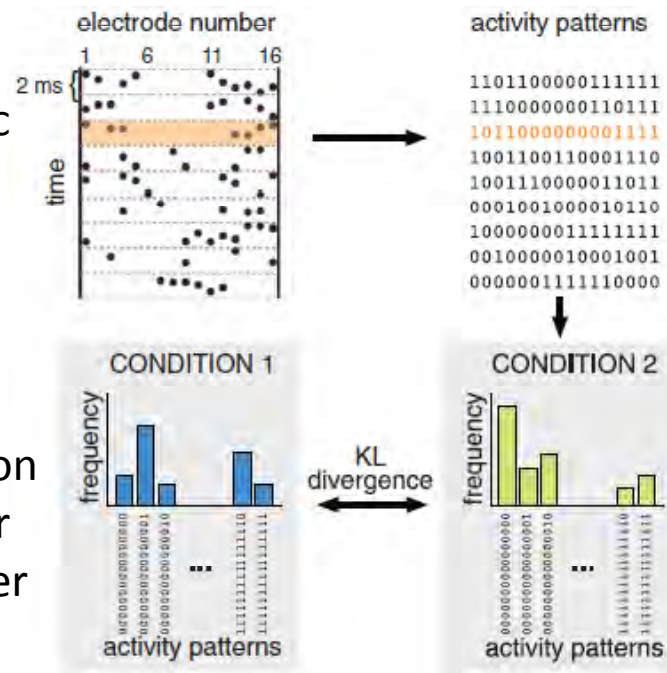
- l'âge des animaux (furets)
- le type de stimuli auxquels ils sont exposés: scènes naturelles, rien, ou stimuli artificiels.



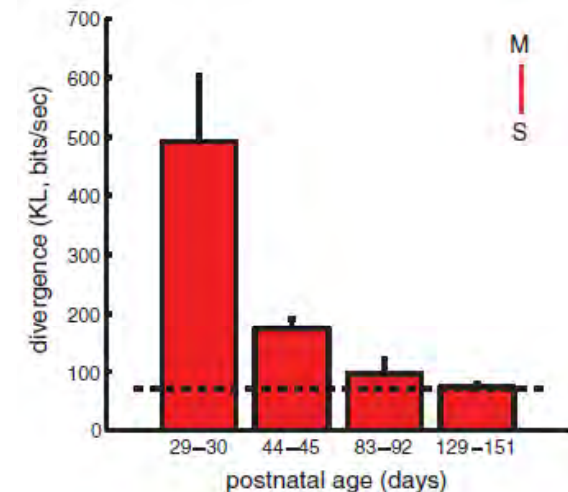
L'activité spontanée converge vers l'activité évoquée par un film

Enregistrement de l'activité (évoquée ou spontanée) avec 16 électrodes dans le cortex visuel

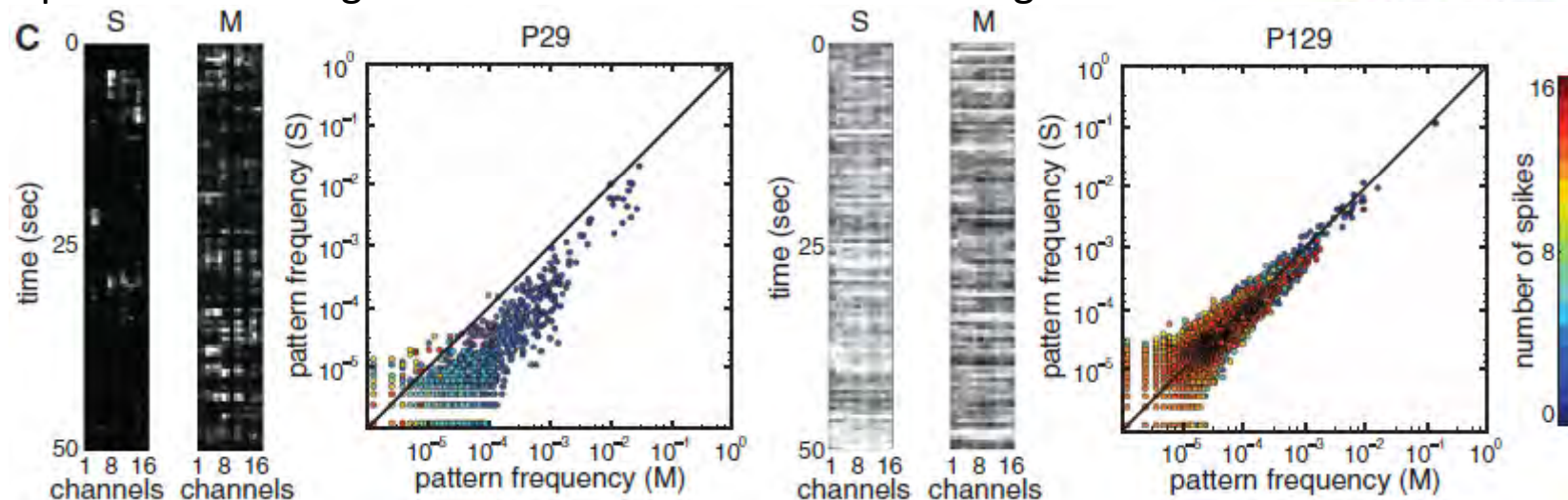
Conversion en informations digitales, calcul d'un histogramme des configurations, et comparaison de deux distributions par leur divergence de Kullback-Leibler



La divergence entre l'activité spontanée et celle évoquée par un film diminue avec l'âge.



La fréquence des configurations neuronales évolue avec l'âge:



L'activité spontanée: un modèle spatio-temporel de l'activité évoquée

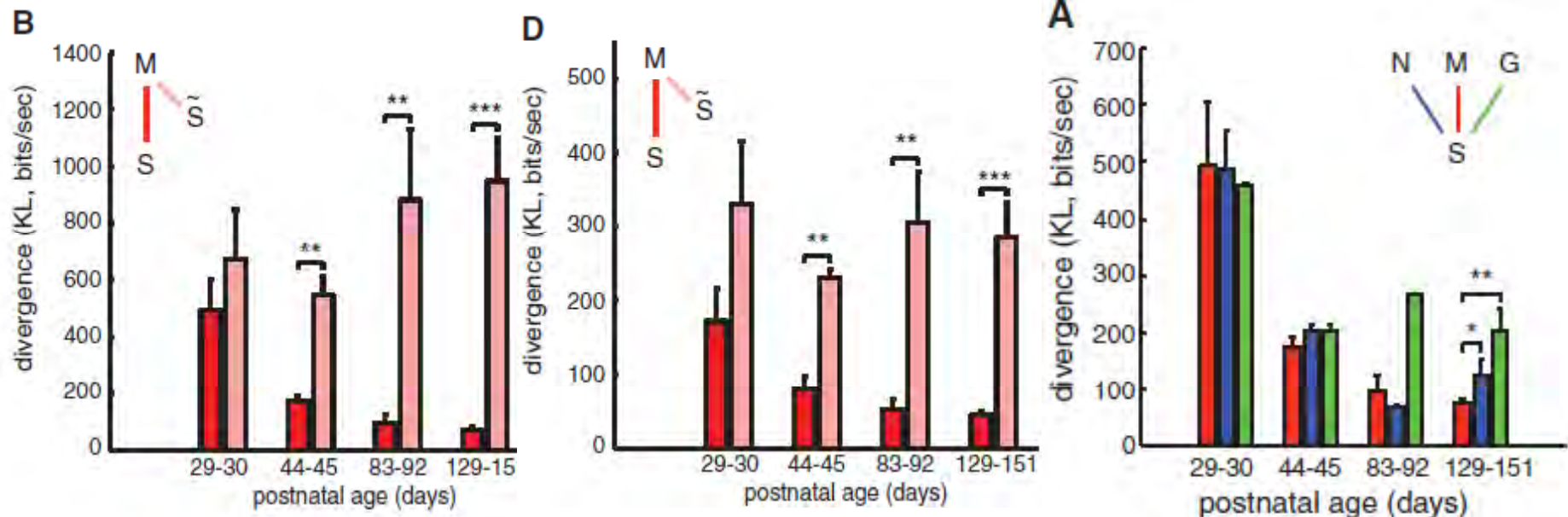
Berkes, P., Orban, G., Lengyel, M., & Fiser, J. (2011). Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science*, 331(6013), 83-87.

La ressemblance croissante de l'activité spontanée avec l'activité évoquée par un film (en rouge) disparaît lorsqu'on perturbe les enregistrements en randomisant, soit leur organisation spatiale (B), soit leur organisation temporelle (D).

Les corrélations entre neurones augmentent avec l'âge, et l'activité neuronale ne devient pas plus éparsée au fil du temps.

Les transitions temporelles dans l'activité neuronale, elles-aussi, se mettent à ressembler à celles évoquées par le film.

La ressemblance croissante n'est pas observée lorsqu'on compare l'activité spontanée à celle évoquée par du bruit (N) ou des grilles (G).



Conclusions

Le cerveau du nouveau-né construit des modèles mentaux hiérarchiques du monde extérieur.

Dès la première année de vie, ces modèles sont très sophistiqués et incluent une combinaison de contraintes issues de la physique et des probabilités. On peut légitimement parler d'un « sens des probabilités » chez l'enfant de quelques mois.

L'examen de l'activité corticale spontanée pourrait fournir une nouvelle manière d'évaluer les modèles mentaux de l'enfant.

