

L'IA peut-elle jouer à la science ?

Le défi d'Eleusis face aux LLMs

Anima Sciences Tom François, 27 février 2026

Présentation d'un article de blog proposé par David Louapre (chaîne Youtube *Science Etonnante*).

David Louapre (2026). "Can LLMs Play the Game of Science?".

Plan

- Eleusis : le jeu de cartes
- Comment les LLMs jouent à Eleusis ?
 - méthode
 - résultats
- Discussion

Jeu de cartes Éleusis

Jeu inventé en 1956 par [Robert Abbot](#), popularisé en France grâce à Bernard Werber et *[L'Encyclopédie du savoir relatif et absolu](#)*

Jeu de cartes Éleusis

Jeu inventé en 1956 par [Robert Abbot](#), popularisé en France grâce à Bernard Werber et *[L'Encyclopédie du savoir relatif et absolu](#)*

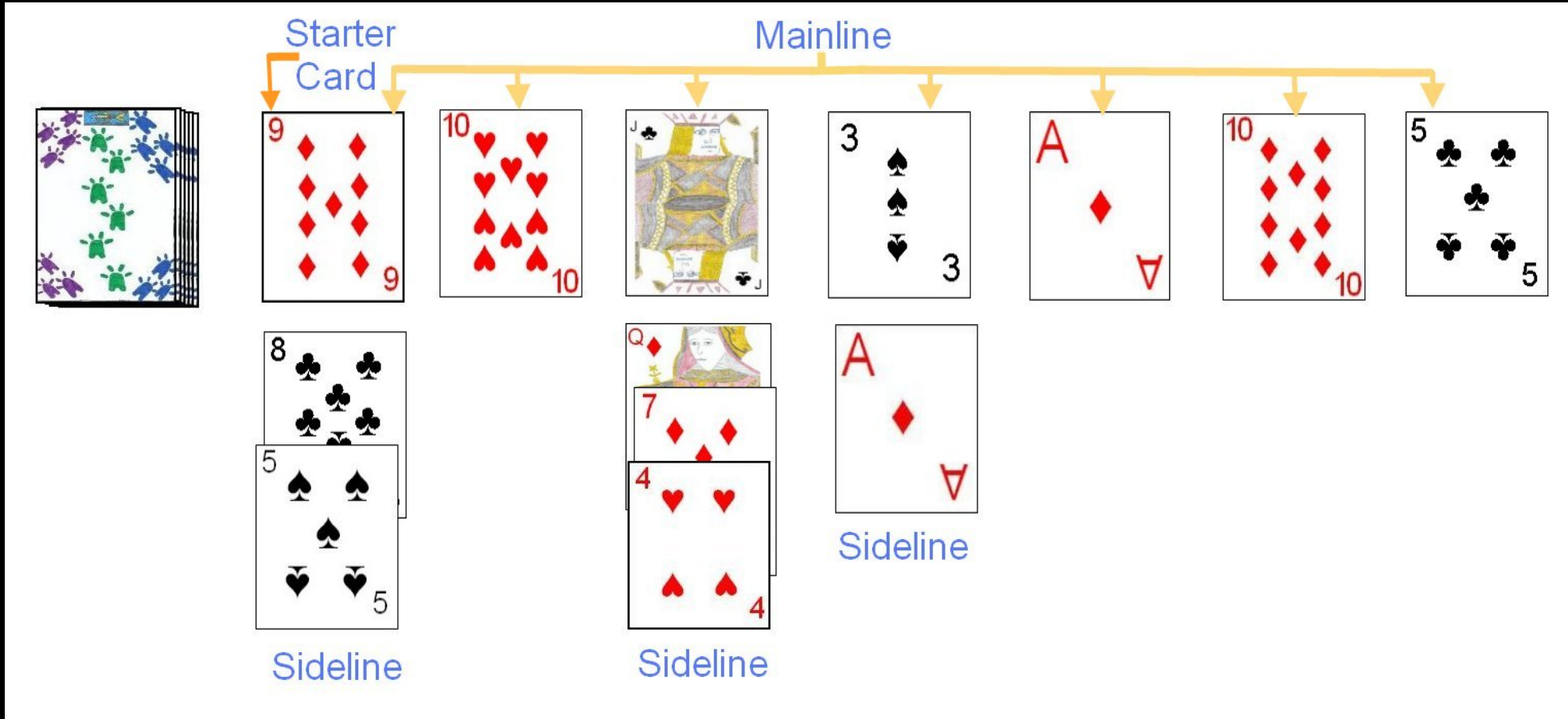
Rôles : un "Dieu", des "chercheurs".

But : trouver la règle de placement inventée par un joueur

Exemple de règles

- les cartes doivent successivement être rouges ou noires
- on joue dans l'ordre croissant

Exemple d'une partie en cours



Pourquoi ce jeu pour tester les LLMs?

C'est un jeu qui simule bien l'approche scientifique [1, Romesburg, 1979].

- observation
- analyse
- confiance

Trouver le bon compromis entre observer l'ensemble des possibles et jouer toutes les cartes et faire des prédictions erronées.

Méthode pour faire jouer l'IA

- liste de règles
- liste de modèles LLMs
- un prompt commun aux modèles

On fait jouer chaque modèle avec 3 initialisations différentes 30 tours par règle.

Il utilise un ensemble de règles pré-établies séparées en différentes catégories

Catégorie de règles	Exemples
Statique	"seulement les rouges" ; "uniquement les têtes"
Combiné	"seulement les cartes coeur inférieur à 7"
Sequentielle	"la carte doit être plus grande que la précédente"
Cyclique	"alterner entre pair et impair"
Conditions complexes	"écart entre 2 cartes doit être de 2 exactement"

Format de la réponse de chaque LLM

```
{  
  "reasoning_summary": "To test if the rule depends on rank, I play  
a 4♣ (same rank as the starter 4♠) hoping to see if same-rank cards  
are accepted.",  
  "card": "4♣",  
  "tentative_rule": "A card is accepted if it has the same rank as  
the previous mainline card.",  
  "confidence_level": 3,  
  "guess_rule": false  
}
```

Calcul du score

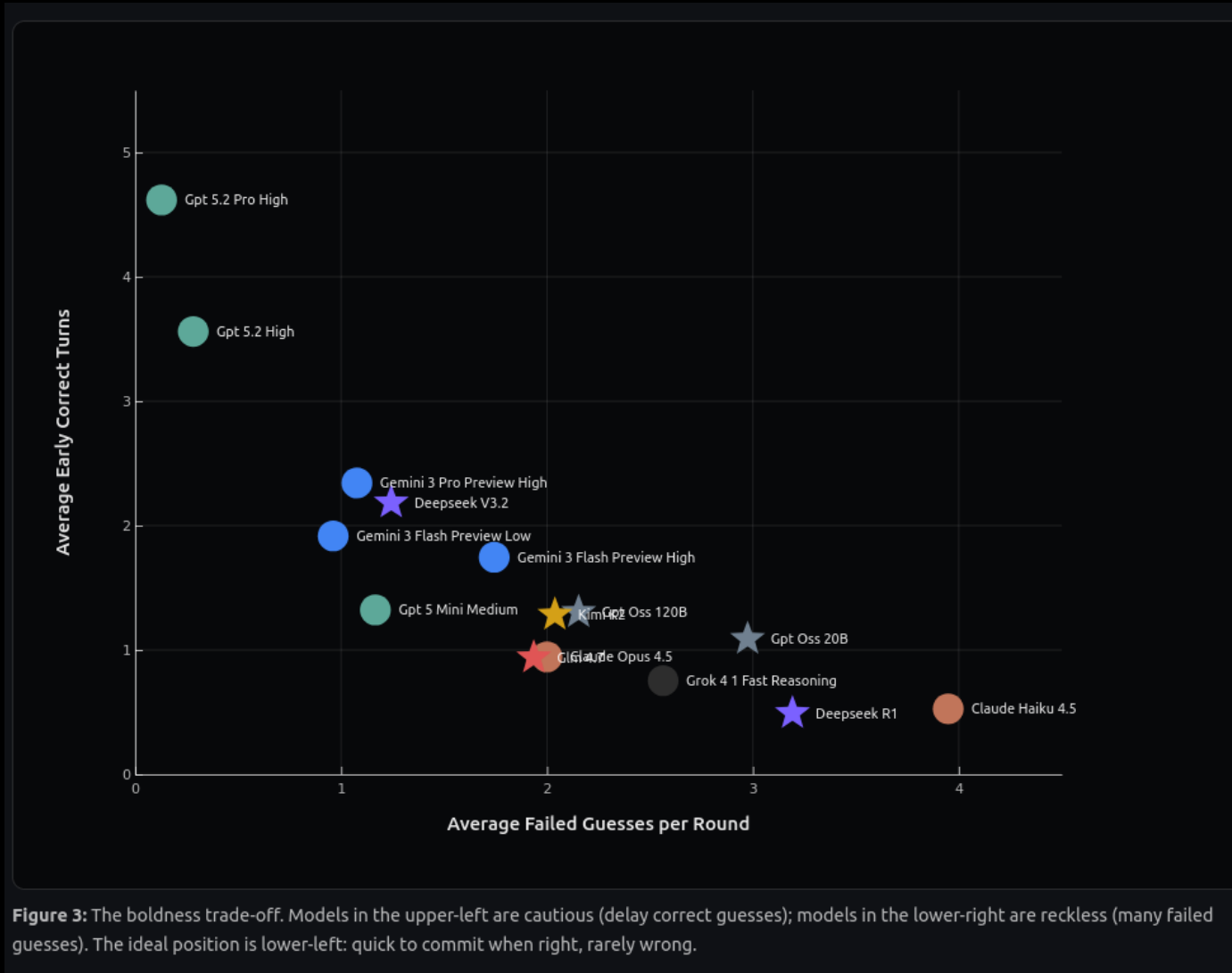
Le nombre maximum de tour est 30

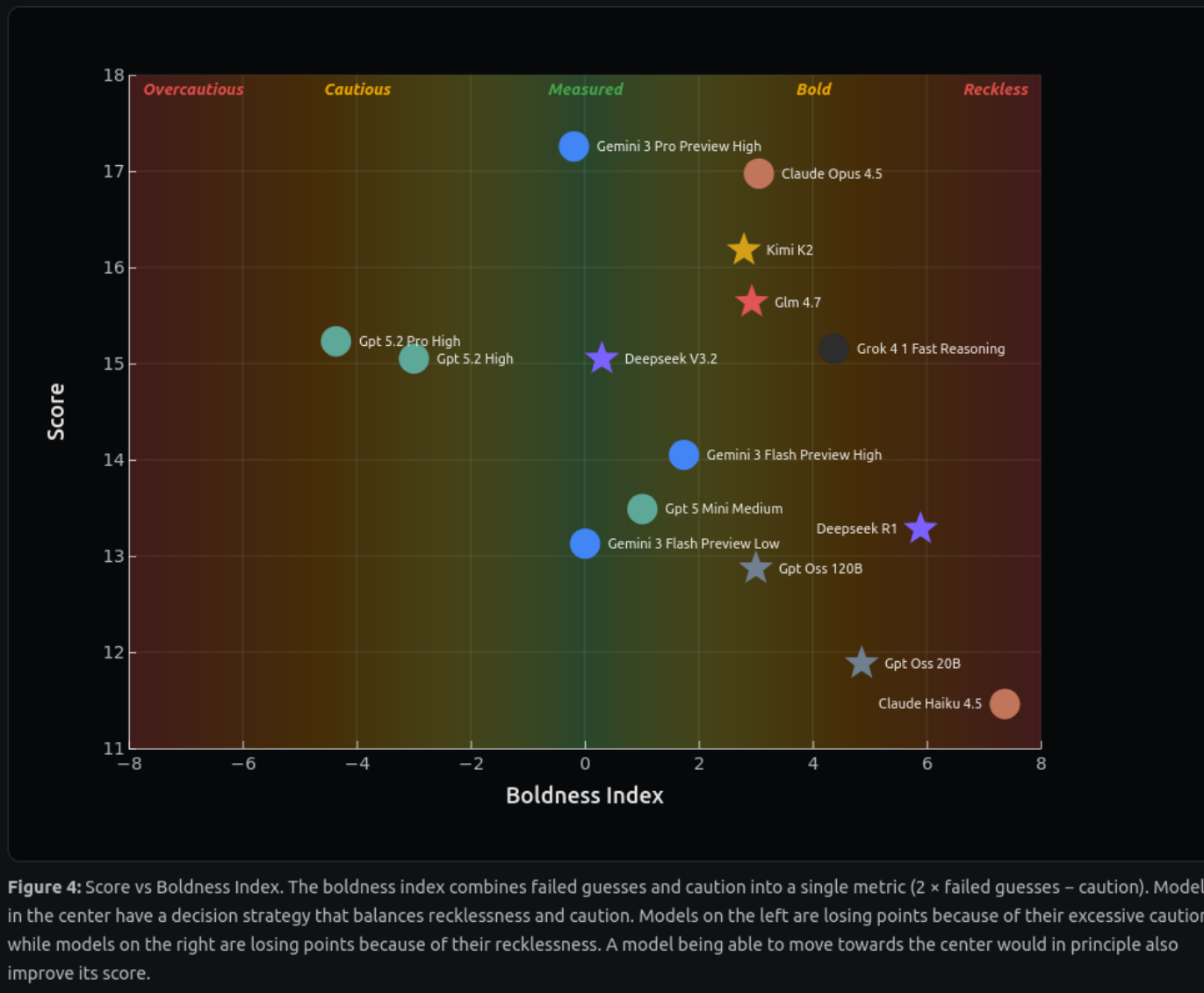
$$\textit{score} = (31 - \textit{nombre de tours passé}) - 2 \times \textit{mauvais paris}$$

Ce score pénalise le fait d'attendre trop longtemps pour deviner la règle et le fait de faire trop de mauvais paris.

Résultats







Conclusion

Cas d'usage des LLMs

- raisonnement inductif (capacité de déduire la règle)
- calibration méta cognitive (comment savoir si je suis sûr)

Merci de votre attention

Sources/Bibliographie

- [1]: Simulating scientific inquiry with the card game eleusis, Charles Romesburg, 1979
<https://onlinelibrary.wiley.com/doi/10.1002/sce.3730630505>
- David Louapre (2026). "Can LLMs Play the Game of Science?".
- Dépôt du code utilisé pour cette étude
- Règles Eleusis
- Vidéo de la chaîne Hygiène Mentale sur Eleusis

Backup slides

L'humain est-il bon en logique ?

Hygiène Mentale - Ep10 Raisonner de façon
correcte (Testez votre logique)



Hypothèse

Si une carte a une Dame d'un côté, alors l'autre côté est Bleu.

1 2 3 4

Quelle(s) carte(s) faut il retourner pour vérifier l'hypothèse ?



Hypothèse

Si une carte a une Dame d'un côté, alors l'autre côté est Bleu.

1 2 3 4

Modèles testés

Model	Lab	Provider	Reasoning	Weights	
Kimi K2 Thinking	Moonshot	🙌 Inference providers	Default	Open	
GLM 4.7	Z.ai	🙌 Inference providers	Default	Open	
DeepSeek R1	DeepSeek	🙌 Inference providers	Default	Open	
DeepSeek V3.2	DeepSeek	🙌 Inference providers	Default	Open	

Comment évaluer la règle formulée

les règles peuvent être formulées différemment et pourtant être identique :
"alterner les couleurs" et "rouge après noir, et noir après rouge". Pour vérifier si la règle proposée est équivalente à celle à deviner, on convertit les règles en code Python et on simule ce qui pourrait se passer dans la suite du jeu. Si les résultats sont les mêmes alors on considère que les règles sont équivalentes.