AU-DELA DU « BLUE DOT EFFECT » : L'IA COMME LEVIER DE REDUCTION DES BIAIS COGNITIFS EN ORTHOPHONIE

BEYOND THE "BLUE DOT EFFECT": LEVERAGING AI TO MITIGATE COGNITIVE BIASES IN SPEECH-LANGUAGE PATHOLOGY.

Dr. Kawther BENHAMMADA*
Université Alger2, laboratoire langage et cognition : Développement et pathologie ; COGLANG (Algérie)
kawther.benhammada@univ-alger2.dz

Date de soumission: 27/03/2026 ***Date d'acceptation: 10/05/2026***

Résumé: Le jugement clinique en orthophonie est exposé à des distorsions cognitives systématiques, notamment le glissement conceptuel induit par la prévalence (effet Blue Dot), l'ancrage et le biais de confirmation. Ces biais peuvent entraîner des erreurs diagnostiques, des évaluations parcellaires et des disparités dans la prise en charge. Cet article examine comment l'Intelligence Artificielle (IA) peut fonctionner comme un partenaire de débiaisage structuré. En fournissant des retours perceptifs standardisés, en permettant une comparaison objective des données et en agissant comme un second regard via des modèles de raisonnement multi-agents, l'IA aide les cliniciens à remettre en question leurs hypothèses initiales. Toutefois, l'intégration de l'IA introduit de nouveaux risques, tels que le biais d'automatisation. Nous concluons qu'une approche collaborative, où l'IA soutient l'expertise clinique sans s'y substituer, est indispensable pour accroître la précision diagnostique et garantir l'équité des soins.
Mots clés: Biais cognitifs, Intelligence Artificielle, Neuropsychologie, Orthophonie, Blue Dot Effect, Prise de décision clinique

Abstract: Clinical judgment in speech-language pathology is vulnerable to systematic cognitive distortions, including prevalence-induced concept change (the "Blue Dot Effect"), anchoring, and confirmation bias. These biases can lead to misdiagnosis, incomplete assessments, and inequitable patient care. This article explores how Artificial Intelligence (AI) can serve as a "structured debiasing partner." By providing standardized perceptual feedback, offering objective data comparison, and acting as a "second mind" through multi-agent reasoning models, AI helps clinicians challenge their initial assumptions. However, the integration of AI also introduces new risks, such as automation bias. We conclude that a collaborative "human-in-the-loop" approach, where AI supports rather than replaces clinical expertise, is essential for improving diagnostic accuracy and ensuring fairness in clinical practice.
key words: Cognitive Bias, Artificial Intelligence, Neuropsychology, Speech-Language Pathology, Blue Dot Effect, Clinical Decision-Making.

* Auteur correspondant

1.Introduction :

Un biais cognitif désigne une distorsion systématique dans le traitement de l'information qui conduit à des jugements ou des décisions erronées. Ces biais sont des tendances naturelles de l'esprit humain à traiter et interpréter les informations de manière subjective, souvent influencée par des émotions, des expériences passées ou des heuristiques. Les biais cognitifs peuvent affecter la perception la réalité, la prise des décisions. Ils peuvent être liés à des facteurs comme l'attachement émotionnel, les attentes personnelles ou des processus de pensée simplifiés qui favorisent des conclusions rapides, mais parfois inexactes.

Le processus diagnostique en rééducation cognitive est intrinsèquement exposé à des biais subjectifs. Parmi ceux-ci, le phénomène de **glissement conceptuel induit par la prévalence** (communément appelé *Blue Dot Effect*) décrit une distorsion cognitive où la perception des stimuli change à mesure que leur fréquence diminue : plus un symptôme devient rare, plus le clinicien tend à élargir sa définition, risquant ainsi de pathologiser des variantes de la norme (Levari et al., 2018). En contexte clinique, ce biais influence la manière dont le thérapeute perçoit, interprète et évalue les données neurocognitives, altérant l'objectivité indispensable au diagnostic. L'intelligence artificielle (IA) peut jouer un rôle crucial pour atténuer les biais cognitifs chez les professionnels, notamment en fournissant des outils pour réduire leurs impacts sur les processus d'évaluation, de diagnostic et de traitement des patients. Dans cet article nous allons exposer plusieurs manières dont l'IA peut contribuer à minimiser cet effet.

2. Glissement conceptuel induit par la prévalence (effet « Blue Dot ») :

Le "Blue Dot Effect" réfère à un biais cognitif particulier qui peut influencer la manière dont le thérapeute perçoit, interprète et évalue les informations liées à un patient, notamment lors d'évaluations neurocognitives. Il peut être compris comme un phénomène où des informations ou des symptômes familiers sont perçus comme plus significatifs ou représentatifs qu'ils ne le sont réellement. Ce biais pourrait affecter l'objectivité de l'évaluation neurocognitive, en particulier dans les contextes de diagnostic de troubles cognitifs ou de troubles cérébraux.

Chez un thérapeute, l'effet "Blue Dot" pourrait se manifester par une multiplicité de micro biais, tels qu'une tendance à surévaluer ou à sous-évaluer certains éléments d'information basés sur leur familiarité, leur confort ou leur cohérence avec les croyances personnelles ou les expériences passées. Cela peut avoir plusieurs implications dans le cadre d'une consultation thérapeutique. (Roscoe & Taylor, 2022)

2.1. Biais de Familiarité :

Le biais de familiarité est un élément central du Blue Dot Effect. Un thérapeute pourrait être influencé par des éléments ou des schémas qu'il reconnaît ou comprend facilement, même si ces éléments ne sont pas nécessairement représentatifs de la situation réelle du patient. Par exemple, si un patient présente des symptômes similaires à ceux d'un cas déjà traité dans le passé, le thérapeute pourrait, involontairement, accorder trop d'importence à ce parallélisme et réduire l'objectivité de l'évaluation du patient.

Si un thérapeute a traité plusieurs patients souffrant d'anxiété sociale, il pourrait, de manière inconsciente, reconnaître les signes de cette condition chez un nouveau patient, même si ce dernier présente un ensemble de symptômes légèrement différents. Le thérapeute pourrait alors

évaluer cette situation en fonction de ses précédentes expériences, limitant ainsi son exploration d'autres possibilités diagnostiques. (Speerforck et al., 2024)

2.2. Biais de Confirmation :

Ce biais désigne la tendance à privilégier les informations qui confirment les hypothèses ou les attentes préalables. Si un thérapeute a une première impression forte ou une hypothèse sur un patient (par exemple, une tendance à voir un patient comme ayant un trouble TSA il pourrait inconsciemment chercher des indices ou des réponses qui confirment cette hypothèse, négligeant des éléments de l'histoire du patient qui pourraient suggérer un autre diagnostic. (Levari, 2022).

Exemple :

Un thérapeute qui a une attente préconçue que son patient souffre de TSA pourrait se concentrer principalement sur des symptômes tels que les troubles du comportement et les troubles de la communication, en ignorant des indices qui pourraient suggérer un trouble lié à l'anxiété de séparation ou une anxiété de morcellement.

2.3. Effet de Halo et Premières Impressions :

Ce biais consiste à évaluer globalement un patient sur la base de certaines premières impressions, comme un comportement particulier lors de la première séance. Si un patient donne une impression très positive ou négative, cette première évaluation peut influencer la manière dont le thérapeute perçoit et interprète toutes les autres informations, même si elles sont contradictoires. (Roscoe & Taylor, 2022; Cardoso et al., 2022; King, 2017)

Exemple :

Si un patient semble particulièrement détendu et ouvert lors de la première consultation, le thérapeute pourrait inconsciemment supposer qu'il est facile à traiter ou qu'il n'a pas de trouble avéré tel qu'un enfant s'exprimant bien oralement allant même à faire des jeux de mots, mais souffre d'une dyslexie de surface sévère. Cela pourrait influencer la profondeur de l'évaluation ou la manière dont le thérapeute aborde les difficultés cognitives du patient.

3. Conséquences du glissement conceptuel induit par la prévalence chez un thérapeute :

L'impact du Blue Dot Effect peut avoir des conséquences importantes, notamment :

3.1 Diagnostic erroné ou retardé : L'un des risques majeurs est un diagnostic incorrect, car les biais peuvent amener le thérapeute à interpréter des résultats en fonction de préjugés ou d'attentes personnelles. Cela peut entraîner un retard dans le traitement approprié du patient.

3.2 Évaluation incomplète : Si le thérapeute se concentre sur certains symptômes familiers ou facilement identifiables, d'autres symptômes importants peuvent être ignorés, ce qui limite l'évaluation globale du patient.

3.3 Manque de diversité dans les recommandations de traitement : un biais cognitif peut affecter les suggestions de traitement ou de gestion, qui pourraient être inadaptées au patient si elles sont basées sur des idées préconçues plutôt que sur une évaluation complète et objective.

3.4 Perception déformée de l'état du patient : Les biais cognitifs peuvent également influencer la perception que le thérapeute a de la progression du patient. Par exemple, il pourrait

juger un patient comme étant "plus stable" ou "moins affecté" qu'il ne l'est réellement, en raison d'une première impression trompeuse.

4. L'intelligence artificielle comme levier de réduction des biais cognitifs en orthophonie

L'intelligence artificielle (IA) peut accompagner les orthophonistes dans la remise en question de leurs biais cognitifs, à condition d'être utilisée de manière délibérée et critique.

4.1. Agir comme un « second regard » dans la prise de décision

- Aide à la décision : Les outils d'aide à la décision fondés sur l'IA peuvent mettre en évidence des diagnostics ou des axes thérapeutiques alternatifs. Cela force le praticien à réévaluer sa position lorsqu'il est exposé aux risques d'ancrage, de clôture prématurée ou de biais de confirmation (Brown et al., 2023 ; Ke et al., 2024).
- Approche multi-agents : Les configurations de modèles de langage (LLM) multi-agents, intégrant des rôles tels que « l'avocat du diable », ont significativement amélioré la précision diagnostique par rapport aux humains dans des cas médicaux empreints de biais, en contestant explicitement les hypothèses initiales (Ke et al., 2024).
- Modèles de raisonnement : Des études sur vignettes cliniques suggèrent que les modèles axés sur le raisonnement présentent des taux de biais cognitifs inférieurs à ceux des cliniciens, confirmant leur valeur en tant qu'outils de vérification (Degany et al., 2025).

4.2 Standardisation des jugements et des retours

- Troubles de l'articulation et de la parole : Les systèmes d'IA prédisant les jugements perceptifs cliniques (ex: ChainingAI/PERCEPT-R) offrent un feedback automatisé et cohérent, comparable à celui des évaluateurs humains, réduisant ainsi la variabilité liée aux biais perceptifs individuels (Benway & Preston, 2024).
- Neutralisation de la variabilité : L'analyse automatisée permet de « neutraliser » la variabilité inter-juges observée dans la notation humaine de signaux complexes (sommeil, parole), offrant ainsi un point de référence uniforme (Norori et al., 2021 ; Benway & Preston, 2024).

4.3 Réduction de la charge cognitive et des biais liés à l'épuisement professionnel

Les outils d'IA optimisant la documentation, la recherche et le suivi permettent d'alléger la charge de travail. Indirectement, cela réduit les erreurs dues à la fatigue et les jugements hâtifs (Gandhi et al., 2023 ; Giordano et al., 2021 ; Georgiou, 2025).

5. Limites et stratégies d'utilisation :

5.1 Limites d'utilisation et nouveaux risques de biais

- Héritage des biais : Les systèmes d'IA héritent des biais sociétaux et des données d'entraînement (langue, accent, origine ethnique, statut socio-économique), ce qui peut accentuer les inégalités si l'outil est utilisé sans recul critique (Cross et al., 2024 ; Yang et al., 2024 ; Norori et al., 2021).
- Limites des LLM : Les modèles de langage présentent eux-mêmes des biais cognitifs qui ne sont pas résolus par de simples avertissements ; la prévention ne réduit le biais que de manière modeste (Wang & Redelmeier, 2025).

- Biais d'automatisation : La dépendance excessive à l'IA est un danger documenté ; le clinicien doit impérativement conserver son esprit critique et le jugement final (Brown et al., 2023 ; Challen et al., 2019 ; Aquino, 2023).

5.2 Stratégies pratiques d'utilisation :

1. Utiliser l'IA comme partenaire de débiaisage structuré : Sollicitation des hypothèses alternatives (« Qu'est-ce que cela pourrait être d'autre ? ») ou des raisons pour lesquelles le plan initial pourrait être erroné.
2. Auditer ses propres pratiques : Comparer les évaluations avec les scores générés par l'IA pour identifier les propres schémas de jugement.
3. Privilégier l'explicabilité : Choix d'outils dotés d'interfaces transparentes et centrées sur l'humain, favorisant la réflexion plutôt qu'une acceptation aveugle.
4. Vigilance sur l'équité : Interrogation de la performance du modèle sur les populations spécifiques (dialectes, bilinguisme, cultures) et écarter les outils sous-performants dans ces contextes.

Conclusion

Le Blue Dot Effect est un biais cognitif qui peut affecter la manière dont nous, thérapeutes, évaluons nos patients. Il peut conduire à des diagnostics erronés et des traitements inappropriés. Cependant, l'introduction de l'IA dans la pratique clinique peut réduire considérablement ces biais en fournissant une analyse objective et personnalisée des données.

L'IA peut aider les orthophonistes à identifier et contrer leurs propres biais cognitifs et perceptifs en fournissant des seconds avis cohérents, des retours standardisés et des incitations décisionnelles qui remettent en cause l'ancrage. Pour que l'IA réduise réellement les biais au lieu de simplement les transférer dans un logiciel, le thérapeute doit la considérer comme un soutien à la pensée critique, et non comme une autorité, tout en restant attentif à l'impact des biais algorithmiques sur ses patients.

Pour tirer parti de l'IA, les orthophonistes doivent se former aux outils, comprendre leurs capacités et leurs limites, et les intégrer dans le cadre d'une évaluation neurocognitive rigoureuse. Chercheurs et praticiennes doivent garder a l'esprit que L'IA ne remplace pas l'expertise humaine, mais elle constitue un outil complémentaire qui peut améliorer la précision des évaluations et aider à prendre des décisions plus éclairées.

Références

Aquino, Y. (2023). Making decisions: Bias in artificial intelligence and data-driven diagnostic tools.. Australian journal of general practice, 52 7, 439-442. https://doi.org/10.31128/ajgp-12-22-6630

Benway, N., & Preston, J. (2024). Artificial Intelligence–Assisted Speech Therapy for /ɹ/: A Single-Case Experimental Study. American Journal of Speech-Language Pathology, 33, 2461 - 2486. https://doi.org/10.1044/2024_ajslp-23-00448

Brown, C., Nazeer, R., Gibbs, A., Page, L., & Mitchell, A. (2023). Breaking Bias: The Role of Artificial Intelligence in Improving Clinical Decision-Making. Cureus, 15. https://doi.org/10.7759/cureus.36415

Cardoso, C., Ferreira, Â., Pinto, D., & Ribeiro, E. (2022). Therapist's interventions immediately after exceeding the client's therapeutic zone of proximal development: A comparative case study. Psychotherapy Research, 33, 70 - 83. https://doi.org/10.1080/10503307.2022.2153093

Celi, L., Cellini, J., Charpignon, M., Dee, E., Dernoncourt, F., Eber, R., Mitchell, W., Moukheiber, L., Schirmer, J., Situ, J., Paguio, J., Park, J., Wawira, J., & Yao, S. (2022). Sources of bias in artificial intelligence that perpetuate healthcare disparities—A global review. PLOS Digital Health, 1. https://doi.org/10.1371/journal.pdig.0000022

Challen, R., Denny, J., Pitt, M., Gompels, L., Edwards, T., & Tsaneva-Atanasova, K. (2019). Artificial intelligence, bias and clinical safety. BMJ Quality & Safety, 28, 231 - 237. https://doi.org/10.1136/bmjqs-2018-008370

Clémence Brun. Influence de l'intolérance à l'incertitude sur la prise de décision dans le domaine de la santé.

Cross, J., Choma, M., & Onofrey, J. (2024). Bias in medical AI: Implications for clinical decision-making. PLOS Digital Health, 3. https://doi.org/10.1371/journal.pdig.0000651

Degany, O., Laros, S., Idan, D. et al. Evaluating the o1 reasoning large language model for cognitive bias: a vignette study. Crit Care 29, 376 (2025). https://doi.org/10.1186/s13054-025-05591-5

Degany, O., Laros, S., Idan, D., & Einav, S. (2025). Evaluating the o1 reasoning large language model for cognitive bias: a vignette study. Critical Care, 29. https://doi.org/10.1186/s13054-025-05591-5

Deka, C., Shrivastava, A., Abraham, A., Nautiyal, S., & Chauhan, P. (2024). AI-based automated speech therapy tools for persons with speech sound disorder: a systematic literature review. Speech, Language and Hearing, 28. https://doi.org/10.1080/2050571x.2024.2359274

Gandhi, T., Classen, D., Sinsky, C., Rhew, D., Garde, N., Roberts, A., & Federico, F. (2023). How can artificial intelligence decrease cognitive and work burden for front line practitioners?. JAMIA Open, 6. https://doi.org/10.1093/jamiaopen/ooad079

Georgiou, G. (2025). Transforming Speech-Language Pathology with AI: Opportunities, Challenges, and Ethical Guidelines. Healthcare, 13. https://doi.org/10.3390/healthcare13192460

Giordano, C., Brennan, M., Mohamed, B., Rashidi, P., Modave, F., & Tighe, P. (2021). Accessing Artificial Intelligence for Clinical Decision-Making. Frontiers in Digital Health 3.https://doi.org/10.3389/fdgth.2021.645232

Ke, Y., Yang, R., Lie, S., Lim, T., Ning, Y., Li, I., Abdullah, H., Ting, D., & Liu, N. (2024). Mitigating Cognitive Biases in Clinical Decision-Making Through Multi-Agent Conversations Using Large Language Models: Simulation Study. Journal of Medical Internet Research, 26. https://doi.org/10.2196/59439

King, G. (2017). The Role of the Therapist in Therapeutic Change: How Knowledge From Mental Health Can Inform Pediatric Rehabilitation. Physical & Occupational Therapy In Pediatrics, 37, 121 - 138. https://doi.org/10.1080/01942638.2016.1185508

Levari, D. (2022). Range-frequency effects can explain and eliminate prevalence-induced concept change. Cognition, 226. https://doi.org/10.1016/j.cognition.2022.105196

Nissen-Lie, H., Rønnestad, M., Høglend, P., Havik, O., Solbakken, O., Stiles, T., & Monsen, J. (2015). Love Yourself as a Person, Doubt Yourself as a Therapist?. Clinical psychology & psychotherapy, 24 1, 48-60. https://doi.org/10.1002/cpp.1977

Norori, N., Hu, Q., Aellen, F., Faraci, F., & Tzovara, A. (2021). Addressing bias in big data and AI for health care: A call for open science. Patterns, 2. https://doi.org/10.1016/j.patter.2021.100347

Restrepo, D., Armstrong, K. A., & Metlay, J. P. (2020). Annals Clinical Decision Making: Avoiding Cognitive Errors in Clinical Decision Making. Annals of internal medicine, 172(11), 747–751. https://doi.org/10.7326/M19-3692

Roscoe, J., & Taylor, J. (2022). Maladaptive therapist schemas in CBT practice, training and supervision: A scoping review.. Clinical psychology & psychotherapy. https://doi.org/10.1002/cpp.2802

Rosenbacke, R., Melhus, Å., Mckee, M., & Stuckler, D. (2024). How Explainable Artificial Intelligence Can Increase or Decrease Clinicians' Trust in AI Applications in Health Care: Systematic Review. JMIR AI, 3. https://doi.org/10.2196/53207

Speerforck, S., Jürgensen, V., Göbel, M., Meyer, N., & Schomerus, G. (2024). 'concept creep' in perceptions of mental illness — an experimental examination of prevalence-induced concept change. *European Archives of Psychiatry and Clinical Neuroscience.* Advance online publication. https://doi.org/10.1007/s00406-023-01737-0

Wang, J., & Redelmeier, D. (2025). Forewarning Artificial Intelligence about Cognitive Biases. Medical Decision Making, 45, 913 - 916. https://doi.org/10.1177/0272989x251346788

Yang, Y., Lin, M., Zhao, H., Peng, Y., Huang, F., & Lu, Z. (2024). A survey of recent methods for addressing AI fairness and bias in biomedicine. ArXiv. https://doi.org/10.1016/j.jbi.2024.104646