



HAL
open science

Confiance entre les humains et les machines intelligentes et biais cognitifs induits

Gilles Desclaux

► **To cite this version:**

Gilles Desclaux. Confiance entre les humains et les machines intelligentes et biais cognitifs induits. Bernard Claverie; Baptiste Prébot; François du Cluzel. Cognitive Warfare, La guerre cognitive, NATO Collaboration Support Office, pp.5, 1-6, 2021, 978-98-837-2368-4. hal-03424925

HAL Id: hal-03424925

<https://hal.science/hal-03424925v1>

Submitted on 10 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



COGNITIVE WARFARE LA GUERRE COGNITIQUE

*Première réunion scientifique Cognitive Warfare
Bordeaux (FR) - 21 juin 2021*

Sous la direction scientifique de : B. CLAVERIE, B. PRÉBOT ET F. DU CLUZEL.



*Journée organisée par l'Innovation Hub de NATO-ACT et l'ENSC,
avec le soutien de l'État-Major des Armées (FR – Major Général),
du NATO-CSO et de la Région Nouvelle Aquitaine.*

Ouvrage publié par le CSO - Neuilly (France) © édition 2021

Chapitre 5 – CONFIANCE ENTRE LES HUMAINS ET LES MACHINES INTELLIGENTES ET BIAIS COGNITIFS INDUITS

Général Gilles Desclaux¹

*« L'homme a beaucoup appris des machines qu'il a lui-même construites,
sauf peut-être savoir mieux vivre avec elles. »*

Le domaine stratégique de la gestion de crise repose à la fois sur la connaissance de l'information la plus complète possible, la confiance dans les meilleures technologies qui les délivrent, et l'aptitude décisionnelle du commandeur qui s'appuie sur une organisation forte et efficace.

Dans le contexte de l'information massive, ces trois dimensions nécessitent le développement d'agents logiciels dits « intelligents », capables de sélectionner, fusionner et représenter les informations pertinentes et de proposer à haute vitesse des solutions décisionnelles. Ces agents sont développés par de grands industriels ; ils progressent de façon constante vers une plus grande autonomie. Malgré ces progrès, et face à une complexité croissante de la criticité des situations, le projet de systèmes purement autonomes s'éloigne des perspectives réalistes à court et moyentermes. Les experts de la gestion de crises et ces systèmes artificiels doivent de plus en plus travailler de manière collaborative, chacun apportant au duo humain-système le meilleur de ses compétences. La notion de confiance est alors centrale pour l'Interaction/Intégration Humain-Système (I₂HM), et la collaboration entre humains et machines. La robustesse ou la faiblesse de cette relation de collaboration est un enjeu clé pour la sécurité, et donc l'une des cibles de la guerre cognitive (*cyber warfare*).

5.1 LA COLLABORATION HUMAINS-MACHINES POUR LA GESTION DE CRISE

Le management des systèmes de défense ou d'opérations militaires est un domaine aussi complexe que codifié. La gestion rapide des crises en est un des domaines stratégiques. La doctrine, le droit de la guerre, la responsabilité quant à l'attrition humaine minimale pour une efficacité tactique matérielle adéquate bornent l'action du décideur qui doit pourtant agir vite et bien. Gérer une crise, c'est mobiliser de manière la plus efficace possible des moyens mis à disposition pour imaginer, évaluer et mettre en œuvre les solutions mesurées et mesurables les plus pertinentes menant à une solution favorable la plus rapide possible. Les crises peuvent être ponctuelles, en lieu comme en temps, ou plus globales et durables, nécessitant des ajustements ou des solutions dont la complexité évolue avec de multiples dimensions évolutives à prendre en compte.

Pour cela, la connaissance est le véritable « carburant » de la mesure, de l'anticipation et de la conduite de l'action. Elle est un critère majeur de différenciation pour maîtriser la criticité des situations. Elle s'élabore à partir de masses de données qui dépassent aujourd'hui les capacités humaines de représentation ou de compréhension globales, et nécessite le recours à des techniques utilisant les *Big Data*, l'intelligence artificielle » et la « visualisation » de solutions potentielles et changeantes sur lesquelles repose la décision.

Depuis quelques années, le développement d'agents logiciels « intelligents » progresse vers une plus grande autonomie. De nombreux obstacles restent à surmonter pour atteindre la perspective de véritables systèmes capables de se substituer efficacement aux experts humains. Dans un avenir proche, ces experts

¹ Gilles Desclaux est Général de Corps Aérien (2s), président de RACAM (Réunion Aviation Civile - Aviation Militaire). Il est chercheur du Laboratoire Human Engineering for Aerospace Laboratory (HEAL) – ENSC Bordeaux-INP / THALES. Il y coordonne la recherche sur les processus d'aide IA à la décision humaine « Anticipo » dans le « C2 Air. »

et les systèmes artificiels vont devoir continuer à « travailler en équipe », de manière encore plus collaborative. Le concept de *Human-Autonomy Teaming (HAT)* a été proposé pour cela par des équipes de la NASA en 2018 (O'Neill et al., 2020) pour rendre compte de cette « collaboration étrange », qui mêle Intelligence artificielle (IA) et Intelligence naturelle (IN). Elle contribue à l'émergence de systèmes hybrides, anthropotechnique, une forme d'intelligence duale et partagée, qui n'est pas sans poser de problèmes concrets de fragilité et de fiabilité dans le domaine cognitif.

5.2 UNE COOPERATION BASEE SUR DES PROCESSUS COGNITIFS DIFFERENTS

Le processus de décision mis en œuvre par les humains est radicalement différent de celui des machines intelligentes. Des architectures cognitives identiques pourraient faciliter la communication, mais contrairement aux humains, les machines sont restreintes à des objectifs et des priorités bien définis, sans capacité d'improvisation ou d'adaptation interprétatives, et sans réelle inventivité au-delà de la proposition algorithmique de solutions inattendues. Les humains, cependant, peuvent développer ces qualités mais restent médiocres dans la description précise de leurs intentions, de leurs objectifs et de leurs priorités comme les machines intelligentes l'exigent. De même, leurs capacités en attention, en mémoire ou en fiabilité de raisonnement sont fragiles et fréquemment mises en défaut alors que les systèmes artificiels sont particulièrement fiables en la matière.

Au sein d'un « réseau décisionnel » de type HAT, humains et machines modifient continuellement leurs propres rôles, tâches et relations avec les autres acteurs, naturels comme artificiels, partenaires comme extérieurs. Cette activité est dite « réseaux centrée ». Lorsque les processus habituels semblent ne pas correspondre à leurs attentes, des stratégies nouvelles sont mises en œuvre : les machines ouvrent des procédures de consultation de bases de données extérieures, alors que les humains constituent ou restructurent des groupes de travail informels ou ad hoc, et recherchent de nouveaux experts.

Les machines intelligentes restent et resteront, tout au moins dans un proche avenir, partiellement incompréhensibles pour les humains. Il en est évidemment de même des humains pour les machines. Établir une confiance entre les deux types d'entités est donc difficile. Les machines intelligentes sont sensibles aux cyber-intrusions qui peuvent compromettre leurs « perceptions », la pertinence de leur « prise de décision » et leurs capacités de gestion de données et de communication. Les humains présentent d'autres fragilités, telles que fatigue, mémoire limitée, et capacités cognitives fragiles et influençables. Dans un tel contexte, une solution consiste à favoriser l'établissement de relations constructives de surveillance de la performance, entre experts humains, entre machines et, dans les deux sens, entre experts et machines.

5.3 LE PROBLEME DE L'INTERPRETABILITE

L'interprétabilité présente deux dimensions. Dans un premier aspect, elle correspond, pour l'utilisateur d'un système automatisé ou autonome, à son degré de compréhension de ce que le système fait, comment il le fait et pourquoi il le fait. L'interprétabilité du système peut conduire au développement d'un modèle cognitif aussi complet que possible afin de fournir une compréhension de son fonctionnement, et la capacité de prédire ce qu'il ferait dans certaines circonstances. Deux approches permettent de faciliter l'interprétabilité :

- La rétroaction (*feed-back*) du système améliore l'expérience des interactions avec les utilisateurs et facilite leur sentiment de contrôle. Ceux-là souhaitent habituellement que le système fournisse lui-même des informations compréhensibles sur son propre niveau de fiabilité afin de pouvoir ou non lui faire confiance.
- L'explication post-hoc, connue dans le monde anglo-saxon sous le nom d'IA explicable (Adadi et Berrada, 2018) ou XAI, fournit à l'utilisateur une explication qui justifie la prise de décision, rendant ainsi le système plus interprétable et facilitant le Retour d'expérience (Retex).

Dans un second aspect, l'interprétabilité concerne la limitation, pour l'utilisateur ou le partenaire humain, à des comportements ou des prises de décisions compréhensibles pour la machine, ou cohérentes avec ses propres registres de connaissances. Cette limite est nécessaire au maintien du lien de collaboration efficace. Cette dimension n'est pas sans poser de problèmes d'acceptabilité pour des utilisateurs humains naïfs, qui doivent apprendre à collaborer avec les machines pour faciliter la compétence et le maintien de l'efficacité du système HAT. Là encore, les systèmes apprenants sont soumis à la fréquentation des experts et doivent pouvoir les identifier afin de s'adapter à leurs particularités et les spécificités de leurs caractéristiques cognitives : personnalité, âge, performances mnémoniques plus ou moins grandes, visuelles ou formelles, sensibilité aux sons ou aux images, dépendance ou indépendance au champ, saturations attentionnelles, résistance à la fatigue, maîtrise du stress, etc. Le recours aux technologies portables (*wearable tech.*), aux capteurs et aux auto-quiz sur tablettes est aujourd'hui étudié pour cela par les laboratoires de l'U.S. Army (Buchler et al., 2016) et dans le cadre de collaborations entre certains industriels et des écoles d'ingénieurs universitaires ou de la Défense des pays de l'OTAN.

Bien que cette piste soit encore exploratoire, on prévoit des technologies capables de faciliter la collaboration et l'efficacité du couple humain-système et la performance de la mission en fonction du partenaire humain reconnu et identifié par la machine, et qui informe en continue la machine de l'évolution de son état cognitif et de ses connaissances.

5.4 L'ÉVALUATION DE L'INCERTITUDE

À ce jour, la plupart des automatisations décisionnelles fonctionnent de manière satisfaisante pour des situations spécifiques, et pour lesquelles elles sont conçues, mais nécessitent le recours à une expertise humaine dès qu'il s'agit de gérer des situations en dehors de certains environnements définis ou limités. Notamment, lorsque les algorithmes informatiques sont confrontés à l'incertitude et à l'ambiguïté des données, ils sont souvent dépassés dans le registre de la prise de décision.

Les humains surpassent les machines pour la compréhension du contexte. Les machines restent incapables d'exercer un jugement nuancé sur les environnements complexes ou ambigus et évolutifs. De plus, étant donné que des machines sont programmées ou entraînées à l'aide d'ensembles d'informations pertinentes pour une tâche ou un problème spécifique, rencontrer un nouveau problème à tendance à entraîner des ambiguïtés, voire provoquer un échec. La capacité humaine à s'adapter à de nouvelles situations est bien supérieure et même des réponses incomplètes ou imparfaites sont susceptibles d'être performantes. Les humains utilisent des capacités de substitution mentale et des estimations à partir de compétences ou de tâches familières, et peuvent ainsi fournir des réponses approximatives, ce que les technologies d'IA ne sont pour l'instant pas capables de faire.

Ils surpassent également les machines dans ce qui est leur capacité à évaluer la qualité de leur cognition. La métacognition est une des caractéristiques de l'esprit humain. Elle échappe pour l'instant à la machine. Des travaux sont entrepris afin de comprendre ce phénomène d'expertise sur sa propre expertise cognitive, en donner un formalisme qui puisse être compris par la machine, et doter celle-là de capacités de « méta-programmatique » pour s'évaluer elle-même, être capable d'évoluer, et surtout d'évaluer la cognition humaine pour s'adapter à son évolution ou sa performance dans une relation HAT dynamique.

5.5 LE MANQUE DE TRANSPARENCE

Lorsque les systèmes autonomes manquent de compréhensibilité et de prévisibilité, on évoque un problème de manque de « transparence ». Cette notion désigne l'incapacité qu'a l'humain de comprendre pourquoi le système prend telle mesure ou, au contraire, ne prend pas la décision d'une action attendue. Le manque de transparence produit un manque de conscience, notamment il ne permet pas aux opérateurs de savoir quelles sont les informations utilisées pour effectuer une tâche.

Ce manque de transparence est alors parfois à l'origine d'un manque de confiance qui conduit à la fois à une sous-utilisation du système par méfiance ou au contraire à une surutilisation en raison d'une confiance aveugle (Clark et al., 2014). Ce problème de confiance doit pouvoir être évalué sur des bases objectives, avec des indicateurs clairs.

Ces domaines de difficultés ne sont pas des problèmes indépendants et peuvent se combiner de manière souvent dangereuse (Endsley, 2016). Les systèmes intelligents sont fragiles, et peuvent rapidement passer d'un bon fonctionnement à une dégradation rapide et globale. Il incombe donc à l'opérateur humain de surveiller la survenue de telles défaillances, et d'en anticiper les conséquences. Or surveiller un système qui semble fonctionner correctement est un travail pour lequel les humains sont mal préparés. On évoque ici des phénomènes de « mise hors de la boucle », ou *Out-Of-The-Loop (OOTL)*, en anglais (cf. Suhir, 2021), qui induisent une conscience restreinte, voire très réduite de la situation (Endsley, 2015).

5.6 LA CONFIANCE AU CŒUR DE LA RELATION HUMAIN/MACHINE INTELLIGENTE

Dans le contexte HAT, la confiance doit être examinée à deux niveaux.

Pour la machine, la qualité de la relation se base sur des algorithmes statistiques de surveillance psychophysique ou de qualité et quantité d'information échangée. Le monitoring des partenaires humains peut permettre la mise en œuvre de processus automatisés ou de rappel à l'opérateur. Ce type de processus est particulièrement étudié dans l'aide à la conduite automobile et la détection d'états d'endormissement ou de baisse d'attention du conducteur, mais également la non-détection de dangers imminents (piéton, obstacles, verglas, etc.). Le formalisme computationnel requis nécessite un modèle cognitif du conducteur (Bellet et al., 2011). La cyber défense de ces programmes reste une des préoccupations majeures face à la nécessité d'évolution et de mise à jour permanente des logiciels.

Pour le partenaire humain, la confiance est généralement définie comme « le degré auquel un utilisateur pense qu'un système se comportera comme prévu ». Sans ce niveau de confiance approprié, les opérateurs peuvent refuser l'utilisation des systèmes autonomes ou au contraire se décharger complètement sur eux. Ces phénomènes de sur-dépendance pouvant aller jusqu'à l'échec, suivis d'une sous-dépendance à l'automatisation, sont bien documentés. Les principaux facteurs qui favorisent le développement de la confiance, sont l'acceptabilité, la tolérance, la transparence et la bidirectionnalité de la communication Humain/Système.

La confiance dépend du contexte spécifique d'une interaction humain/système intelligent, et est influencée par l'environnement et l'état mental de l'opérateur. L'utilité perçue d'un système autonome, en matière de capacité à effectuer une tâche difficile ou exigeante, influence la décision de l'individu de lui faire confiance. Mais les opérateurs ayant une charge de travail élevée ont aussi tendance à s'appuyer davantage sur la machine, quel que soit leur niveau de confiance réel dans le système. L'automate, en dehors des tâches simples, ne remplace généralement pas complètement l'humain. Au contraire, il change la nature de son travail en le déléguant de certaines tâches pour lesquelles il est plus performant. Cela pose clairement le problème de l'acceptabilité réciproque. La compréhension, la capacité d'utilisation et l'attente des utilisateurs d'un système intelligent sont corrélées à la probabilité de faire confiance.

La confiance se construit au fil du temps et, en conséquence, pour le partenaire humain, la formation et l'entraînement favorisent la familiarité nécessaire à l'utilisation du système. Pour ce qui est du système artificiel, elle doit aujourd'hui être programmée faute de pouvoir disposer d'algorithmes évolutifs, voire de machines adaptatives.

5.7 LES BIAIS COGNITIFS DANS LE DUO HUMAIN-SYSTEME

La transparence est ce qui permet à l'opérateur de déterminer si la machine autonome est susceptible de fournir la bonne réponse dans une situation complexe donnée, et ce qui permet à la machine de savoir si les informations données par l'humain sont dignes de confiance ou présentent des incongruités qu'il convient d'éclaircir.

Mais, cette transparence va au-delà de la simple mise à disposition des informations à l'opérateur humain ou au partenaire artificiel autonome. Pour être transparent, l'automate doit présenter les informations de manière adaptée au modèle mental de l'opérateur, en tenant compte de ses préférences et contraintes cognitives, alors que, à l'inverse, le partenaire humain doit s'adapter au modèle mental du concepteur du programme. Là réside un premier biais cognitif : la machine n'est pas un partenaire comme un autre, elle a été programmée par quelqu'un. Elle peut d'ailleurs être déprogrammée, reprogrammer, subir l'influence de patches ou de programmes complémentaires, et donc de virus, chevaux de Troie et autres malwares. Ce biais de dissonance cognitive est d'autant plus grâce qu'il s'impose sans réelle solution, face à des informaticiens ou industriels persuadés que leur mode de pensée est le meilleur pour les autres.

Les biais cognitifs sont des distorsions spontanées de la pensée rationnelle qu'adoptent les humains et qui sont à la sources de nombreuses erreurs (Kahneman et Tversky, 1974). Ils sont étudiés par les économistes et les psychologues, notamment pour ce qui est de la prise de décision, mais ils font l'objet d'une attention nouvelle de ces experts et de ceux du traitement de l'information, avec l'étude des biais des machines (Bertail et al., 2019) et la création algorithmique d'inéquité, voire de discrimination posant des problèmes éthiques incontournables.

Dans un contexte d'informations massives, et pour ce qui est des usagers des systèmes, les humains se concentrent le plus souvent sur des sources et des méthodes de sélection qu'ils connaissent bien et auxquelles ils font confiance, introduisant de ce fait un autre type redoutable de biais. C'est un domaine où les machines sont pourtant très performantes, apportent une vitesse élevée d'acquisition et de traitement de gros volumes d'informations, ainsi qu'une gestion cohérente, rigoureuse et impartiale des données. Mais sans un niveau de transparence qui permet de reconnaître les sources d'information et d'en analyser la qualité, l'efficacité de tels systèmes restera insuffisante, et le doute reste sous-jacent à la relation de l'humain face à la machine.

Un exemple permet d'illustrer cette notion. Un système semi-autonome présente plusieurs options qu'il a générées, accompagnées d'évaluations d'efficacité potentielle quant à l'adéquation de chacune d'elles. Un tel dispositif de facilitation de la transparence doit être accompagné de la possibilité, pour l'opérateur, d'ajouter des informations que le dispositif autonome ne connaît pas. L'opérateur doit pouvoir proposer des solutions et les faire évaluer par l'automate. La résolution collaborative de problèmes est alors un processus de va-et-vient, de type *wargaming*. Ce type de communication bidirectionnelle favorise le partenariat et permet d'évaluer les solutions favorables de résolution potentielle de problème.

Un troisième type de biais concerne le sentiment spontané de supériorité de l'humain sur la machine. Un niveau d'engagement cognitif faible rend intrinsèquement difficile, pour un opérateur, de comprendre ce qui se passe quand il n'exerce qu'une surveillance passive d'un système autonome. La passivité dans l'exécution d'une tâche est alors un obstacle à l'efficacité de l'interaction humain-machine intelligente. Ce défi dépend de ce que certains auteurs (Endsley, 2016) désignent comme « énigme de l'automatisation ». Ainsi, « plus on ajoute d'automatisation à un système, et plus cette automatisation est fiable et robuste, moins il est probable que les opérateurs humains la supervisent. Ils seront alors incapables de comprendre la situation, et auront tendance à reprendre le contrôle du système. Le système devient alors dégradé, restreint aux simples capacités limitées des opérateurs, ce qui est évidemment un avantage significateur pour l'ennemi potentiel. L'énigme de l'automatisation crée un obstacle majeur à l'autonomie dans les domaines où la sécurité est critique.

5.8 CONCLUSION

La complexité de la gestion de crises demande aujourd'hui de traiter une quantité de données, et de prendre des décisions souvent critiques dans des temps de plus en plus courts et dans des contextes de plus en plus contraints. Les décideurs à la tête d'organisations de gestion de crise doivent donc de plus en plus se reposer sur des systèmes hybrides. L'aide de systèmes intelligents est devenue indispensable. Malgré les performances incontestables de tels systèmes, ils sont encore incertains dans plusieurs domaines, et les humains, qui continueront de jouer un rôle important dans cette collaboration avec les machines, ont une tendance à ne pas maîtriser un ensemble de biais générés par l'échange HAT. Des voies de progrès résident d'une part dans la capacité de ces machines à mieux expliquer, à établir une confiance étayée, à communiquer plus aisément, voire à comprendre les intentions dissimulées et les émotions des acteurs humains, et d'autre part dans une nouvelle culture d'acceptation des machines par les humains.

Dans un article fondateur (2017), Kott et Alberts, écrivaient : « Bienvenue à bord des choses intelligentes. Quelles que soient nos lacunes respectives, nous serons plus forts et plus agiles en travaillant ensemble dans les organisations décisionnelles. »

5.9 BIBLIOGRAPHIE

- Adadi, A., Berrada, M. (2018). Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access*, 6, 52138-52160.
- Alberts, D.S., Haye, R.E (2006). *Understanding Command and Control*. Washington (DC-USA) : DoD CCRP Publication Series.
- Bellet, T. Mayenobe, P. Gruyer, D. Bornard, J.C. Claverie, B. (2011). *The Living Cognition Paradigm: An Application to Computational Modeling of Drivers' Mental Activities*. *US-China Education Review*, 1, 4, 568-578.
- Bertail, P., Bounie, D., Cléménçon, S., Waelbroeck, P. (2019). *Algorithmes : Biais, Discrimination et Équité*. Rapport de la Fondation Abeona et de Télécom ParisTech. Paris (France) : Telecom ParisTech.
- Buchler, N., Fitzhugh, S.M., Marusich, L.R., Ungvarsky, D.M., Lebiere, C., Gonzalez, C. (2016). *Mission Command in the Age of Network-Enabled Operations: Social Network Analysis of Information Sharing and Situation Awareness*. *Frontiers in Psychology*, 7, 937, 1-15.
- Clark, B.B., Robert, C., Hampton, S.A. (2014). *The Technology Effect: How Perceptions of Technology Drive Excessive Optimism*. *Journal of Business and Psychology*, 15, 4-18.
- Claverie, B. (2005). *Cognitive, Science et Pratique des Rapports à la Machine à Penser*. Paris (France): L'Harmattan.
- Claverie, B., Desclaux, G. (2015). *La Cybernétique : Commande, Contrôle et Comportement dans la Gestion des Systèmes D'information et de Communication*. *Hermès*, 71, 72-79.
- Endsley, M. (2015). *Situation Awareness Misconceptions and Misunderstanding*. *Journal of Cognitive Engineering and Decision Making*, 9, 1, 4-32
- Endsley, M. (2017). *From Here to Autonomy : Lessons Learned from Human-Automation Research*. *Human Factors*, 59, 1, 5-27.

- Gutzwiller, S.R., Espinosa, S.H., Kenny, C., Lange, D. (2018). A Design Pattern for Working Agreements in Human-Autonomy Teaming. In D.N. Cassenti (Ed.) *Advances in Human Factors in Simulation and Modeling : Proceedings of the AHFE 2017 International Conference on Human Factors in Simulation and Modeling*. New-York (NY, USA) : Springer, 12-24.
- Kahneman D., Tversky, A. (1974). Judgment Under Uncertainty: Heuristics and Biases. *Science, New Series*, 185, 4157, 1124-1131.
- Kott, A., David S.A. (2017). How Do You Command an Army of Intelligent Things? *Computer*, 12, 96-100.
- Le Guyader, H., Eshelman-Hayne, C., Irandour, H., Lange, D., Genchev A., Cakir M., Verstraete, E., Brill, J.C., Desclaux, G. (2021 – à paraître – H. Leguyader Ed.). *Human Considerations for Artificial Intelligence in Command and Control*. Technical Report of the NATO Science and Technology Organization Research Group IST-157, NATO. Paris (France) : NATO-STO Collaboration Support Office.
- O’Neill, T., McNeese, N., Barron, A., Schelble, B. (2020). *Human-Autonomy Teaming: A Review and Analysis of the Empirical Literature*. *Human Factors*. 2020 octobre 22, 18720820960865.
- Shively, R., Lachter, J., Brandt, S.L., Matessa, M., Battiste, V., Johnson, W. (2018). Why Human-Autonomy Teaming? *Proceedings of the AHFE 2017 International Conference on Neuroergonomics and Cognitive Engineering*, July 17–21, 2017, Los Angeles (CA, USA). *Advances in Neuroergonomics and Cognitive Engineering*, 586, 3-11.
- Suhir, E. (2021). *Human-In-The-Loop: Probabilistic Modeling Approach in Aerospace Engineering*. Boca Raton (FL, USA) : CRC Press.

REPORT DOCUMENTATION PAGE			
1. Recipient's Reference	2. Originator's References	3. Further Reference ISBN 978-98-837-2368-4	4. Security Classification of Document PUBLIC RELEASE
5. Originator Science and Technology Organization North Atlantic Treaty Organization BP 25, F-92201 Neuilly-sur-Seine Cedex, France			
6. Title <i>Cognitive Warfare : La guerre cognitive</i>			
7. Presented at/Sponsored by Première réunion scientifique Cognitive Warfare Bordeaux (FR) – 21 juin 2021. Journée organisée par l'Innovation Hub de NATO-ACT et l'ENSC, avec le soutien de l'État-Major des Armées (FR – Major Général), du NATO-CSO et de la Région Nouvelle Aquitaine.			
8. Author(s)/Editor(s) B. Claverie, B. Prébot, F. Du Cluzel.			9. Date October 2021
10. Author's/Editor's Address Multiple			11. Pages 116
12. Distribution Statement There are no restrictions on the distribution of this document. Information about the availability of this and other STO unclassified publications is given on the back cover.			
13. Keywords/Descriptors Biais cognitifs; Cognition; Cognitive warfare; Cyberpsychologie; Domaine cognitive; Guerre cognitive; Humain			
14. Abstract Cet ouvrage publié par le CSO de l'OTAN réunit des articles reprenant les principales interventions de la première réunion Cognitive Warfare tenue à Bordeaux en juin 2021 à l'initiative de l'Innovation Hub du Commandement pour la Transformation de l'OTAN et de l'École Nationale Supérieure de Cognitive avec la collaboration de l'État-major des armées françaises, du CSO et de l'ACT de l'OTAN, et de la Région Nouvelle Aquitaine. Cette première initiative fait un point sur la cognition humaine, sa force et ses faiblesses, son organisation collaborative pour la décision militaire, ses rapport et dépendance à la technologie numérique et ses dimensions sociales et politiques notamment dans la dure compétition internationale. Elle donne la parole au Major général des armées et au Commandeur pour la transformation de l'OTAN, et sévira de point de départ à une suite de rencontres d'approfondissement, à l'initiative du CSO et de l'ACT.			