

Lutter contre la désinformation

Penser autrement l'action publique à l'aune des sciences cognitives

Rapport sur la désinformation

Institut Jean Nicod
Département d'Études Cognitives, École Normale Supérieure - PSL; École des hautes études en sciences sociales

Auteur : Grégoire DARCY.

Éditeurs : Hugo MERCIER, Alda MARI, Gloria ORIGGI, Lydia YAHIA, Roberto CASATI.

Paris, 5 juin 2025

Institut | Nicod



Abstract :

La désinformation est perçue comme un risque sociétal majeur, mais les politiques publiques actuelles peinent à y répondre efficacement en raison de trois impensés structurants : une approche essentiellement réactive centrée sur la gestion de crise (fact-checking, debunking, communication), un modèle erroné de la psychologie humaine qui réduit le problème à un déficit de rationalité individuelle, et l'absence d'un cadre d'analyse intégratif articulant les dynamiques technologiques, sociales et cognitives. Or, la désinformation prospère sur des vulnérabilités multiples - cognitives, affectives, sociales, économiques et institutionnelles - et sur un écosystème informationnel profondément dégradé. Ce rapport propose un cadre analytique renouvelé, fondé sur les apports des sciences cognitives, et recommande une stratégie intégrée articulée autour de trois volets : renforcer de manière calibrée la résilience cognitive individuelle, transformer structurellement les environnements numériques et médiatiques, et traiter les causes sociales profondes qui alimentent la réceptivité aux récits trompeurs.

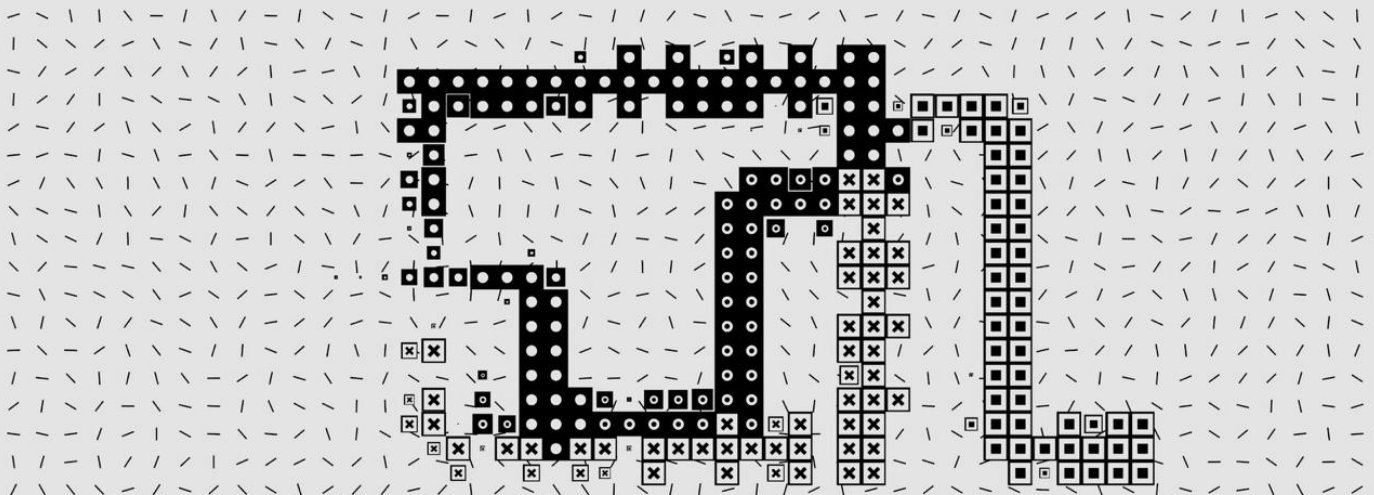
À propos de l'Institut Jean Nicod :

Unité mixte de recherche du CNRS créée en 2002 et dont les deux tutelles universitaires sont l'École normale supérieure (ENS) et l'École des hautes études en sciences sociales (EHESS), l'Institut Jean-Nicod est un laboratoire interdisciplinaire de sciences cognitives composé d'environ une centaine de membres. Le thème unificateur étant l'esprit humain et la nature des représentations (linguistiques, mentales, et sociales). www.institutnicod.org.

Auteur correspondant : Grégoire Darcy (gregoire.darcy@ens.psl.eu)

Les opinions exprimées dans ce rapport n'engagent que leurs auteurs et ne sauraient en aucun cas être attribuées aux institutions dont ils relèvent.

Pour citer ce rapport : Darcy, G. et al. 2025. "Lutter Contre La Désinformation : Penser Autrement L'action Publique À L'aune Des Sciences Cognitives." OSF Preprints. May 28. doi:10.31219/osf.io/fu9cz_v1.



Sommaire :

Résumé Exécutif	4
Introduction	8
1. La lutte contre la désinformation repose sur un modèle implicite et simpliste de la psychologie humaine parfois contre-productif.	10
1.1. La désinformation n'est pas (seulement) un problème d'irrationalité et de crédulité	10
1.2. La désinformation n'est qu'un aspect relativement marginal d'une crise informationnelle plus vaste.	11
1.3. La désinformation n'est lisible qu'à travers le prisme de l'environnement dans lequel elle se déploie	12
1.4. L'adhésion à la désinformation engage des processus cognitifs et comportementaux multiples et différenciés.	13
1.5. L'utilisation du modèle naïf de la psychologie de la désinformation a des effets potentiellement délétères.	14
2. Il faut élargir le cadrage des causes de la désinformation : au-delà des campagnes de manipulation délibérées, elle prospère sur des vulnérabilités individuelles et des facteurs structurels.	16
2.1. Les causes individuelles	16
2.2. Les causes liées à l'écologie informationnelle	18
2.3. La crise de l'information trouve ses racines dans des facteurs sociaux, institutionnels et matériels profondément ancrés.	19
3. Lutter efficacement contre la désinformation suppose de dépasser les approches traditionnelles, en construisant une stratégie intégrée fondée sur l'articulation des leviers cognitifs, informationnels et structurels.	22
Volet 1 : Les politiques de renforcement de la résilience cognitive individuelle, bien qu'amplement plébiscitées, restent peu efficaces et risquent d'entraîner des effets rebond.	22
Volet 2 : Les réponses les plus efficaces à la désinformation passent par une transformation structurelle de l'écosystème informationnel.	24
Volet 3 : La réduction des vulnérabilités sociales, économiques et institutionnelles constitue l'approche la plus structurante pour lutter contre la désinformation	27
Conclusion :	30
Ressources et Bibliographie indicative	32

Résumé Exécutif :

La désinformation, sous ses multiples formes, est perçue comme un risque sociétal majeur. En réponse, la France déploie progressivement des politiques publiques dédiées. Toutefois, la réponse institutionnelle demeure fragmentaire, et plusieurs limites structurelles en entravent l'efficacité. Trois faiblesses principales méritent d'être soulignées :

- **Une approche essentiellement réactive se concentrant sur les symptômes**, centrée sur la gestion de crise immédiate (fact-checking, modération, réponses électorales ponctuelles), au détriment d'une stratégie systémique et anticipatrice visant à restaurer durablement l'écosystème informationnel.
- **Une focalisation excessive sur les vecteurs de diffusion**, négligeant les mécanismes cognitifs, affectifs et sociaux qui favorisent l'adhésion aux récits trompeurs, et les effets concrets de la désinformation une fois intégrée par les individus.
- **Une absence de cadre d'analyse intégratif**, articulant les dynamiques technologiques, géopolitiques et cognitives. Cette lacune empêche de comprendre les vulnérabilités informationnelles dans leur complexité, et conduit à des réponses inefficaces, parfois contre-productives.

Ce rapport propose un cadre conceptuel renouvelé, appuyé sur les apports des sciences cognitives et sociales, et formule des recommandations opérationnelles pour renforcer l'impact des politiques françaises de lutte contre la désinformation. Les résultats principaux en sont les suivants :

1. La lutte contre la désinformation repose aujourd'hui sur un modèle implicite et simpliste de la psychologie humaine. Ce cadre inadéquat conduit à des diagnostics partiels et à des réponses souvent inefficaces. Pire, certaines interventions mal calibrées peuvent renforcer les croyances erronées, aggraver la défiance, ou ralentir la diffusion des informations fiables.

- **La désinformation n'est pas (seulement) un problème de crédulité ou d'irrationalité.** Contrairement au modèle de "psychologie naïve" aujourd'hui répandu, la désinformation répond à des besoins sociaux, émotionnels et identitaires plus qu'à de simples déficits de rationalité. Ainsi, corriger les erreurs factuelles ne suffit pas : il faut s'attaquer aux conditions qui rendent ces récits socialement fonctionnels.
- **La désinformation n'est qu'un aspect relativement marginal d'une crise informationnelle plus vaste.** La désinformation constitue la partie émergée d'un iceberg : visible mais peu fréquente, et à l'impact comportemental limité. C'est la dégradation globale de l'écosystème informationnel - exploitation de l'ambiguïté, malinformation, fragmentation, dégradation de l'écosystème médiatique - qui constitue la principale menace de la crise informationnelle.
- **L'efficacité de la désinformation dépend de son environnement de diffusion.** Les vulnérabilités face à la désinformation ne tiennent pas qu'aux dispositions individuelles, mais s'ancrent dans des environnements sociaux, économiques et médiatiques spécifiques : isolement social, précarité, homogamie idéologique et défiance institutionnelle sont des facteurs clés expliquant l'adhésion, bien au-delà des seuls algorithmes ou biais cognitifs. Comprendre la désinformation impose donc une lecture contextuelle, attentive aux conditions de vie, à la structure de l'espace public et à l'architecture des interactions sociales.
- **L'adhésion à la désinformation engage des processus cognitifs et comportementaux multiples et différenciés.** Croire, partager et agir ne reposent pas nécessairement sur les mêmes logiques (*e.g.* croyances intuitives ou réflexives, partages par provocation ou adhésion sincère) et ne se succèdent pas de manière binaire ni mécaniques. Les indicateurs traditionnels de visibilité - likes, partages, commentaires, impressions ou taux d'engagement - ne rendent pas compte des convictions véritables ni des actions effectives, d'où la nécessité d'analyser les usages sociaux concrets et le fossé entre convictions déclarées et comportements observés.

- **L'utilisation d'un modèle "naïf" de la psychologie de la désinformation a des effets potentiellement délétères.** Une représentation inadaptée des mécanismes psychologiques sous-jacents à la formation des croyances et à la diffusion de la désinformation peut générer des effets contre-productifs. Par exemple, le debunking peut provoquer un effet rebond, et une stratégie d'inoculation mal calibrée risque d'accroître la défiance envers les sources fiables. Les décideurs doivent donc privilégier la restauration de normes sociales claires et la mobilisation de signaux crédibles, plutôt que de se limiter à réfuter d'autorité les contenus erronés.

2. Il faut élargir le cadrage des causes de la désinformation : au-delà des campagnes de manipulation délibérées, elle prospère sur des vulnérabilités individuelles et des facteurs structurels.

Au-delà des campagnes délibérées de manipulation, la désinformation prospère sur des vulnérabilités individuelles (sociodémographiques, psychologiques, émotionnelles et identitaires) et des facteurs structurels (écosystèmes informationnels dégradés, défiance institutionnelle, isolement social et précarité économique). Ces dimensions, encore trop marginales dans l'action publique, sont pourtant centrales pour comprendre et traiter les mécanismes de diffusion des récits trompeurs.

- **Facteurs individuels** : Les déterminants individuels n'expliquent qu'une part limitée du phénomène et constituent des leviers d'action restreints pour les politiques publiques. Les variables **sociodémographiques** (faible niveau d'éducation, précarité économique, isolement social) accroissent la vulnérabilité mais restent des facteurs modérateurs plutôt que déterminants majeurs. Certains **traits de personnalité** (dogmatisme, besoin d'unicité, traits de la triade sombre) favorisent l'adhésion et la diffusion des récits trompeurs. Les **vulnérabilités cognitives** (raisonnement réflexif faible, forte réceptivité aux discours creux, prédominance de l'intuition) augmentent la crédulité, tandis que la pensée activement ouverte constitue un facteur protecteur. Enfin, les **dynamiques identitaires** (polarisation *ingroup/outgroup*, défiance envers les institutions) orientent l'évaluation de l'information selon des logiques sociales et émotionnelles, au détriment des critères épistémiques.
- **Facteurs de l'écosystème informationnel** : L'**environnement numérique**, structuré par des algorithmes d'optimisation de l'engagement, privilégie les contenus émotionnels, polarisants et simplificateurs, amplifiant la formation de bulles informationnelles et de chambres d'écho. Parallèlement, la **dégradation des médias traditionnels** - concentration des acteurs, fragilisation des modèles économiques et fragmentation des audiences - affaiblit la qualité de l'information, érode le pluralisme et nourrit la défiance, ouvrant un espace croissant aux récits alternatifs peu régulés.
- **Facteurs environnementaux** : La désinformation trouve un terrain favorable dans les **fractures sociales** (solitude, délitement des solidarités locales, polarisation affective), qui renforcent la quête de récits d'appartenance et de sens. La **défiance institutionnelle**, alimentée par la perception d'une déconnexion des élites et l'inefficacité de l'action publique, affaiblit l'impact des messages institutionnels et réduit l'efficacité des stratégies de correction. Enfin, la **précarité économique et l'aggravation des inégalités** renforcent la demande en récits simplificateurs, désignant des responsables et apportant une reconnaissance symbolique aux populations marginalisées.

3. Lutter efficacement contre la désinformation suppose de dépasser les approches traditionnelles, en construisant une stratégie intégrée fondée sur l'articulation des leviers cognitifs, informationnels et structurels.

La lutte contre la désinformation gagnerait à s'articuler autour de trois volets complémentaires, allant du plus immédiat au plus structurant à long terme : (1) le renforcement prudent de la résilience cognitive individuelle, dont l'efficacité reste incertaine et les effets parfois contre-productifs ; (2) la transformation structurelle des environnements informationnels par la régulation et le soutien aux médias fiables ; et (3) la réduction des vulnérabilités sociales, économiques et institutionnelles qui nourrissent la réceptivité aux récits trompeurs.

Volet 1 : Renforcer la “résilience cognitive” des individus constitue une piste séduisante et largement soutenue, mais dont l’efficacité réelle reste souvent surestimée — et qui, mal calibrée, peut s’avérer contre-productive, en perturbant la diffusion d’informations fiables.

- **Levier 1 - Développer les compétences informationnelles et l’esprit critique des citoyens par des actions de formation.** Les formations à l’esprit critique ou à la littératie médiatique sont des mesures populaires et largement soutenues. Elles peuvent cependant produire un effet inverse si elles renforcent une méfiance généralisée sans distinguer la fiabilité des sources. Pour être efficaces, elles doivent être régulières, bien contextualisées et accompagnées d’une exposition continue à des contenus de qualité.
- **Levier 2 - Structurer le fact-checking et les dispositifs de débunking pour corriger sans renforcer la désinformation.** Populaires et largement déployés, les dispositifs de débunking peuvent limiter la diffusion des fausses informations à condition que les contre-arguments soient rapides, clairs et émis par des sources perçues comme légitimes. Mal calibrés, ils deviennent contre-productifs : en réactivant les infox, ils renforcent leur familiarité ou déclenchent un rejet chez les publics méfiants, alimentant parfois les récits conspirationnistes.
- **Levier 3 - Renforcer la résilience cognitive par des campagnes d’inoculation ciblées.** De plus en plus populaires, les campagnes d’inoculation anticipent les mécanismes de manipulation en exposant le public à des exemples types de désinformation, pour stimuler une vigilance réflexe. Leur efficacité immédiate est reconnue, mais sans réactivation ou si elles sont mal calibrées, elles peuvent devenir contre-productives en renforçant une méfiance généralisée, y compris envers les sources fiables.
- **Levier 4 - Freiner la diffusion impulsive et encourager la réflexion au moment du partage par des nudges comportementaux.** Encore insuffisamment intégrés aux plateformes, les nudges comportementaux visent à déclencher une pause réflexive avant le partage, en rappelant à l’utilisateur de vérifier le contenu ou sa source. Leur effet peut toutefois devenir contre-productif s’ils sont perçus comme intrusifs ou répétitifs, suscitant lassitude ou rejet.

Volet 2 : Les réponses les plus efficaces à la désinformation passent par une transformation structurelle de l’écosystème informationnel, que seule l’action publique peut permettre, en orchestrant à la fois la régulation algorithmique et le renforcement des médias fiables.

- **Levier 5 - Introduire une friction structurelle pour freiner mécaniquement la viralité des contenus douteux.** Inspirée du principe de « Safety by Design », la friction structurelle repose sur des limitations techniques (délais avant partage, plafonds de diffusion) pour ralentir la viralité sans filtrer les contenus. Efficace en contexte de crise, elle peut toutefois devenir contre-productive si elle freine aussi la circulation d’informations fiables, notamment en cas d’application indifférenciée.
- **Levier 6 - Restituer la maîtrise du fil par une transparence algorithmique accrue et une possibilité de maîtriser la constitution de son feed.** Rendre visibles les critères de recommandation et proposer par défaut un fil chronologique permettrait de réduire les manipulations attentionnelles sans recourir à la censure. Pour être efficace, cette réforme structurelle exigerait une volonté politique forte, une harmonisation européenne et une coopération technique avec les plateformes.
- **Levier 7 - Assainir l’espace numérique et valoriser l’information fiable sans restreindre le pluralisme.** L’assainissement du numérique passerait par la suppression des agents de manipulation et la promotion active des sources fiables via des critères transparents d’indépendance et de qualité éditoriale. Cette approche reste toutefois sensible aux faux positifs et dépend d’une gouvernance indépendante et rigoureuse pour éviter toute atteinte au pluralisme.
- **Levier 8 - Assurer un financement stable pour garantir l’indépendance des médias et du service public d’information.** Un financement pérenne des médias d’intérêt général, via subventions ciblées, crédits d’impôt ou contributions obligatoires des plateformes, renforcerait l’indépendance éditoriale et la qualité de l’information, dont la mauvaise qualité peut favoriser la diffusion de la désinformation. Pour être légitime et

efficace, ce soutien devrait être strictement conditionné à des critères vérifiables de transparence, d'éthique et de gouvernance, garantis par des audits indépendants.

- **Levier 9 - Revitaliser le journalisme local pour ancrer l'information de qualité dans les territoires.** Soutenir le journalisme local par des aides ciblées, des incitations fiscales et des modèles coopératifs permettrait de renforcer une information de proximité crédible et adaptée aux réalités locales. Cette stratégie, déjà éprouvée à l'étranger, contribuerait à restaurer la cohésion sociale et l'engagement civique dans les zones délaissées par les grands médias nationaux.

Volet 3 : La réduction des vulnérabilités sociales, économiques et institutionnelles constitue l'approche la plus structurante pour lutter contre la désinformation, en s'attaquant aux facteurs qui nourrissent la réceptivité aux contenus trompeurs - précarité, marginalisation, polarisation et défiance envers les institutions.

- **Levier 10 - Renforcer les politiques sociales, économiques et de santé publique pour réduire la vulnérabilité à la désinformation.** Les inégalités, la précarité et le sentiment d'abandon constituent un terreau fertile pour la désinformation en affaiblissant les repères et la confiance dans les institutions. Des politiques publiques renforcées en matière de protection sociale, de santé mentale, d'éducation et d'emploi contribueraient à restaurer une stabilité propice à l'analyse critique et à la résilience informationnelle.
- **Levier 11 - Réinvestir les espaces de sociabilité pour renforcer la cohésion sociale et réduire la polarisation.** La désinformation prospère sur l'isolement social et les clivages intergroupes ; réactiver les lieux de rencontre (maisons de quartier, tiers-lieux) et organiser des dialogues structurés permettrait de recréer des liens concrets et durables. En favorisant la reconnaissance mutuelle et la circulation apaisée de l'information, ces espaces partagés renforceraient la cohésion sociale et la résilience collective face aux récits clivants.
- **Levier 12 - Renforcer l'intégrité publique perçue pour restaurer la crédibilité des institutions.** La perception d'opacité ou d'impunité, qu'elle soit fondée ou non, alimente la défiance et rend les citoyens plus réceptifs à la désinformation de disqualification des institutions. Des engagements déontologiques lisibles, des sanctions crédibles et des dispositifs de redevabilité ancrés localement contribueraient à restaurer une légitimité perçue essentielle à la lutte contre la désinformation.

Conclusion :

Combattre la désinformation sans en traiter les causes profondes revient à écopier un navire percé. Tant que les politiques publiques se contenteront de réponses réactives, centrées sur les symptômes visibles et ignorantes des dynamiques cognitives, sociales et structurelles à l'œuvre, elles risquent d'aggraver ce qu'elles prétendent corriger. En cause : un modèle implicite, souvent naïf, de la psychologie humaine - un schéma linéaire et individualisant, qui réduit l'adhésion aux contenus trompeurs à un simple déficit d'information ou de rationalité. Ce cadre conduit à des politiques fragmentées, peu efficaces, parfois même contre-productives : certaines interventions mal calibrées peuvent renforcer les croyances erronées, aggraver la défiance, ou freiner la circulation d'informations fiables. Il est temps de repenser la lutte contre la désinformation comme une écologie cognitive intégrée, agissant de concert sur trois niveaux interdépendants : renforcer la résilience individuelle par des mesures fondées sur les sciences comportementales ; assainir l'environnement informationnel en régulant les plateformes et en promouvant des contenus fiables ; et traiter les vulnérabilités sociales qui alimentent la demande en récits trompeurs. Cette approche systémique, fondée sur la rigueur empirique et une vision politique de long terme, seule permettra de restaurer un équilibre démocratique durable dans un paysage informationnel aujourd'hui fragmenté et instable.

Lutter contre la désinformation

Penser autrement l'action publique à l'aune des sciences cognitives

Introduction

La désinformation est aujourd'hui perçue comme un risque stratégique de premier ordre. Le *Global Risks Report 2025* du Forum économique mondial la classe comme la principale menace à court terme, tandis que se multiplient les discours sur une « guerre de l'information » déclarée et l'avènement de la soi-disant "ère de la post-vérité". Dans ce contexte, la France, à l'instar de ses partenaires européens, déploie progressivement un arsenal de politiques publiques visant à contrer cette menace : dispositifs numériques contre les ingérences informationnelles, encadrement juridique des agents d'influence étrangère, ou encore campagnes d'inoculation (*prébunking*). Pourtant, l'écosystème français de lutte contre la désinformation, en cours de structuration, présente trois tendances limitant potentiellement son efficacité :

1. Une approche réactive et opérationnelle, au détriment du stratégique et du systémique :

Portée par des acteurs issus des milieux de la gestion de crise et de la communication, la réponse institutionnelle privilégie des actions immédiates, focalisées sur l'urgence et le **traitement ponctuel** des crises informationnelles. Cette orientation **rend difficile l'élaboration de stratégies de long terme**, capables de traiter les causes profondes et systémiques de la vulnérabilité informationnelle des sociétés démocratiques.

2. Une focalisation sur l'amont de la chaîne de la désinformation, au détriment de l'aval :

Les dispositifs actuels concentrent leurs efforts sur l'identification des producteurs de contenus trompeurs et l'analyse de leur diffusion dans l'espace informationnel, **négligeant la compréhension fine de ce qui se joue après la diffusion** : comment et pourquoi certains contenus modifient-ils les croyances des individus ? Dans quelle mesure ces altérations se traduisent-elles en changements de comportements politiques ou sociaux (par exemple, un vote différent, un changement de "camp" dans un conflit géopolitique) ? Cette invisibilisation de l'aval est d'autant plus paradoxale que la légitimité même des politiques de lutte contre la désinformation repose sur l'hypothèse de tels effets comportementaux. Surtout, elle prive les décideurs de leviers d'action essentiels, alors même qu'il est impossible, dans des sociétés ouvertes et démocratiques, de prévenir totalement la pénétration de narratifs extérieurs : la question n'est pas seulement d'empêcher leur circulation, mais surtout d'en limiter l'impact.

3. Le chaînon manquant cognitif entre approches cyber et analyses macro-sociales :

Les stratégies actuelles concentrent leurs investissements, d'un côté, sur les sciences du numérique, de la captologie et de la cybersécurité, de l'autre, sur les analyses macrosociales et géopolitiques, **sans**

qu'existe de véritable interface pour articuler ces deux niveaux d'analyse. Cette étanchéité résulte en partie d'une dépendance historique au sentier institutionnel français et d'une culture administrative et stratégique centrée sur une lecture géopolitique des menaces. Faute d'un investissement suffisant dans les sciences empiriques - en particulier les sciences cognitives-, le chaînon analytique permettant de relier la dynamique de production technologique de l'information aux mécanismes de réception, de traitement et d'intégration des contenus par les individus fait défaut. Cette lacune conduit à l'adoption de dispositifs fondés sur des hypothèses implicites, rarement explicitées ou validées empiriquement, et issues de modèles linéaires de la communication ou de paradigmes behavioristes : des croyances sont ainsi imputées à de simples expositions, et des relations de causalité non vérifiées sont supposées entre exposition et adhésion.

Le présent document vise donc à fournir aux acteurs français de la lutte contre la désinformation une synthèse claire et accessible, en langue française, de l'état de la recherche en sciences cognitives sur l'adhésion, la diffusion et les effets comportementaux de la désinformation. Élaboré par un collectif de scientifiques issus des principales disciplines constitutives des sciences cognitives - psychologie, intelligence artificielle, évolution culturelle, philosophie et linguistique - il apporte un éclairage spécifique sur l'aval de la chaîne de la désinformation, afin d'éclairer les leviers d'action une fois celle-ci présente dans l'espace informationnel national. Sur la base des modèles théoriques et des résultats empiriques les plus récents, il propose des recommandations concrètes pour orienter les dispositifs opérationnels et stratégiques de lutte contre ses effets.

À titre d'exemple, la diffusion massive de contenus de désinformation pro-Trump et anti-Clinton durant l'élection présidentielle de 2016 a souvent été perçue comme un facteur déterminant dans la victoire de Donald Trump. Toutefois, une telle corrélation n'implique pas nécessairement un lien de causalité direct. Ce rapprochement pourrait, en réalité, découler d'un troisième ensemble de causes sous-jacentes, rarement prises en compte dans les analyses. En effet, des facteurs comme le sentiment d'abandon par les élites, la précarité croissante ou encore la défiance envers les institutions peuvent engendrer à la fois une inclination au vote populiste et une plus grande réceptivité aux contenus de désinformation. Lorsqu'un individu se sent trahi par le système, lorsqu'il nourrit déjà une aversion pour une figure politique, il sera plus enclin à partager des contenus qui renforcent sa position - qu'il les pense vraies ou non. Autrement dit, c'est cette constellation de ressentiments socio-politiques qui pourrait expliquer à la fois le soutien électoral à Trump et la diffusion virale de récits mensongers en sa faveur.

1. La lutte contre la désinformation repose sur un modèle implicite et simpliste de la psychologie humaine parfois contre-productif.

En raison du "chaînon manquant cognitif" et du relatif déficit d'intégration des sciences empiriques, la plupart des initiatives de lutte contre la désinformation reposent aujourd'hui sur un modèle implicite de la psychologie humaine, dit de psychologie naïve (*folk psychology*). Ce modèle postule souvent une chaîne causale simpliste et mécanique : **l'exposition à une fausse information façonnerait mécaniquement des croyances, qui détermineraient ensuite les comportements.** Intuitif et séduisant, ce cadre alimente des politiques focalisées sur la correction des erreurs individuelles - *fact-checking, formations à l'esprit critique, détection des producteurs de fake news* - en supposant **qu'il suffirait de rétablir la vérité pour endiguer la désinformation et amoindrir ses effets négatifs.**

Or, cette lecture réductrice néglige l'influence décisive des contextes sociaux, des vulnérabilités économiques, de la défiance institutionnelle et des logiques identitaires. Elle ignore également que la désinformation peut parfois agir sans même convaincre, en modifiant la perception des normes sociales et en répondant à des besoins d'appartenance et de reconnaissance. À défaut d'une approche systémique, fondée sur les apports des sciences cognitives et sociales, ces stratégies resteront largement inefficaces face aux racines structurelles de la vulnérabilité informationnelle.

À rebours de ce modèle de psychologie naïve, nous présenterons donc **quatre grands principes de cadrage issus du consensus scientifique sur le fonctionnement de la cognition humaine**, avant d'analyser les effets potentiellement contre-productifs qu'entraîne l'adoption d'un modèle

simplificateur dans la conception des politiques publiques.

1.1. La désinformation n'est pas (seulement) un problème d'irrationalité et de crédulité

La désinformation est souvent interprétée comme le produit d'un déficit de rationalité ou d'une crédulité excessive. Cette lecture intuitive structure de nombreux discours et réponses institutionnelles - de l'enseignement de l'esprit critique à la prolifération des dispositifs de fact-checking - fondées sur l'idée qu'il suffirait de corriger des erreurs de raisonnement pour tarir la diffusion des fausses nouvelles. Or, si les limites de certaines de nos capacités cognitives jouent bien un rôle, elles n'expliquent qu'une partie du phénomène. Croire ou partager de la désinformation ne tient pas seulement à une prétendue incapacité à distinguer le vrai du faux, mais répond aussi à des motivations sociales, émotionnelles et identitaires. Loin d'être de simples erreurs épistémiques, ces croyances traduisent souvent une manière de se situer dans le monde et d'interagir avec les autres : elles peuvent être une réponse fonctionnelle découlant d'une méfiance face à des institutions perçues comme corrompues, un besoin de cohérence dans un environnement ressenti comme chaotique, ou encore une sensibilité exacerbée à la menace due à certaines caractéristiques de l'environnement.

Dans ce cadre, **la véracité d'une information importe parfois moins que l'usage social que permet sa diffusion.** On exprime des idées erronées par erreur, mais aussi pour séduire, provoquer, rassurer ou impressionner, parfois bien plus que parce qu'on y adhère. Par exemple, certains affichent publiquement leurs croyances comme des signaux d'appartenance. **Plus ces croyances sont coûteuses socialement** - croire à une Terre plate ou à des élites pédo-satanistes - **plus elles pourraient renforcer la loyauté perçue au groupe**, jusqu'à « brûler les ponts » avec le monde extérieur.

Ces croyances semblent particulièrement attrayantes pour des individus isolés, précarisés ou défiants vis-à-vis des institutions : en relayant ces récits, ils trouvent un espace d'appartenance, du sens face à l'incertitude, et parfois même un soutien matériel. En ce sens, croire et propager de la désinformation manifeste une sorte de "rationalité" sociale. Dès lors, lutter contre la désinformation exige plus qu'une correction des faits : il faut comprendre et traiter les besoins sociaux et psychologiques auxquels ces récits offrent des réponses.

1.2. La désinformation n'est qu'un aspect relativement marginal d'une crise informationnelle plus vaste.

La désinformation - la diffusion délibérée de fausses informations - et les fake news monopolisent l'attention médiatique et politique. Pourtant, elles ne constituent qu'un aspect marginal d'une crise informationnelle bien plus vaste, trop souvent réduite à ces seules questions. En France, la désinformation représente ainsi seulement 0,16 % de la consommation médiatique, et sa diffusion est très concentrée : lors d'une étude menée en 2021, 61 % des participants français n'ont consulté aucune source peu fiable durant les 30 jours de l'étude, et une petite minorité était responsable de l'essentiel des contenus trompeurs vus et partagés en ligne. De même, lors de l'élection présidentielle américaine de 2016, pourtant emblématique de l'influence supposée des fake news, celles-ci ne représentaient que 0,15 % du régime médiatique américain global (et moins de 1% des nouvelles au sens strict). Leur consommation était en outre extrêmement concentrée : 75 % des électeurs n'ont consulté aucun site de fake news, et les 10 % des internautes avec la diète médiatique la plus "conservatrice" (républicains) ont généré à eux seuls 60 % du trafic associé. Le même phénomène s'observe durant la pandémie de COVID-19, où 65 % des contenus anti-vaccins diffusés sur Facebook provenaient de seulement douze comptes. Par ailleurs, la majorité de la population s'informe peu sur les affaires publiques : quand elle le fait, elle privilégie

les médias traditionnels, et l'on surestime largement le rôle des réseaux sociaux dans la formation des opinions politiques et scientifiques. Surtout, la désinformation circule avant tout dans des sphères marginales, **touchant des individus déjà prédisposés à y adhérer** du fait de leurs orientations idéologiques, de traits psychologiques particuliers ou d'identités sociales marquées. **Son impact comportemental à l'échelle de la population reste donc limité.**

En réalité, le mensonge explicite est rare, coûteux et souvent inutile. La désinformation n'est que la partie émergée de l'iceberg : la plus visible, la plus facile à désigner, mais ni la plus répandue, ni la plus dangereuse. Les manipulations les plus fréquentes reposent moins sur des contre-vérités que sur des procédés bien plus subtils : sélection, interprétation biaisée ou omission stratégique de faits - autant de pratiques largement répandues, y compris dans les médias traditionnels ou les conversations ordinaires.

Imaginons par exemple une société de communication chargée d'influencer un référendum sur le Brexit. Exploitant les outils de ciblage publicitaire de Facebook, elle peut segmenter les audiences en fonction des groupes et pages que les individus ont likés. Aux défenseurs de la cause animale, elle diffuserait exclusivement des contenus mettant en avant les projets européens facilitant les tests pharmaceutiques sur les animaux ; aux passionnés de chasse, elle présenterait uniquement des propositions européennes restreignant la pratique de la chasse. Aucune information fautive n'est diffusée, et pourtant, nous sommes ici bien en présence d'une forme de manipulation.

Réduire la crise informationnelle à la seule question des fake news et du mensonge explicite est non seulement erroné, mais dangereux du point de vue des politiques publiques. **Cette focalisation, aussi tentante que spontanée, a l'avantage trompeur d'incarner la menace sous une forme claire et intuitive, et de faire croire à des solutions immédiates, peu contestables**

sur le plan éthique ou juridique - fact-checking, debunking, suppression de contenus - mais en réalité insuffisantes et strictement réactives, incapables de corriger en profondeur les causes structurelles de la désinformation. Elle détourne ainsi l'attention d'autres enjeux tout aussi déterminants : la malinformation - c'est-à-dire la diffusion d'informations, même exactes, mais utilisées pour nuire, manipuler ou tromper - et la dégradation continue de l'écosystème informationnel. Pourtant, la notion même de désinformation reste mal définie : elle trouve ses racines dans la doctrine soviétique de *desinformatsiya* (1923), qui ne visait pas seulement la fausseté factuelle, mais tout contexte informationnel destiné à produire de l'ignorance, de la confusion et une incertitude radicale quant aux sources. La crise actuelle des autorités épistémiques traditionnelles, conjuguée à l'effondrement généralisé de la confiance, alimente des dynamiques de désinformation qui débordent largement le seul contenu des messages. En ne voyant dans cette crise qu'un enchaînement de mensonges isolés, on risque de minimiser la gravité systémique et de se priver des leviers d'action réellement efficaces : réformer les plateformes numériques, revitaliser un journalisme de qualité, encadrer les algorithmes, et reconstruire un environnement informationnel capable de restaurer la confiance et de garantir l'accès à des contenus fiables.

1.3. La désinformation n'est lisible qu'à travers le prisme de l'environnement dans lequel elle se déploie

L'efficacité de la désinformation est souvent analysée de manière désincarnée, sans considération pour les environnements dans lesquels elle prend racine. L'adhésion ou le rejet d'une fake news est alors attribué à des différences de compétences cognitives ou de traits de personnalité, comme si ces différences relevaient d'une simple disposition individuelle. Or, ces croyances peuvent également s'enraciner dans des réponses adaptatives à des

environnements sociaux, informationnels et matériels différents. Adopter une croyance erronée marginale expose à des coûts sociaux élevés; l'adopter lorsqu'elle est largement partagée dans son groupe offre potentiellement, au contraire, reconnaissance et appartenance. Loin d'être irrationnelles, ces dynamiques relèvent souvent de calculs inconscients de positionnement social et de bénéfices symboliques.

De plus, lorsque l'environnement est pris en compte, il est trop souvent réduit à sa seule dimension numérique : algorithmes de recommandation, viralité des contenus, ou décisions éditoriales des grandes plateformes. L'étude de ces éléments est essentielle, mais l'écosystème informationnel dépasse largement Internet. La concentration des médias, le déclin progressif de la presse de proximité, et l'érosion du pluralisme contribuent tout autant à fragiliser l'espace public. S'y ajoutent d'autres dimensions environnementales, par exemple des facteurs sociaux : l'isolement relationnel, l'homogamie sociale et idéologique et la faible exposition à la diversité des points de vue alimentent des chambres d'écho bien réelles, hors des réseaux numériques. Enfin, les environnements matériels et institutionnels - précarité économique, défiance envers les institutions, sentiment d'abandon - structurent en profondeur les vulnérabilités face à la désinformation, comme le développe la seconde partie de ce rapport. Penser ces enjeux impose ainsi de dépasser le seul prisme du numérique pour appréhender l'ensemble des contraintes sociales, économiques et médiatiques qui façonnent l'espace de décision cognitive des individus.

1.4. L'adhésion à la désinformation engage des processus cognitifs et comportementaux multiples et différenciés.

L'adhésion à la désinformation est souvent pensée comme un acte binaire et stable : on serait « complotiste » ou non, trompé par une fake news ou doté d'un esprit critique efficace. Cette simplification, très présente dans les discours publics et les mesures (sondages évaluant le « taux de complotisme » ou le nombre de personnes affirmant « que la Terre est plate »), masque la complexité réelle des mécanismes d'adhésion et le caractère contextuel et relatif du statut épistémique des informations.

La notion même de « croyance » recouvre d'abord plusieurs réalités cognitives distinctes. On distingue ainsi les croyances intuitives, souvent acceptées sans examen conscient et pouvant guider immédiatement l'action (par exemple, la certitude que mon pied ne traversera pas la prochaine marche d'escalier), des croyances réflexives, affirmées publiquement mais sans incidence directe sur les comportements (par exemple, affirmer que la Terre est plate).

Cette distinction explique pourquoi certains dénoncent publiquement des complots mondiaux supposément protégés par des polices secrètes, tout en discutant librement de ces théories sur internet ou lors de conférences, sans adopter aucune précaution. À l'inverse, dans des pays où une police secrète existe réellement, les individus convaincus de son existence évitent soigneusement toute critique publique et affichent plutôt leur loyauté.

Ce contraste montre que la même croyance apparente peut entraîner des comportements opposés selon qu'elle relève d'une conviction intuitive agissante ou d'une affirmation réflexive à portée sociale, même lorsque l'individu est authentiquement convaincu dans les deux cas.

C'est par exemple le cas des attitudes d'hésitation vaccinale : beaucoup de Français se sont fait vacciner malgré leurs réserves, suggérant que cette position pouvait parfois relever moins d'une réelle conviction que d'une attitude protestataire.

Par ailleurs, croire une information et la partager ne sont pas nécessairement liés. Une part significative de la mésinformation sur les réseaux sociaux circule par provocation, dérision ou parce que son contenu est intéressant, indépendamment de sa véracité : l'information est diffusée non pour sa crédibilité mais pour l'ampleur qu'aurait ses conséquences si elle était vraie. À l'inverse, certains adhèrent sincèrement à des récits trompeurs, mais s'abstiennent de les diffuser, par crainte du discrédit social et sachant bien que la plupart les tournerait en ridicule s'ils affirmaient, par exemple, que la Terre est plate. Or, cette distinction est ignorée par les approches traditionnelles qui évaluent l'impact de la désinformation à partir des indicateurs de visibilité (likes, partages). Non seulement ces métriques ne mesurent ni les croyances réelles ni les comportements effectifs, mais elles peuvent aussi induire une estimation trompeuse de l'influence réelle de la désinformation, d'autant plus que la consommation d'information via les réseaux sociaux demeure minoritaire comparativement aux médias traditionnels et aux échanges directs entre individus.

Enfin, l'existence d'une croyance ne préjuge en rien de sa traduction en action. La littérature scientifique qualifie cet écart structurel de « fossé entre les croyances et les comportements » (*Belief-Action Gap*) soulignant qu'admettre une vérité n'implique ni l'engagement ni le renoncement qu'elle exigerait.

Si croire suffisait à agir, la crise climatique serait surmontée depuis longtemps: bien que la majorité des Français reconnaisse l'urgence environnementale, trop peu sont ceux qui modifient concrètement leurs habitudes de consommation ou de mobilité.

Il en va de même pour la désinformation : l'adhésion sincère à certains récits n'entraîne pas mécaniquement des comportements alignés sur ces convictions. Cette dissociation impose de dépasser l'illusion des croyances déclarées pour **concentrer l'action publique sur l'observation rigoureuse des comportements effectifs**, seul moyen d'évaluer avec précision les risques et de calibrer des réponses adaptées, au-delà des postures et des intentions affichées.

1.5. L'utilisation du modèle naïf de la psychologie de la désinformation a des effets potentiellement délétères.

Une représentation naïve de la psychologie de la désinformation n'est pas seulement inexacte : elle peut conduire à **des erreurs majeures dans la conception des politiques publiques** en négligeant la diversité des canaux par lesquels la propagande agit.

La désinformation ne produit pas toujours ses effets en modifiant directement les croyances, mais en envoyant des signaux qui transforment les perceptions des normes sociales et des conséquences attendues des comportements.

L'exemple de la propagande nazie est éclairant: l'intensification des campagnes antisémites n'a que peu converti et a parfois renforcé leur défiance envers le régime. En revanche, pour les antisémites déjà convaincus, cette propagande a agi comme un signal institutionnel indiquant que l'État ne sanctionnerait plus les actes hostiles envers les Juifs. Ce n'est donc pas seulement la persuasion, l'augmentation des croyances antisémites qui a déclenché la hausse des violences, mais la conviction qu'il était désormais socialement et juridiquement « sûr » de passer à l'acte.

Une contre-propagande efficace ne devait pas simplement réfuter les croyances, **mais restaurer la perception d'une norme sociale et institutionnelle claire** en rendant visibles, par exemple, les sanctions appliquées

contre ces comportements, ou le fait que la majorité de la population les suit.

Plus largement, la désinformation peut agir non seulement en changeant directement les convictions individuelles, mais en modifiant la perception de ce que croient les autres. **Elle influence tant ce qu'un individu croit que ce qu'il croit que les autres croient.** Or, cette représentation des croyances majoritaires détermine les coûts et bénéfiques sociaux de l'adhésion : une opinion perçue comme marginale expose à la stigmatisation, mais peut devenir valorisante si elle est perçue comme dominante. C'est ainsi que des récits marginaux **peuvent s'imposer non par leur plausibilité, mais par l'illusion de leur large diffusion.** Ce levier, central dans la dynamique des normes sociales, invite à repenser les stratégies publiques : plutôt que de se concentrer uniquement sur la correction des croyances, il est **crucial d'agir sur la perception de leur prévalence**, en rendant visibles les opinions réellement majoritaires et les comportements effectivement valorisés.

Cette réduction du problème à un modèle de psychologie naïve peut également favoriser des effets rebond. Les politiques publiques pourraient non seulement échouer, mais aggraver la situation qu'elles entendent corriger. Le **débunking** est un premier exemple de potentiel effet rebond : **en répétant une information erronée pour la réfuter, on en augmente mécaniquement la familiarité** avec celle-ci. Surtout, ce risque est exacerbé lorsque la correction émane d'une source perçue comme illégitime ou antagoniste par les publics cibles. Dans ces situations, loin de favoriser la réévaluation cognitive, la correction est interprétée à travers le prisme d'une défiance préexistante et vient renforcer le schéma narratif de la croyance initiale. Par exemple, un individu adhérant à l'idée que « les médias sont contrôlés par des élites occultes » percevra toute tentative de debunking émanant de ces mêmes médias comme une manœuvre de dissimulation, renforçant ainsi sa conviction initiale. Or, ces individus sont une des cibles principales du message. Ce mécanisme illustre le rôle central des

dynamiques identitaires et de la crédibilité perçue des sources dans l'efficacité - ou l'échec - des corrections informationnelles.

De même, le **prebunking** - ou inoculation - vise à renforcer la vigilance face à la désinformation en exposant préalablement les individus à des exemples de techniques manipulatoires. Cette approche montre des effets positifs sur la capacité à détecter les fake news, en rendant les individus plus attentifs aux signes de contenus douteux. Toutefois, plusieurs études soulignent un effet secondaire potentiel : cette vigilance accrue ne s'exerce pas toujours de manière sélective. **Elle peut également conduire à une forme de méfiance généralisée**, réduisant la confiance non seulement dans les informations trompeuses, mais aussi dans des contenus fiables et de qualité. Ce déplacement du "seuil de vigilance" n'est pas sans conséquence : en diminuant la confiance dans les sources légitimes, il pourrait contribuer, de manière involontaire, à renforcer la défiance envers les institutions et les médias traditionnels. Or, cette défiance constitue l'un des principaux facteurs de vulnérabilité face à la désinformation et aux récits complotistes. Ces résultats n'invalident pas l'intérêt du prebunking, mais soulignent la nécessité d'en calibrer soigneusement les effets pour éviter qu'une intervention conçue pour protéger les esprits critiques n'alimente indirectement la dynamique même qu'elle cherche à contenir.

2. Il faut élargir le cadrage des causes de la désinformation : au-delà des campagnes de manipulation délibérées, elle prospère sur des vulnérabilités individuelles et des facteurs structurels.

Ce *Rapport* ne traite pas de l'amont de la chaîne de la désinformation: des stratégies géopolitiques ou des campagnes coordonnées de désinformation, déjà largement étudiées. Il s'intéresse à un angle complémentaire et encore trop marginal dans l'action publique : les facteurs cognitifs, sociaux et contextuels qui favorisent l'adhésion à ces récits et leur diffusion. Comprendre ces déterminants est essentiel pour concevoir des politiques plus efficaces, ciblant non seulement les contenus, **mais aussi les vulnérabilités sur lesquelles ils prospèrent**. L'analyse portera d'abord sur les facteurs individuels, avant d'aborder les déterminants structurels et environnementaux. L'accent sera mis sur les leviers environnementaux, plus directement accessibles à l'action publique que certains composants individuels comme les traits de personnalité.

2.1. Les causes individuelles

Concrètement, **quels facteurs socio-démographiques, traits de personnalité, capacités cognitives et dimensions identitaires prédisposent à l'adhésion à la mésinformation ?** Bien que ces facteurs présentent une forte variabilité selon les contextes géographiques, le type d'information ou de théorie du complot, et les environnements sociaux, certains déterminants sont récurrents.

Les facteurs socio-démographiques

Plusieurs variables socio-démographiques exercent une influence stable mais modérée sur la vulnérabilité à la désinformation. Par exemple, **le niveau d'éducation est le facteur le plus constant :**

l'augmentation du niveau d'études diminue l'adhésion à la désinformation. L'âge suit une courbe non linéaire : la Génération Z (nés entre 1997 et 2012) est plus exposée, tandis qu'une hausse de la vulnérabilité est également observée après 65 ans. **Le genre présente des effets différenciés** : les hommes réussissent en moyenne mieux aux tests de discernement mais surestiment davantage leurs compétences, tandis que les femmes et personnes non-binaires obtiennent des scores légèrement inférieurs mais évaluent plus justement leurs capacités. **Enfin, la précarité économique est associée à une plus grande vulnérabilité**, corrélée à une moindre confiance institutionnelle et à une exposition accrue à des réseaux sociaux homogènes idéologiquement. Si ces effets désignent des groupes particulièrement exposés - jeunes adultes peu diplômés, économiquement fragiles et socialement isolés - les effets socio-démographiques restent individuellement de faible amplitude.

Les traits de personnalité

Les traits de personnalité constituent des prédictors robustes mais différenciés de la vulnérabilité à la désinformation. Les individus présentant de faibles niveaux de **conscienciosité** (manque de rigueur, faible capacité d'autocontrôle) et d'**agréabilité** (faible prédisposition à la coopération et à la confiance) sont plus enclins à accepter et diffuser des récits trompeurs. Une forte extraversion favorise également la diffusion, en raison d'une plus grande propension à partager des contenus, indépendamment de leur véracité. Ces tendances sont amplifiées par les traits de la "**Triade sombre**" - narcissisme, machiavélisme et psychopathie - qui prédisposent à des comportements manipulateurs et à une recherche de domination symbolique à travers des récits polarisants. À un niveau plus spécifique, un dogmatisme élevé (adhésion rigide à ses croyances) et un fort besoin d'unicité (volonté de se distinguer du groupe majoritaire) augmentent la perméabilité aux discours conspirationnistes. À l'inverse, une estime de soi individuelle bien calibrée agit comme un facteur protecteur, en réduisant la

nécessité de recourir à des explications simplificatrices pour renforcer l'image de soi. Ces facteurs, relativement stables à l'échelle individuelle, **ne constituent ni une cible prioritaire ni un levier d'action pertinent** pour les politiques publiques, tant pour des raisons pratiques que normatives.

Les facteurs cognitifs, épistémiques et émotionnels

Au niveau cognitif, l'adhésion à la désinformation est d'abord liée à des **facteurs épistémiques**. Une faible numératie, une moindre capacité de raisonnement réflexif et un faible besoin de cognition augmentent la crédulité face aux contenus trompeurs. Les individus s'appuyant davantage sur l'intuition que sur l'analyse, et ceux présentant une forte réceptivité aux discours creux (*bullshit receptivity*), sont particulièrement vulnérables. Des biais cognitifs spécifiques renforcent cette tendance, comme la perception illusoire de motifs dans des événements aléatoires, la détection exagérée d'intentions cachées (illusory agency detection) ou l'attribution d'intentions humaines à des entités non humaines (anthropomorphisme). Enfin, la pensée activement ouverte, définie comme la disposition à considérer des arguments contraires et à réviser ses opinions, agit comme un puissant facteur de protection, tout comme un niveau d'intelligence plus élevé.

Sur le plan émotionnel, des **états affectifs durables** augmentent également la perméabilité aux récits trompeurs. La perception d'un monde dangereux, le cynisme social et le sentiment d'impuissance forment un terreau psychologique propice à la réception de récits conspirationnistes ou extrêmes. **L'anxiété** - y compris face à la mort -, le besoin de clôture cognitive et l'intolérance à l'ambiguïté renforcent cette dynamique, en favorisant la recherche de récits simples et rassurants. À l'inverse, un fort sentiment de contrôle perçu réduit cette vulnérabilité, tandis que l'efficacité personnelle n'apparaît pas comme un prédicteur significatif. Ces facteurs émotionnels et existentiels jouent un rôle

dans l'adhésion à la désinformation, en créant un climat mental propice à l'acceptation de récits explicatifs, même au détriment de leur plausibilité.

Les facteurs identitaires

Sur le plan identitaire et social, l'adhésion à la désinformation est étroitement liée aux dynamiques d'appartenance et de polarisation. Les individus évaluent les informations en fonction de leur compatibilité avec les croyances de leur groupe. Ce biais est renforcé par les mécanismes d'opposition *ingroup/outgroup*, qui favorisent la crédibilité des contenus valorisant son propre groupe et discréditent ceux favorables à l'*outgroup*. L'orientation politique joue ainsi souvent un rôle majeur : dans plusieurs pays, plus les individus se positionnent à droite de l'échiquier politique, plus leur vulnérabilité à la désinformation est élevée et plus ils surestiment leur capacité à discerner le vrai du faux. La susceptibilité à la désinformation corrèle également avec des niveaux accrus d'autoritarisme et d'orientation à la dominance sociale. Ces effets, bien documentés aux États-Unis, en Irlande, en Espagne et au Mexique, varient cependant selon les contextes nationaux. Des expériences montrent par exemple qu'en Hongrie, la susceptibilité à la désinformation est davantage liée au soutien au gouvernement en place qu'à l'orientation politique stricte, tandis qu'en Ukraine, elle dépend du degré de perception du conflit intergroupes. De plus, sur le plan groupal, un narcissisme collectif prononcé et une fierté identitaire élevée favorisent l'adhésion à des récits valorisant le groupe d'appartenance, souvent au détriment d'extragroupes perçus comme menaçants. Enfin, la défiance institutionnelle et sociale, mesurée par de faibles niveaux de confiance, constitue un prédicteur transversal, tout comme les sentiments d'anomie et d'aliénation, qui renforcent la recherche de récits explicatifs alternatifs. Ces facteurs identitaires et sociaux agissent comme des filtres interprétatifs puissants, structurant l'évaluation de l'information en fonction des appartenances plus qu'en fonction de critères épistémiques.

2.2. Les causes liées à l'écologie informationnelle

Les causes et comportements individuels liés à la désinformation sont souvent des réponses à des environnements sociaux, économiques et informationnels dégradés. Il est donc essentiel d'analyser aussi les causes structurelles profondes de cette crise de l'information.

L'émergence des médias numériques et des réseaux sociaux

L'écologie informationnelle numérique contemporaine constitue un facteur central de la crise de l'information, en remodelant en profondeur les modalités de production, de circulation et de consommation des contenus. Dominé par des plateformes dont les algorithmes sont conçus pour maximiser leurs revenus - et donc l'engagement - cet écosystème favorise systématiquement les contenus émotionnels, sensationnalistes et polarisants. Ces logiques économiques, fondées sur l'économie de l'attention, tendent à enfermer les utilisateurs dans des bulles de filtre, limitant leur exposition à des points de vue divergents (même si des formes de chambres d'écho émergent également de manière organique - et parfois plus intensément encore - dans les interactions sociales hors ligne). **La viralité des contenus repose non sur leurs qualités épistémiques, mais sur des caractéristiques affectives et sociales** : tonalité négative, forte charge émotionnelle, simplicité narrative et cadrage ingroup/outgroup. Cette dynamique alimente la polarisation, fragilise la confiance dans les institutions et les médias, et accélère la circulation de contenus de faible qualité, en particulier en période de crise.

Par ailleurs, la configuration même des plateformes - **opacité des algorithmes, absence de contrôle utilisateur sur la personnalisation des flux, et logique de propagation sans friction** - participe à une amplification systémique de la crise de l'information. Cette diffusion sans friction, permise par des interfaces réduisant à l'extrême le coût d'action (un simple clic),

permet à un contenu d'atteindre très rapidement des audiences massives et hétérogènes, bien au-delà des cercles sociaux traditionnels. Elle introduit ainsi une asymétrie radicale entre l'effort requis pour produire et partager un contenu et l'ampleur potentielle de sa diffusion dans des délais extrêmement courts. La difficulté pour les utilisateurs de comprendre ou de paramétrer les mécanismes qui sélectionnent les contenus qu'ils voient alimente la méfiance. À cela s'ajoutent l'exploitation des bots et des comportements inauthentiques par des acteurs mal intentionnés, qui simulent des consensus artificiels, et le micro-ciblage politique opaque, permettant de diffuser des messages adaptés à chaque segment de l'électorat en échappant à tout contrôle démocratique. Enfin, les biais algorithmiques tendent à reproduire et renforcer les discriminations sociales et l'exclusion des groupes minoritaires, contribuant ainsi à l'aggravation des inégalités informationnelles.

La dégradation de l'environnement médiatique traditionnel

La révolution numérique et l'attention portée à la désinformation ne doivent pas occulter le rôle central des médias traditionnels. Aux États-Unis, l'information télévisée reste cinq fois plus consultée que l'information en ligne, les réseaux sociaux n'étant que rarement une source unique et relayant majoritairement des contenus issus des médias mainstream. De surcroît, ceux qui s'informent principalement via ces plateformes sont aussi les moins engagés politiquement. En France, cette dynamique s'accompagne d'une érosion marquée de la confiance : seuls 32 % des citoyens déclarent encore faire confiance aux médias d'information, 37 % à l'information télévisuelle ou radiophonique, 36 % à la presse écrite, et plus de la moitié estiment qu'une large part des informations diffusées par ces médias est erronée, alimentant ainsi un recours croissant à des canaux alternatifs, souvent plus biaisés. Ce phénomène s'inscrit dans une dégradation plus large de l'environnement journalistique : la part des pays de l'OCDE considérés

comme offrant un environnement « bon » pour la presse a chuté de 49 % en 2015 à 21 % en 2023, et de 21 % à 4 % à l'échelle mondiale. Enfin, seuls 38,8 % des répondants à l'enquête sur la confiance dans les institutions publiques déclaraient, dès 2021, avoir confiance dans les médias, une tendance qui s'est encore dégradée depuis. **On impute ainsi souvent aux fake news des effets qui relèvent en réalité de la dégradation perçue de la qualité informationnelle des médias traditionnels**, la confiance dans ces derniers et dans les institutions constituant l'un des premiers facteurs prédisant la vulnérabilité face à la désinformation et aux récits complotistes.

Cette perception de la dégradation de l'environnement informationnel n'est pas uniquement spéculative. Trois dynamiques structurelles y contribuent effectivement : **la concentration capitaliste des groupes de presse et d'édition, l'érosion du modèle économique traditionnel et la fragmentation des audiences**. En France, cette concentration atteint un niveau inédit : onze milliardaires contrôlent aujourd'hui 80 % de la presse quotidienne généraliste, 60 % de l'audience télévisuelle et la moitié du marché radiophonique. Si l'impact de cette configuration sur le pluralisme éditorial reste difficile à quantifier, elle favorise des stratégies de captation de l'attention fondées sur des logiques commerciales et d'influence, souvent au détriment du pluralisme éditorial et de l'investigation. Sur le plan économique, les recettes publicitaires de la presse écrite ont chuté de 50 % en une décennie, sous l'effet de la captation croissante des revenus par les grandes plateformes numériques. En France, ces dernières concentrent désormais 68 % du marché publicitaire digital, imposant leurs standards de rentabilité et de performance. Cette pression économique a entraîné une réduction continue des moyens consacrés à la production de l'information, avec une baisse significative des effectifs de journalistes dans la presse écrite et une montée en puissance des contenus produits à partir de dépêches d'agences et de communiqués institutionnels.

Parallèlement, la transition numérique de la presse traditionnelle reste marquée par une forte précarité des modèles économiques : les taux de conversion des lecteurs gratuits en abonnés payants demeurent faibles, et le coût d'acquisition d'un abonné numérique continue d'absorber une part significative des recettes issues des abonnements. Ce déséquilibre incite les rédactions à privilégier les contenus à fort potentiel de viralité, souvent émotionnels ou polémiques, au détriment des formats longs et des enquêtes de fond. Enfin, la fragmentation accélérée des audiences complique encore la mission des médias traditionnels : l'offre informationnelle se répartit désormais entre une multitude de canaux, et les pratiques se polarisent selon les générations et les profils socioculturels. Faute de pouvoir maintenir un lien de confiance stable avec des publics éclatés et défiants, les médias traditionnels voient leur capacité à structurer l'espace public et à offrir une information de qualité, vérifiée et pluraliste, progressivement affaiblie, laissant ainsi un terrain favorable à l'expansion des "médias alternatifs".

2.3. La crise de l'information trouve ses racines dans des facteurs sociaux, institutionnels et matériels profondément ancrés.

La prolifération de la désinformation ne constitue pas seulement une pathologie du champ informationnel ; elle est d'abord le symptôme d'un malaise social plus profond, enraciné dans la dégradation des conditions sociales, matérielles et économiques dans un environnement de plus en plus polarisé.

L'environnement social

La désinformation trouve un terrain particulièrement fertile **dans la solitude sociale et psychologique**, qui affaiblit les défenses cognitives et renforce la quête de récits explicatifs face à un monde perçu comme chaotique. Là où les liens communautaires se délitent et les formes traditionnelles d'engagement disparaissent, l'individu isolé compense son sentiment de

perte de contrôle en adhérant à des narratifs simplificateurs qui offrent à la fois sens et appartenance. Cette dynamique est solidement étayée : une étude longitudinale montre que la solitude, notamment lorsqu'elle s'installe dès l'adolescence, prédit fortement l'adhésion ultérieure à des théories complotistes. Ce mécanisme s'est exacerbé durant la pandémie de COVID-19, où l'isolement forcé a intensifié la vulnérabilité aux récits complotistes, en particulier chez les individus présentant déjà des tendances paranoïaques légères. Mais au-delà de cette conjoncture, l'isolement résulte de tendances sociales structurelles : déclin des solidarités locales, multiplication des foyers monopersonnels, éclatement des réseaux de proximité et raréfaction des espaces de sociabilité. Dans ce contexte, **la désinformation n'est pas un simple produit de l'ignorance, mais une réponse à la souffrance sociale, comblant le vide relationnel et existentiel par des récits qui réintroduisent cohérence et reconnaissance.**

De même, la désinformation prolifère plus intensément dans les environnements sociaux caractérisés par une forte polarisation, où les appartenances idéologiques structurent l'évaluation des faits et réduisent la réceptivité aux informations correctives. Ce phénomène, bien documenté en psychologie politique, repose sur une dynamique de polarisation affective : les individus développent des affects négatifs intenses envers les groupes opposés et valorisent leur propre camp au point de privilégier systématiquement les récits qui renforcent leur identité collective, même lorsqu'ils sont manifestement faux. Cette distorsion cognitive, amplifiée chez les individus à forte implication politique, **transforme l'adhésion à certains récits en acte de loyauté sociale, neutralisant les mécanismes de correction factuelle.** La polarisation partisane et affective est en progression en France, favorisant ainsi la diffusion de récits désinformants qui exploitent ces lignes de fracture. **Plus la société se fragmente autour de clivages idéologiques et identitaires, plus l'espace public devient perméable à la désinformation.**

L'environnement institutionnel

La dégradation de la confiance institutionnelle constitue un facteur central de vulnérabilité à la désinformation. En France, seuls 34 % des citoyens déclarent faire confiance à leur gouvernement, un niveau significativement inférieur à la moyenne des pays de l'OCDE (39 %), traduisant une perception largement répandue d'inefficacité et d'éloignement des élites politiques. Cette défiance fragilise la capacité des institutions à produire un cadre de référence commun et cohérent, laissant place à des récits concurrents qui prétendent combler les incertitudes laissées sans réponse. Ce mécanisme a été particulièrement visible lors de la crise sanitaire : dans une enquête conduite dans 21 pays, la méfiance envers les autorités gouvernementales et sanitaires s'est révélée être l'un des déterminants les plus puissants de l'adhésion aux fausses informations liées à la pandémie. Plus largement, **dans un contexte où l'autorité publique est perçue comme illégitime ou déconnectée, toute communication officielle tend à perdre en crédibilité, indépendamment de la solidité de ses arguments.** Les tentatives de correction ou de clarification, loin de rétablir la confiance, sont alors souvent interprétées comme des stratégies de dissimulation, alimentant un cercle de suspicion qui renforce la réceptivité aux narrations alternatives. Cette dynamique explique en partie pourquoi la désinformation trouve une telle résonance dans les contextes de défiance accrue : elle s'impose non seulement par ses contenus, mais parce qu'elle répond à un besoin de compréhension et de réappropriation du sens face à des institutions perçues comme distantes ou impuissantes.

L'environnement économique et matériel

La dégradation des conditions matérielles constitue enfin un déterminant structurel central de la vulnérabilité à la désinformation, en agissant à la fois directement sur les individus et indirectement en nourrissant la défiance

institutionnelle et le sentiment d'injustice. L'aggravation des inégalités, la précarité économique et l'absence de perspectives alimentent un climat de frustration propice à la recherche d'explications globales, simplificatrices et souvent fallacieuses. Ainsi, les croyances conspirationnistes sont plus répandues dans les sociétés marquées par une faible croissance et de fortes inégalités, où les récits désinformants offrent des cadres interprétatifs accessibles, identifient des coupables désignés et apportent une reconnaissance symbolique des expériences de déclassement. Cette dynamique est particulièrement marquée au sein des populations précaires, où le capital éducatif et le temps limité pouvant être consacré à l'information entrave l'exercice du discernement critique. L'effet cumulatif de ces fragilités économiques, éducatives et sociales favorise l'adhésion à des narratifs alternatifs qui font fonction à la fois d'explications causales et de revendications identitaires. À mesure que les inégalités s'approfondissent, l'affaiblissement des mécanismes traditionnels de protection - éducation, stabilité économique, mobilité sociale - réduit encore les capacités de résilience face à ces récits, installant durablement la désinformation dans l'espace public.

3. Lutter efficacement contre la désinformation suppose de dépasser les approches traditionnelles, en construisant une stratégie intégrée fondée sur l'articulation des leviers cognitifs, informationnels et structurels.

Propos liminaire. Les mesures proposées dans ce chapitre ont été choisies parce qu'elles font consensus dans la littérature scientifique et qu'elles ciblent les mécanismes cognitifs de la vulnérabilité informationnelle, expliquant ainsi la mise en avant de certaines approches et la prudence concernant d'autres. Contrairement aux deux premiers volets, ces recommandations dévient du noyau de compétence au sens strict de l'Institut Jean Nicod; ainsi, elles nécessitent une coordination étroite avec des administrateurs et des experts en ingénierie des politiques publiques pour leur mise en œuvre concrète. Bien que l'Institut Jean Nicod rassemble des chercheurs travaillant sur de la conception des politiques publiques, y compris sociales et sanitaires, son cœur d'expertise demeure en sciences cognitives (linguistique, action/perception, cognition sociale). De même, si les enjeux éthiques, juridiques et techniques liés à ces interventions ne sont pas développés ici, ils restent centraux, en particulier face aux tensions entre liberté d'expression et encadrement des contenus.

La lutte contre la désinformation requiert des solutions articulées à trois niveaux :

- la correction au niveau individuel par l'amélioration des capacités cognitives,
- la régulation et l'assainissement de l'écologie informationnelle,
- et le traitement à la racine des sources de la « demande » sociale en récits trompeurs.

Aujourd'hui, les politiques publiques se concentrent essentiellement sur le premier levier et, de manière partielle, sur la seule

dimension numérique du second ; l'ambition de ce rapport est d'offrir un panorama intégratif couvrant l'ensemble des strates d'action possibles.

Volet 1 : Les politiques de renforcement de la résilience cognitive individuelle, bien qu'amplement plébiscitées, restent peu efficaces et risquent d'entraîner des effets rebond.

La correction au niveau individuel, centrée sur l'amélioration des capacités cognitives face à la désinformation, est aujourd'hui l'approche la plus répandue. Son succès tient autant à son attractivité conceptuelle - elle correspond à l'idée intuitive que la désinformation prospère sur la crédulité et l'irrationalité - qu'à des avantages institutionnels : ces mesures sont visibles, peu coûteuses, faciles à mettre en place et font peser la responsabilité sur les individus plutôt que sur les structures. Toutefois, la recherche récente appelle à une grande prudence. Ces dispositifs, qu'ils relèvent du fact-checking ou de la formation à l'esprit critique, s'attaquent avant tout aux symptômes sans traiter les causes profondes. Ils souffrent par ailleurs d'une faible résilience dans le temps : les effets des interventions éducatives s'estompent rapidement et leur capacité à modifier durablement les évaluations de crédibilité reste incertaine.

Comme nous l'avons précédemment expliqué, ces approches peuvent également produire des effets contre-productifs, en renforçant les croyances erronées et en alimentant une défiance accrue envers les sources d'information fiables. Enfin, ils réduisent la désinformation à un problème de crédulité, alors qu'elle relève bien souvent d'un manque "d'intention épistémique" - c'est à dire que le souci de la vérité est souvent loin d'être le premier motivant le partage ou l'adhésion - et d'une absence de mobilisation de capacités cognitives pourtant disponibles. En dépit de leur popularité, ces interventions doivent donc être intégrées avec précaution dans

des stratégies plus larges, ciblant les vulnérabilités sociales, institutionnelles et environnementales sur lesquelles prospèrent réellement les récits désinformants. **Ces approches ne doivent pas être nécessairement écartées, mais intégrées avec discernement, sans les considérer à tort comme des solutions universelles** ; nous en présentons ci-après les principales déclinaisons.

Levier 1 - Développer les compétences informationnelles et l'esprit critique des citoyens par des actions de formations :

Un premier ensemble de mesures, déjà partiellement déployé en France, vise à renforcer les compétences informationnelles et l'esprit critique des citoyens. Il s'agit d'apprendre à évaluer la fiabilité des sources, repérer les biais cognitifs, décrypter les mécanismes de diffusion de l'information et identifier les stratégies de manipulation. Ces objectifs sont poursuivis dans une pluralité de contextes : à l'école via l'Éducation aux Médias et à l'Information (ÉMI), dans des formations à l'analyse critique, et au sein d'actions de médiation locale portées par les bibliothèques, associations ou collectivités. Concrètement, cela prend la forme d'ateliers de fact-checking, de confrontation de récits contradictoires ou de cafés-vérif' organisés dans les médiathèques et lieux de vie.

Ces dispositifs favorisent une réelle montée en compétences, mais présentent deux fragilités bien documentées. D'une part, leurs effets tendent à s'estomper sans réactivation régulière. D'autre part, mal conçus ou mal contextualisés, ils peuvent générer un effet rebond : en éveillant une vigilance généralisée sans distinction de fiabilité, ils risquent d'alimenter la méfiance envers l'ensemble des sources d'information, y compris les plus légitimes. Pour en renforcer l'impact tout en évitant ces dérives, il serait utile d'y adosser des mesures incitatives promouvant une pratique active, régulière et valorisée de l'information - par exemple via l'intégration de l'actualité dans les routines pédagogiques ou la mise en visibilité de contenus fiables sur les plateformes numériques fréquentées par les

jeunes. L'enjeu ne réside pas seulement dans la correction des erreurs, mais dans la construction d'une culture partagée de l'information : seule une exposition durable à des contenus de qualité permet de marginaliser les récits trompeurs.

Levier 2 - Structurer le fact-checking et les dispositifs de débunking pour corriger sans renforcer la désinformation :

La correction publique des fausses informations repose sur des dispositifs dits de « débunking », qui consistent à produire et diffuser des contre-arguments factuels pour contrecarrer des récits trompeurs. Ces corrections peuvent être publiées par les médias, les institutions ou des acteurs indépendants, et relayées sur les réseaux sociaux, moteurs de recherche ou plateformes de messagerie. Lorsqu'ils sont bien conçus - c'est-à-dire rapides, clairs, mémorables, et émanant de sources perçues comme crédibles par les publics visés -, ces dispositifs peuvent parfois limiter la propagation et l'adhésion à certaines fausses informations. Mais ces opérations de correction peuvent produire des effets inverses à ceux recherchés. En répétant la fausse information pour la réfuter, elles risquent de la rendre plus familière et donc plus crédible aux yeux du public (effet d'« illusion de vérité »). Chez les individus déjà méfiants envers les institutions, elles peuvent aussi provoquer un rejet instinctif du message ou être perçues comme une tentative d'imposition autoritaire de la vérité. Dans certains cas, elles nourrissent même des récits conspirationnistes, en apparaissant comme des instruments de censure déguisée ou de manipulation informationnelle.

Levier 3 - Renforcer la résilience cognitive par des campagnes d'inoculation ciblées :

Les campagnes d'inoculation sont des dispositifs de prévention cognitive qui visent à exposer le public, de manière anticipée, aux principales stratégies de manipulation utilisées dans les contenus trompeurs. Inspirées des mécanismes psychologiques

de vaccination, elles consistent à présenter de façon synthétique et engageante des exemples de désinformation typique - accompagnés d'un décryptage explicite de leurs ressorts - afin de développer une vigilance réflexive chez les citoyens. Déployées en amont des périodes sensibles (élections, crises sanitaires, tensions internationales), elles prennent la forme de vidéos courtes, de quiz interactifs ou de contenus pédagogiques diffusés sur les réseaux sociaux, dans les écoles ou les médias. Cependant, leur efficacité à court terme est bien établie, mais les effets tendent à s'estomper rapidement sans réactivation régulière. Par ailleurs, si elles sont mal calibrées, ces campagnes peuvent nourrir une défiance généralisée envers l'ensemble de l'information, y compris celle issue de sources pourtant fiables.

Levier 4 - Freiner la diffusion impulsive et encourager la réflexion au moment du partage par des nudges comportementaux :

Les nudges - ou "coups de pouces" sont des « architectures de choix » issues des sciences comportementales, conçues pour orienter les décisions sans restreindre la liberté d'action. Ces dispositifs partent du constat que le partage de contenus douteux n'est pas toujours dû à une ignorance, mais souvent à une non-activation des capacités critiques déjà acquises ; ils visent donc à déclencher une pause réflexive ciblée pour freiner la diffusion impulsive et réduire la viralité de la désinformation. Sur les plateformes, ils se traduisent par des rappels contextuels (par exemple, « Avez-vous lu cet article avant de le partager ? ») ou des invites de vérification de sources au moment de la publication. Leur efficacité dépend de deux variables clés : l'acceptabilité par l'utilisateur (formulation, fréquence, impression de paternalisme) et la qualité de l'intégration UX (timing du rappel, placement dans le parcours), sous peine de voir leur impact neutralisé par une perception d'intrusion ou de lassitude.

Volet 2 : Les réponses les plus efficaces à la désinformation passent par une transformation structurelle de l'écosystème informationnel.

Deux grands types de leviers permettent de transformer l'écosystème informationnel pour éviter la prolifération de la désinformation : d'un côté, les interventions sur les plateformes et leurs algorithmes, et de l'autre, le soutien structuré aux médias fiables et à la pratique journalistique.

Les mesures concernant les réseaux sociaux et les dispositifs algorithmiques.

La régulation et l'assainissement de l'écologie informationnelle visent à transformer en profondeur les environnements de production et de diffusion des contenus, en agissant directement sur les conditions structurelles qui favorisent la désinformation. **Probablement les plus efficaces et efficaces en termes de rapport coût/bénéfice, ces mesures reposent essentiellement sur des leviers de politique publique et de régulation économique.** Leur principal obstacle n'est pas tant lié à des effets contre-productifs ou à une érosion de leur efficacité dans le temps, mais à la **difficulté de leur mise en œuvre** face à des intérêts politiques et économiques puissants, à des enjeux juridiques complexes de territorialité, et à un lobbying particulièrement actif des acteurs visés. On compte par exemple les mesures suivantes :

Levier 5 - Introduire une friction structurelle pour freiner mécaniquement la viralité des contenus douteux :

À la différence des nudges comportementaux qui sollicitent une pause réflexive de l'utilisateur, la friction structurelle viserait à ralentir mécaniquement la viralité des contenus en intégrant des limitations techniques directement dans l'architecture des plateformes. Inspirée du principe de « Safety by Design », cette approche reposerait sur des dispositifs tels que des

délais obligatoires avant repartage, des plafonds sur la transmission en cascade, ou des restrictions sur le nombre de destinataires simultanés. Ces mesures exploiteraient le ralentissement algorithmique non pour filtrer le contenu, mais pour casser la dynamique de diffusion instantanée qui alimente la désinformation. Déjà éprouvées sur certaines messageries dans des contextes de crise, elles ont démontré leur efficacité pour freiner la contagion informationnelle. Toutefois, leur impact dépendrait d'un calibrage fin : une friction excessive, appliquée de façon indifférenciée, risquerait de ralentir également la circulation de contenus fiables, en particulier en situation d'urgence, et pourrait alors produire un effet contre-productif.

Levier 6 - Restituer la maîtrise du fil par une transparence algorithmique accrue et une possibilité de maîtriser la constitution de son feed :

La transparence algorithmique viserait à rendre compréhensibles et configurables les systèmes de recommandation en publiant leurs critères : **pondérations appliquées, signaux mobilisés** (clics, partages, durée de visionnage), objectifs poursuivis (engagement, rétention). Cette mesure permettrait de lever l'opacité des tris automatisés et d'offrir aux utilisateurs un accès clair à la logique décisionnelle qui structure leur environnement informationnel. Concrètement, un « Tableau de bord de transparence » synthétiserait ces paramètres dans une interface lisible, facilitant l'identification des biais de classement et des logiques de captation attentionnelle. Cette transparence devrait être complétée par un droit à la maîtrise du fil d'actualité, en proposant par défaut un affichage chronologique, sans tri algorithmique initial. Ce mode neutre pourrait ensuite être activement personnalisé via un opt-in (tri thématique ou ciblé) et désactivé à tout moment via un opt-out. Ce mécanisme exploiterait les biais d'inertie cognitive pour favoriser l'adoption du mode non filtré, tout en garantissant une liberté totale de reconfiguration.

En combinant ces deux approches - intelligibilité des critères et liberté de tri - le dispositif pourrait permettre de réduire les risques de manipulation de l'attention, de freiner la polarisation algorithmique et de limiter la diffusion virale de contenus trompeurs, en rééquilibrant les signaux de popularité vers des critères temporels neutres. Une obligation légale en la matière renforcerait aussi la résilience démocratique en rendant les systèmes moins vulnérables à l'orientation politique de leurs propriétaires, notamment en cas de rachat d'une plateforme par un acteur poursuivant un agenda idéologique. Surtout, cette modification est au niveau architectural : elle n'impliquerait aucune modération de contenu ni censure et agirait uniquement sur les modalités de hiérarchisation, respectant ainsi pleinement la liberté d'expression. Pour être pleinement efficace, cette réforme exigerait un volontarisme politique fort pour surmonter les pressions de certains acteurs économiques et faire évoluer le cadre réglementaire européen (DSA, RGPD), une coordination technique entre États pour harmoniser les exigences d'implémentation, et l'ouverture de négociations opérationnelles avec les plateformes afin de garantir l'interopérabilité et l'effectivité des outils de transparence. Enfin, de tels dispositifs n'ayant jamais pleinement été déployés, leur efficacité n'est pas encore une certitude empirique.

Levier 7 - Assainir l'espace numérique et valoriser l'information fiable sans restreindre le pluralisme.

L'amélioration de l'espace informationnel passerait par une double approche : d'une part, l'obligation pour les plateformes de détecter et supprimer rapidement les agents de manipulation - bots, faux comptes coordonnés, deepfakes - à l'aide d'outils automatisés ; d'autre part, la promotion des sources fiables via des labels de qualité fondés sur des critères objectifs tels que l'indépendance éditoriale, la transparence des sources, et l'existence de mécanismes de correction publics. Ces labels, attribués par une entité indépendante, seraient affichés directement sur les navigateurs, référencés dans les résultats de recherche,

et intégrés aux outils de ciblage publicitaire pour orienter les choix des utilisateurs comme des annonceurs. Certaines de ces mesures sont déjà partiellement déployées à l'échelle européenne (e.g. *Journalism Trust Initiative*). Toutefois, cette approche soulève plusieurs défis : la détection automatisée reste sujette aux faux positifs et contournements ; la définition juridique de ce qui constitue un contenu malveillant ou un média « fiable » implique des arbitrages délicats. Une gouvernance rigoureuse, transparente et indépendante serait donc indispensable pour garantir l'efficacité de cette approche tout en préservant la liberté d'expression et le pluralisme médiatique.

Leviers complémentaires : De nombreuses autres politiques complémentaires peuvent renforcer l'assainissement de l'écosystème informationnel.

- **La démonétisation des sites de désinformation** prive les diffuseurs de récits trompeurs de leurs incitations économiques en restreignant l'accès aux régies publicitaires et services de paiement, bien que son efficacité dépende de la solidité des listes noires et du contrôle des réseaux alternatifs.
- **La transparence des publicités politiques, le signalement des comptes ou médias opérés par des acteurs étatiques et la création de registres du lobbying étranger** permettent de rendre visibles les campagnes d'influence, en imposant aux annonceurs de déclarer leurs budgets et sources de financement, mais rencontrent des résistances politiques et des enjeux de respect de la vie privée.
- **L'accès sécurisé aux données pour la recherche et les audits externes** offre un levier essentiel pour évaluer objectivement l'impact des mesures de régulation et éclairer les décisions publiques sur des bases empiriques, sous réserve de garantir la confidentialité et la sécurité des données.
- **Enfin, l'archivage sécurisé des contenus supprimés** assure la traçabilité des décisions de modération et facilite les enquêtes judiciaires, au prix de défis

techniques en matière de gestion des accès et de protection des données sensibles.

Les mesures concernant les médias traditionnels et les journalistes

Une part significative de la prolifération de la désinformation trouve son origine dans la perte de confiance, parfois à juste titre, envers la qualité et l'indépendance des médias traditionnels. Or, la mesure la plus efficace pour lutter contre la désinformation est de lutter pour la bonne information : celle qui est accessible, rigoureuse, pluraliste et digne de confiance. Plutôt que de concentrer les efforts uniquement sur la correction des fausses nouvelles, il est crucial de renforcer l'offre informationnelle de qualité, en soutenant activement les médias indépendants, en favorisant la transparence des pratiques journalistiques et en réinvestissant dans les ancrages locaux de l'information. C'est dans cette logique que s'inscrivent les mesures suivantes.

Levier 8 - Assurer un financement stable pour garantir l'indépendance des médias et du service public d'information.

Le renforcement durable de l'intégrité de l'information passerait par un financement pérenne des médias indépendants, en particulier ceux d'investigation et de service public. Ce soutien pourrait reposer sur un mix d'outils : subventions publiques ciblées, crédits d'impôt dédiés et reversements obligatoires des grandes plateformes numériques, proportionnels à leurs revenus publicitaires. Pour préserver la légitimité de ces financements, leur attribution devrait être strictement conditionnée à des critères d'indépendance éditoriale, de transparence financière et de gouvernance, contrôlés par des audits annuels indépendants. Des dispositifs comparables ont été mis en œuvre dans les pays scandinaves et au Canada, permettant de renforcer la production d'enquêtes de fond, de limiter la dépendance à la publicité privée, et d'offrir aux citoyens une information d'intérêt

général accessible, pluraliste et indépendante des logiques commerciales.

Levier 9 - Revitaliser le journalisme local pour ancrer l'information de qualité dans les territoires.

Les titres de presse locale, encore consultés par plus de 40 millions de Français et bénéficiant d'un niveau de confiance élevé (67 %), sont en déclin. Ils constituent pourtant un levier stratégique pour renforcer l'intégrité de l'information à l'échelle territoriale. Son soutien passerait par des dispositifs ciblés mobilisant les fonds d'innovation territoriale, des allègements fiscaux sélectifs, et le développement de modèles coopératifs ou associatifs adaptés aux spécificités locales. Ces aides devraient également favoriser la transition numérique des formats et encourager les partenariats durables avec les collectivités, notamment pour couvrir la vie démocratique locale et fournir des repères communs. Expérimentée dans certains pays nordiques ou au Canada, cette approche permettrait de revitaliser les écosystèmes médiatiques locaux, de renforcer la cohésion sociale et de restaurer l'engagement civique dans les zones moins couvertes par les grands médias nationaux.

Leviers complémentaires : De nombreuses autres politiques structurelles peuvent contribuer à restaurer la crédibilité des médias et renforcer leur rôle dans l'assainissement de l'écosystème informationnel.

- **Lutter contre la concentration pour restaurer la confiance dans l'indépendance des médias.** La défiance envers les médias traditionnels alimente la crédulité face aux contenus trompeurs ; garantir leur indépendance structurelle est donc un levier central de lutte contre la désinformation. Cela supposerait un renforcement de la régulation du secteur de la presse et de l'édition, en abaissant les seuils de concentration autorisés et en soumettant toute opération de fusion significative à l'avis conjoint de l'ARCOM et de l'Autorité de la concurrence. Pour rendre ce contrôle effectif, l'ARCOM

pourrait par exemple être dotée de pouvoirs d'enquête accrus.

- **Créer un statut de « société de presse à mission » :** intégrer des engagements statutaires en matière d'indépendance éditoriale, de diversité des contenus, de déontologie et de médiation avec le public, contrôlés par une instance indépendante.
- **Mettre en place un ombudsman médiatique externe :** offrir un recours accessible pour le public en cas de dérive éditoriale, avec la capacité de formuler des recommandations contraignantes hors du cadre judiciaire.
- **Renforcer la gouvernance interne des rédactions :** imposer la création de comités d'éthique, la publication des conflits d'intérêts et l'organisation de formations régulières à l'intégrité journalistique.
- **Adopter des lois anti-SLAPP ciblées :** garantir un examen prioritaire des procédures-bâillons et des sanctions contre les recours abusifs visant à intimider les journalistes.

Ces dispositifs, complémentaires, visent à restaurer la confiance du public dans les médias tout en réaffirmant leur rôle de garant d'un débat public informé et pluraliste.

Volet 3 : La réduction des vulnérabilités sociales, économiques et institutionnelles constitue l'approche la plus structurante pour lutter contre la désinformation

La lutte contre la désinformation ne peut être pleinement efficace sans agir sur ses racines sociales, économiques et institutionnelles. Malgré les évolutions technologiques, la désinformation continue de s'appuyer sur des facteurs structurels qui ont historiquement alimenté sa diffusion. Inégalités, précarité, isolement social et défiance institutionnelle en constituent les principaux moteurs, rendant certaines populations "perméables" aux récits trompeurs. La réponse la plus durable consiste moins à traquer sans fin les récits trompeurs qu'à assécher leur terrain en renforçant la cohésion sociale, en réduisant

la misère économique et en restaurant la confiance civique. Ces enjeux ne sont pas périphériques mais centraux : la meilleure manière de combattre la désinformation est de promouvoir activement la bonne information dans une société plus juste, éduquée et solidaire. Alors que l'on privilégie souvent les dépenses régaliennes face aux crises, il faut rappeler que l'État social constitue lui aussi un levier essentiel de résilience cognitive et démocratique. Ces leviers relèvent d'ailleurs pour beaucoup de politiques publiques déjà engagées pour d'autres objectifs - lutte contre la pauvreté, promotion de l'éducation, de la santé mentale ou de l'inclusion sociale - qu'il convient ici de considérer également comme des instruments stratégiques de lutte contre la désinformation.

Levier 10 - Renforcer les politiques sociales, économiques et de santé publique pour réduire la vulnérabilité à la désinformation.

La résilience informationnelle dépend largement des conditions sociales, économiques et psychologiques dans lesquelles évoluent les individus. Pour la consolider, il serait nécessaire de renforcer les politiques publiques en matière de protection sociale, d'accès à l'emploi, à la santé mentale et à l'éducation, notamment dans les territoires marqués par le déclin industriel, la précarité ou un sentiment d'abandon. Des dispositifs concrets - fonds de transition juste, formations professionnelles adaptées, accompagnement médico-social renforcé - pourraient être déployés pour restaurer une stabilité matérielle et cognitive propice à l'analyse critique de l'information. En réduisant les inégalités structurelles et les fractures territoriales, ces mesures limiteraient l'adhésion aux récits conspirationnistes et renforceraient la confiance dans les institutions, condition essentielle à la lutte contre la désinformation.

Levier 11 - Réinvestir les espaces de sociabilité pour renforcer la cohésion sociale et réduire la polarisation.

La lutte contre la désinformation passerait par la reconstruction de liens sociaux concrets et la réduction des clivages intergroupes, souvent exacerbés par l'environnement numérique. Cela supposerait un investissement massif dans la rénovation et l'animation d'espaces de rencontre accessibles - maisons de quartier, médiathèques, tiers-lieux - qui offriraient des cadres physiques de dialogue, de diffusion d'information locale et de participation citoyenne. En parallèle, des programmes de rencontres intergroupes structurés (sociaux, culturels, politiques) pourraient être mis en place pour favoriser la confrontation directe à l'altérité, corriger les représentations erronées entre groupes sociaux, raciaux ou idéologiques, et désamorcer les mécanismes de polarisation. En restaurant des espaces partagés de discussion et de reconnaissance mutuelle, ces dispositifs contribueraient à renforcer la résilience collective face aux récits clivants, à réduire les effets de chambre d'écho, et à retisser une base commune de confiance sociale indispensable à une information partagée.

Levier 12 - Renforcer l'intégrité publique perçue pour restaurer la crédibilité des institutions.

La vulnérabilité à la désinformation est fortement liée à la perception que les citoyens ont de leurs institutions, notamment en matière d'intégrité, de probité, de transparence et de proximité. L'impression - qu'elle soit fondée ou non - d'opacité, d'impunité et de distance sociale des responsables publics affaiblit la légitimité de la parole institutionnelle et alimente l'adhésion aux récits manipulateurs - notamment de décrédibilisation des acteurs publics. Pour restaurer la confiance, il serait essentiel de renforcer la lisibilité et la traçabilité des engagements déontologiques : déclarations d'intérêts systématiques, encadrement strict des allers-retours entre secteur public et privé, sanctions crédibles et effectivement appliquées, indépendamment des effets

d'annonce. Ces dispositifs devraient être complétés par des formats de redevabilité accessibles et localement ancrés. Ces mesures, bien qu'insuffisantes à elles seules, devraient s'inscrire dans un effort plus large de renforcement effectif de la probité publique.

Conclusion :

Agir contre la désinformation sans s'attaquer à ses causes profondes revient à vider un tonneau des Danaïdes : à force de combattre inlassablement les symptômes en surface, on risque d'alimenter les conditions mêmes de leur réapparition. Les écueils identifiés en introduction - une approche réactive et fragmentée privilégiant l'urgence au détriment du long terme, une focalisation excessive sur l'amont de la chaîne informationnelle sans considération pour ses effets réels à l'aval, et l'absence d'un chaînon cognitif permettant d'articuler sciences du cyberspace, géopolitique et sciences empiriques de la cognition - persistent et fragilisent l'efficacité des dispositifs publics. Pire, ces angles morts exposent les politiques publiques à des effets rebond, où les instruments censés corriger la désinformation contribuent, par méconnaissance des mécanismes cognitifs et sociaux, à la renforcer. Sortir de ce cycle suppose de replacer la lutte contre la désinformation dans une écologie cognitive intégrée, articulant trois niveaux d'action interdépendants :

- **Au niveau individuel**, dépasser les modèles naïfs de la psychologie fondés sur une vision linéaire et simpliste de l'exposition et de la crédulité. Il convient de renforcer la résilience cognitive par une éducation ancrée dans les résultats des sciences comportementales, calibrée pour éviter la méfiance généralisée, et accompagnée d'outils pédagogiques permettant d'exercer un véritable discernement dans des environnements informationnels saturés.
- **Au niveau informationnel**, assainir en profondeur l'écosystème médiatique en régulant les plateformes et les algorithmes, en instaurant des dispositifs de transparence et de redevabilité algorithmique, et en soutenant les médias de qualité. Cela implique également de revitaliser les espaces de débat pluralistes, d'encadrer les logiques de monétisation favorisant la viralité des contenus toxiques, et de renforcer l'offre de contenus dignes

de confiance pour rétablir une norme sociale de fiabilité informationnelle.

- **Au niveau social et institutionnel**, traiter les racines de la « demande » sociale de récits trompeurs en agissant sur les vulnérabilités économiques, sociales et psychologiques. La lutte contre la précarité, la réduction de l'isolement relationnel, la désescalade de la polarisation et la restauration de la confiance civique doivent être envisagées comme des leviers stratégiques de résilience démocratique, au même titre que les politiques de sécurité ou de défense. Loin de s'opposer, l'État social et l'État régalien sont ici complémentaires : l'un protège les frontières physiques et institutionnelles, l'autre renforce l'immunité cognitive de la société.

Ces trois dimensions ne sauraient être dissociées : sans un socle de justice sociale, la défiance continuera de nourrir la perméabilité aux récits trompeurs ; sans un écosystème informationnel régulé et pluraliste, l'éducation critique sera submergée par les flux toxiques ; sans une approche cognitive éclairée par les sciences empiriques, la régulation et l'éducation échoueront à cibler les véritables mécanismes d'adhésion et de diffusion. **La désinformation n'est pas une guerre à gagner, mais un équilibre systémique à restaurer.** Relever ce défi impose de dépasser les réflexes gestionnaires pour bâtir une stratégie transdisciplinaire, fondée sur la rigueur empirique et une vision politique de long terme. C'est à ce prix que l'on rééquilibrera durablement un paysage informationnel aujourd'hui fragmenté, saturé et vulnérable.

À propos des auteurs et contributeurs :

Grégoire Darcy est doctorant en sciences cognitives à l'École Normale Supérieure, spécialisé en cognition sociale. Ses recherches portent sur la confiance, les normes et les politiques publiques.

Hugo Mercier est psychologue et directeur de recherches au CNRS. Cognitiviste, il est spécialisé dans la psychologie du raisonnement et l'étude de la communication.

Alda Mari est linguiste et directrice de recherche au CNRS. Elle est spécialiste de la sémantique formelle, avec un intérêt particulier pour la sémantique modale et les attitudes propositionnelles.

Gloria Origgi est philosophe et directrice de recherche au CNRS. Ses travaux portent sur l'épistémologie sociale, la confiance, la réputation et l'impact des technologies numériques sur la connaissance.

Lydia Yahia-Cherif est neuroscientifique et ingénieure de recherche au CNRS. Spécialisée en neuroimagerie fonctionnelle et en statistiques appliquées, elle travaille notamment sur l'activité cérébrale et les technologies de communication

Roberto Casati est philosophe et directeur de recherche au CNRS, ainsi que directeur de l'Institut Jean Nicod. Ses recherches portent sur la perception, la représentation spatiale et les artefacts cognitifs.

Crédits photo de couverture : Alina Grubnyak

Ressources et Bibliographie indicative

Ressources et Rapports :

- Académie des technologies. (2024). *IA générative & mésinformation*. Rapport de l'Académie des technologies.
- Commission « Les Lumières à l'ère numérique ». (2022). *Les Lumières à l'ère numérique : Rapport de la commission pour le Président de la République française*. République française.
- États généraux de l'information. (2024). *Protéger et développer le droit à l'information : Une urgence démocratique*. Rapport des États généraux de l'information.
- OCDE. (2024). *Les faits sans le faux : Lutter contre la désinformation, renforcer l'intégrité de l'information*. Éditions OCDE.
- Parlement européen. (2022). *Rapport sur les ingérences étrangères dans l'ensemble des processus démocratiques de l'Union européenne, y compris la désinformation (A9-0022/2022)*.
- Pasquinelli, E., Farina, M., Bedel, A., & Casati, R. (2020). *Définir et éduquer l'esprit critique*. [Rapport de recherche]. Institut Jean Nicod.

Bibliographie indicative

- Acerbi, A., Altay, S., & Mercier, H. (2022). Research note: Fighting misinformation or fighting for information? *Harvard Kennedy School Misinformation Review*.
- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*.
- Altay, S. (2022). How effective are interventions against misinformation? *PsyArXiv*.
- Altay, S., Acerbi, A., & Berriche, M. (2023). Misinformation on misinformation: Conceptual and methodological challenges. *Social Media & Society*.
- Altay, S., Berriche, M., Heuer, H., Farkas, J., & Rathje, S. (2023). A survey of expert views on misinformation: Definitions, determinants, solutions, and future of the field. *Harvard Kennedy School Misinformation Review*.
- Anderson, C. W. (2021). Propaganda, misinformation, and histories of media techniques. *Harvard Kennedy School Misinformation Review*.
- Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford University Press.
- Bronner, G. (2013). *La démocratie des crédules*. Paris : Presses Universitaires de France.
- Camargo, C. Q., & Simon, F. M. (2022). Mis- and disinformation studies are too big to fail: Six suggestions for the field's future. *Harvard Kennedy School Misinformation Review*.
- Chan, M. P. S., Jones, C. R., Jamieson, K. H., & Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological Science*, 28(11), 1531–1546.
- Chater, N., & Loewenstein, G. (2023). The i-frame and the s-frame: How focusing on individual-level solutions has led behavioral public policy astray. *Behavioral and Brain Sciences*.
- Cinelli, M., Quattrocchi, W., Galeazzi, A., ... & Scala, A. (2020). The COVID-19 social media infodemic. *Scientific Reports*.
- Citton, Y. (2014). *Pour une écologie de l'attention*. Le Seuil.
- Cordonier, L., & Brest, A. (2021). How do the French inform themselves on the Internet? Analysis of online information and disinformation behaviors. *Fondation Descartes*. <https://hal.archives-ouvertes.fr/hal-03167734/document>
- Del Vicario, M., Bessi, A., Zollo, F., ... & Quattrocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*.
- Dieguez, S., & Delouée, S. (2021). *Le complotisme : cognition, culture, société*. Bruxelles : Éditions Mardaga.

- Ecker, U. K. H., Lewandowsky, S., Cook, J., ... & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*.
- Flynn, D. J., Nyhan, B., & Reifler, J. (2017). The nature and origins of misperceptions: Understanding false and unsupported beliefs about politics. *Political Psychology*.
- Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*.
- Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*.
- Guess, A. M., Nyhan, B., & Reifler, J. (2020). Exposure to untrustworthy websites in the 2016 US election. *Nature Human Behaviour*.
- Johnson, N. F., Velásquez, N., Restrepo, N. J., ... & Lupu, Y. (2020). The online competition between pro- and anti-vaccination views. *Nature*.
- Lazer, D. M. J., Baum, M. A., Benkler, Y., ... & Zittrain, J. L. (2018). The science of fake news. *Science*.
- Lewandowsky, S., Ecker, U. K. H., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the “post-truth” era. *Journal of Applied Research in Memory and Cognition*.
- Loomba, S., de Figueiredo, A., Piatek, S. J., de Graaf, K., & Larson, H. J. (2021). Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nature Human Behaviour*.
- Mercier, H. (2022). *Pas né de la dernière pluie*. Paris : humenSciences.
- Mercier, H., & Sperber, D. (2021). *L'énigme de la raison*. Paris : Odile Jacob.
- Morales, P. R., Berriche, M., & Cointet, J.-P. (2023). The geometry of misinformation: Embedding Twitter networks of users who spread fake news in geometrical opinion spaces. *Proceedings of the International AAAI Conference on Web and Social Media*.
- O'Connor, C., & Weatherall, J. O. (2019). *The misinformation age: How false beliefs spread*. Yale University Press.
- Origgi, G. (2017). *Reputation: What It Is and Why It Matters*. Princeton University Press.
- Origgi, G., & Sperber, D. (2000). Evolution, communication, and the proper function of language. In P. Carruthers & A. Chamberlain (Eds.), *Evolution and the Human Mind: Language, Modularity and Social Cognition* (pp. 140–169). Cambridge University Press.
- Panizza, F., Camminga, T., Gaillard, S., & Grüning, D. J. (2025). Adopt a cultural-historical perspective to adapt misinformation interventions: Reflecting on Harjani et al. *Journal of Trial and Error*.
- Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*.
- Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*.
- Roozenbeek, J., & van der Linden, S. (2024). *The Psychology of Misinformation* (1^{er} éd.). Cambridge University Press.
- Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*.
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language*.
- Sultan, M., Tump, A. N., Ehmann, N., Lorenz-Spreen, P., Hertwig, R., Gollwitzer, A., & Kurvers, R. H. J. M. (2024). Susceptibility to online misinformation: A systematic meta-analysis of demographic and psychological factors. *Proceedings of the National Academy of Sciences*.
- Tandoc, E. C., Jr., Lim, Z. W., & Ling, R. (2017). Defining “fake news”: A typology of scholarly definitions. *Digital Journalism*.
- Uscinski, J. E. (Ed.). (2018). *Conspiracy theories and the people who believe them*. Oxford University Press.
- van der Linden, S. (2023). *Foolproof: Why misinformation infects our minds and how to build immunity*. W. W. Norton.
- Voigtländer, N., & Voth, H.-J. (2015). Nazi indoctrination and anti-Semitic beliefs in Germany. *Proceedings of the National Academy of Sciences*.