

# Fake news et deepfakes : une approche cyberpsychologique

## *Fake news and deepfakes: a cyberpsychological approach*

< Julien NELSON <sup>1</sup> >

1. Centre Giovanni Borelli – Université de Paris  
julien.nelson@u-paris.fr

DOI : 10.25965/interfaces-numeriques.4830

### < RÉSUMÉ >

Une *fake news* est définie comme un élément médiatique factuellement inexact mais présenté par son émetteur comme ayant une valeur de véracité avec une volonté explicite de tromper le récepteur de ce dernier. Le développement massif des réseaux sociaux couplé à un certain nombre d'évènements démocratiques de grande ampleur ces dernières années ont sans doute contribué à une plus grande sensibilisation du grand public aux dangers posés par les fake news. Cet article propose un bref état de l'art des recherches en cyberpsychologie – notamment des travaux relevant d'une approche expérimentale – sur les facteurs influant sur la crédibilité des fake news. Il aborde aussi la question des *deepfakes*, que l'on peut considérer comme des vecteurs iconographiques de ces fake news.

### < MOTS-CLÉS >

fake news, deepfakes, cyberpsychologie, effet de vérité illusoire

### < ABSTRACT >

Fake news are defined as fabricated information that is factually inaccurate but presented by its creator as having the value of truth, with an explicit goal of deceiving its receiver. The mass development of social networks, along with a number of major democratic events in recent years, have perhaps contributed to a greater sensitivity of the public to the dangers of "fake news". The present paper proposes a short review of existing research in cyberpsychology – particularly works following an experimental-psychological approach focusing on the factors influencing and processes involving belief in fake news. It also addresses the emerging topic of deepfakes, which can be thought of as iconographic vectors for fake news.

< **KEYWORDS** >

fake news, deepfakes, cyberpsychology, illusory truth effect

---

## 1. Introduction

Le présent article est une réponse à un appel fait par la recherche en sémiotique pour caractériser un ensemble de phénomènes (textuels et/ou visuels) ayant fait l'objet d'une attention médiatique particulière ces dernières années, regroupés sous le terme global de *fake news*. Plus particulièrement, il s'agit ici de faire le point sur les apports d'une perspective théorique et empirique spécifique, à la compréhension de ces phénomènes : celle de la *cyberpsychologie*, et notamment de la psychologie expérimentale sur laquelle elle s'appuie. Nous aborderons en premier lieu la question des *fake news* et des connaissances issues de la *cyberpsychologie* relatives à leur crédibilité. Dans un second temps, nous aborderons une évolution technologique récente : celle des *deepfakes*, qui posent des questions nouvelles sur les facteurs influençant la crédibilité des *fake news*, mais aussi sur les stratégies qui permettent de s'en prémunir.

Une *fake news* est ainsi définie comme un élément médiatique factuellement inexact mais présenté par son émetteur comme ayant une valeur de véracité avec une volonté explicite de tromper le récepteur de ce dernier (Pennycook & Rand, 2021). Si le phénomène est ancien – Kalsnes (2018) donne plusieurs exemples datant de la fin du XIX<sup>ème</sup> siècle visant par exemple à augmenter les chiffres de vente d'un journal, à divertir un public, ou encore à attiser des sentiments d'hostilité à l'égard d'une minorité – ce terme est réapparu au cours des cinq dernières années, à la faveur de deux phénomènes distincts : (1) la survenue de plusieurs événements démocratiques historiques, par exemple le référendum sur l'appartenance du Royaume-Uni à l'Union Européenne, qui a donné lieu au Brexit, ou encore l'élection de Donald Trump à la présidence des Etats-Unis d'Amérique, qui en a incontestablement popularisé l'usage ; et (2) l'irruption des réseaux sociaux dans ces événements, comme vecteurs de ces messages erronés mais susceptibles d'infléchir sur les résultats des votes. L'intérêt de la psychologie pour ce thème émergent est double. En premier lieu, il permet d'explorer des

thèmes classiques en psychologie sociale tels que la formation de croyances et ses liens avec la prise de décision, sous un angle nouveau : l'étude de la crédulité et des facteurs individuels et contextuels qui la favorisent (Forgas & Baumeister, 2019). En second lieu, naturellement, l'objectif de telles études est souvent de proposer des pistes pour prévenir de telles tromperies ou d'en atténuer les conséquences.

## **2. Exposition aux fake news et prise de décision**

### ***2.1. L'approche classique de la prise de décision en psychologie cognitive***

L'intérêt récent pour les *fake news* tient à leurs effets supposés sur la prise de décision humaine. En psychologie cognitive, un programme de recherche initié dans les années 1960 et 1970 a contribué à montrer en quoi l'être humain ne pouvait être considéré comme un agent décisionnel parfaitement rationnel (Kahneman et al., 1982). En effet, selon le modèle classique défendu notamment en sciences économiques, lorsque l'être humain est confronté à une situation où (1) il peut choisir entre plusieurs actions possibles et (2) de ce choix découlera une issue plus ou moins avantageuse, alors sa décision reposera sur un calcul intégrant, d'une part, le caractère plus ou moins avantageux de chaque issue et d'autre part, la probabilité que telle décision donnera effectivement lieu à telle issue. Dans une situation simpliste dite de jugement sous certitude, où telle action donnera lieu, dans tous les cas, à une issue donnée, la prise de décision est fondée uniquement sur l'évaluation des bénéfices comparés de chaque issue. En pratique, cependant, le lien entre un choix et une issue n'est pas certain. La qualité de la prise de décision dépendra de celle de l'information sur laquelle elle s'appuie. Cependant, l'être humain dispose de ressources limitées pour la prise de décision, que celles-ci soient internes – c'est à dire cognitives, par exemple attentionnelles ou mnésiques – ou externes à l'individu, par exemple lorsqu'une décision est prise sous contrainte temporelle. Dans ces situations de prise de décision sous incertitude, l'agent décisionnel ne traite pas toutes les informations dont il dispose, mais juste suffisamment d'informations pour parvenir à une issue satisfaisante (Simon, 1957). Ce faisant, il s'appuie sur des heuristiques, c'est-à-dire des principes visant à simplifier les traitements

d'information impliqués dans le calcul des probabilités de chaque issue et les utilités associées (Tversky & Kahneman, 1974). La notion de *biais* renvoie au fait que ces jugements subjectifs sont parfois erronés, et ce de manière prévisible.

## 2.2. L'effet de vérité illusoire

Une première piste pour comprendre les effets des *fake news* sur la prise de décision a été mise en évidence par Hasher et al. (1977). Ces auteurs ont exposé leurs participants à quarante affirmations plausibles et leur ont demandé d'en évaluer la vraisemblance. Ils ont répété l'expérience trois fois à deux semaines d'intervalle. D'une session à l'autre, la moitié des affirmations présentées était nouvelle, tandis que les autres étaient répétées d'une fois sur l'autre. Les résultats ont montré que la simple exposition répétée à une affirmation augmentait sa vraisemblance perçue. Cet effet dit de *vérité illusoire* pourrait expliquer la croyance dans la véracité des *fake news*, dès lors que celles-ci sont diffusées massivement *via* les réseaux sociaux et les canaux médiatiques traditionnels.

Deux explications ont été avancées pour expliquer l'effet de vérité illusoire. La première, dite *référentielle*, repose sur le fait que l'individu est capable de se souvenir qu'il a déjà été exposé à une affirmation. De plus, si l'affirmation était accompagnée d'une information sur sa valeur de vérité – vraie ou fausse – l'individu peut se *référer* à cette information (Brown & Nix, 1996). La seconde, dite *non-référentielle*, intervient même lorsque l'individu ne dispose pas d'une information de véracité, et repose sur le fait que la simple exposition répétée à une même affirmation facilite son traitement cognitif ultérieur. Cette facilité de traitement est perçue à son tour comme un indice de la véracité de l'affirmation (Unkelbach & Stahl, 2009). Autrement dit, la simple exposition répétée à une *fake news* pourrait suffire à rendre celle-ci crédible aux yeux du lecteur. Ce phénomène est d'autant plus préoccupant que l'effet de vérité illusoire persiste même lorsque l'individu dispose déjà de connaissances qui permettent de contredire l'information erronée (Fazio et al., 2015).

L'effet de vérité illusoire apparaît donc comme un processus explicatif clé de l'efficacité des *fake news*. Heureusement, la simple exposition

répétée à une information ne suffit pas pour augmenter sa véracité perçue. Pennycook et al. (2018, étude 1) ont montré que cet effet n'existait qu'à la condition que l'information présentée, quoique fausse, restait néanmoins plausible. Une information peu plausible – dans l'article cité, « la Terre est un carré parfait » - ne donnera pas lieu à cet effet.

### **2.3. Vérité illusoire et raisonnement partisan**

L'exemple que nous venons de citer est édifiant. L'affirmation « la Terre est un carré parfait » est manifestement fausse – en l'occurrence, contredite par un consensus scientifique, lui-même construit sur la base d'une expérience commune de plusieurs millénaires. Pourtant, ce type d'affirmation continue de susciter une adhésion au moins de certains groupes sociaux, par exemple autour de la résurgence de la théorie de la « Terre plate ». Ce type de croyance est largement débattu en ligne, avec des arguments avancés en faveur comme en défaveur de cette thèse. Entretenir une croyance minoritaire (ici, que la Terre est plate) n'implique pas seulement d'accorder une crédibilité aux affirmations en faveur de cette thèse, mais aussi d'accorder une plus faible crédibilité à celles qui la contredisent – ou encore à accorder une forte crédibilité aux informations contribuant à discréditer ces sources<sup>1</sup>.

L'intérêt académique récent pour les *fake news* s'est inscrit dans le contexte politique, et plus particulièrement dans le contexte politique américain. Ce système a pour caractéristique de comporter deux partis au idéologies distinctes (Républicain vs. Démocrate). Le débat public repose sur la diffusion d'informations par de multiples canaux médiatiques, y compris des canaux où la diffusion est largement facilitée mais le processus éditorial peut être absent. Cette configuration particulière a été à l'origine de l'hypothèse suivante : l'orientation idéologique pourrait influencer sur la crédibilité qu'un individu apporte à

---

<sup>1</sup> Notons à ce propos que le terme de *fake news* est réapparu dans la conscience populaire lorsque le Président américain Donald Trump l'a utilisé à plusieurs reprises pour *discréditer* des informations véhiculées dans les médias traditionnels (presse, télévision), ayant *a priori* subi un processus éditorial classique.

une information. Cette hypothèse, appelée du *raisonnement motivé*, propose que le traitement de l'information par un individu serait biaisé en fonction de ses convictions idéologiques. Plus particulièrement, l'individu accorderait davantage de crédibilité aux affirmations concordant avec ses convictions, et moins de crédibilité aux affirmations discordantes (Kahan, 2013).

Kahan et al. (2011) ont ainsi évalué la crédibilité perçue de trois individus fictifs, mais présentés comme experts d'un domaine idéologiquement clivant – le changement climatique et son origine anthropique, l'énergie nucléaire, et la législation en faveur du port d'armes, respectivement – présentés comme exprimant une opinion en relation avec leur sujet d'expertise. L'orientation de cette opinion était manipulée expérimentalement. Les auteurs ont mesuré l'expertise perçue de l'individu : ce dernier était perçu comme d'autant plus expert que ses conclusions s'accordaient avec les convictions idéologiques du participant.

Pennycook et al. (2018, étude 2) ont étudié l'impact de ce « raisonnement partisan » sur la perception de la véracité d'informations présentées dans un format semblable à un *post* Facebook (Tableau 1). Les informations étaient de nature politique et idéologiquement orientées, i.e., mettant en valeur ou non un politicien à la tendance politique bien connue. La moitié des informations présentées était fausse, les autres étaient vraies. Enfin, dans le cas des *fake news*, les auteurs ont également manipulé expérimentalement la présence ou l'absence d'un bandeau indiquant que la véracité de l'information était disputée par un tiers.

	Pro-républicain (anti-démocrate)	Pro-démocrate (anti-républicain)
<b>Information fausse (« fake news »)</b>	« Le directeur du FBI, James Comey, vient de démontrer ses orientations en mettant une affiche pro-Trump à l'entrée de sa maison »	« Mike Pence : la thérapie de conversion homosexuelle a sauvé mon mariage »
<b>Information vraie</b>	« Sur Twitter, Trump met en doute (à tort) l'ingérence russe dans les élections »	« un rapport affirme que Vladimir Poutine était personnellement impliqué dans le piratage aux US »

Tableau 1 – Exemples d'informations utilisées dans l'étude de Pennycook et al. (2018, étude 2).

Conformément à la procédure classiquement adoptée dans les travaux sur l'effet de vérité illusoire (Hasher et al., 1977; cf. partie 2.1), la passation expérimentale se déroulait en plusieurs étapes, avec une étape de familiarisation suivie d'une étape d'évaluation, dans laquelle la moitié des informations présentées était familière, et les autres étaient nouvelles. Les résultats ont montré que les informations familières étaient perçues comme plus vraisemblables que les informations nouvelles, suggérant un effet de vérité illusoire. D'autre part, les informations (effectivement) vraies étaient perçues comme plus vraisemblables que les informations fausses. En revanche, aucun effet d'interaction n'a été détecté : l'exposition préalable augmentait la crédibilité perçue des informations, que celles-ci soient vraies ou fausses. Les auteurs ont aussi contrôlé l'orientation idéologique des participants (Démocrate ou Républicaine) à l'aide d'un questionnaire. Pour examiner la présence d'un biais partisan, ils ont classé les informations présentées suivant leur compatibilité avec l'orientation du participant (concordante ou discordante). Les résultats ont montré que les informations présentées étaient perçues comme plus vraisemblables si elles étaient concordantes avec l'orientation du participant que dans le cas contraire. En revanche, il n'y a pas eu d'effet d'interaction : l'exposition répétée à un message faux conduisait les participants à y accorder plus de crédibilité, que ce message soit en accord ou non avec leurs convictions. Ces résultats montrent que le raisonnement partisan pourrait conduire un individu à

accorder de la crédibilité à une *fake news* dès lors que celle-ci est conforme à ses convictions.

#### **2.4. Orientation idéologique ou croyances existantes ?**

L'étude que nous venons de présenter montre que les *fake news* semblent être d'autant plus facilement acceptées qu'elles sont compatibles avec les orientations politiques de l'individu. Pennycook et Rand (2021) ont cependant montré une limite de cette étude, lié au phénomène dit de la « troisième variable » (Shadish et al., 2002). L'orientation politique présente des corrélations élevées avec de nombreuses variables, et il est possible que l'effet apparent de l'orientation politique soit en fait dû à d'autres variables qui lui sont corrélées.

Parmi ces variables, on peut retenir les croyances existantes du participant. Cette explication de l'effet de vérité illusoire a le premier avantage de permettre un lien théorique avec un biais cognitif bien connu : le biais de confirmation (Nickerson, 1998). Ce biais conduit l'individu à accorder plus de crédit à des indices confirmant une hypothèse existante, et moins de crédit à des indices infirmant cette hypothèse.

Un second avantage de cette interprétation est qu'elle permet d'élargir le débat sur la vulnérabilité aux *fake news* au-delà du champ politique. Par exemple, les réseaux sociaux sont le lieu de nombreux débats sur l'efficacité de vaccins et le lien entre la vaccination et la survenue d'autres troubles. L'exemple le plus connu est la croyance en un prétendu lien entre la vaccination à la rougeole, aux oreillons et à la rubéole (dit vaccin ROR) et l'autisme, lien pourtant démenti par la littérature scientifique (DeStefano & Shimabukuro, 2019). Nyhan et al. (2014) ont évalué l'efficacité de différentes stratégies visant à dissiper les croyances erronées sur un lien entre le vaccin ROR et l'autisme. Trois d'entre elles reposaient sur l'exposition à des informations émanant d'une source officielle, le CDC (*Centers for Disease Control and Prevention*) : un démenti sur le lien entre vaccination et autisme, une description des symptômes et risques associés aux maladies que le vaccin permet de prévenir, et un récit d'une mère relatant l'hospitalisation de son enfant atteint de la

rougeole. La dernière consistait à exposer les participants à des photos d'enfants atteints des maladies. L'expérimentation prévoyait enfin une condition contrôle dans laquelle les participants n'étaient exposés à aucune intervention.

L'étude a montré un résultat surprenant : quand bien même les interventions pouvaient réussir à dissiper les croyances erronées sur les effets des vaccins, ceci ne se traduisait pas nécessairement par un effet favorable sur les intentions à vacciner son enfant. Chez les parents nourrissant les attitudes les plus négatives à l'égard du vaccin, l'exposition pouvait même donner lieu à une hostilité plus prononcée au vaccin.

### ***2.5. Les deepfakes : un nouveau medium de désinformation***

Le terme de *deepfake* désigne un type de contenu d'émergence récente exploitant les technologies du *deep learning*, issues du domaine de l'Intelligence Artificielle, pour créer des contenus très réalistes avec l'intention de manipuler le spectateur (Verdoliva, 2020). Si les finalités peuvent être variées – divertir un auditoire, encourager une réflexion citoyenne sur les dérives de la technologie – l'objectif est parfois de tromper l'auditoire, par exemple en fournissant un support réaliste dans lequel un individu connu semble émettre des propos qu'il n'a jamais émis.

La vidéo virale *You won't believe what Obama says in this video*<sup>2</sup> combine deux de ces finalités. Une première partie, falsifiée, met en scène l'ex-président américain Barack Obama qui semble proférer des insultes à l'encontre de son successeur Donald Trump. La seconde partie constitue un démenti, où il est révélé qu'il s'agissait d'un montage, les propos étant prononcés par un imitateur. Cette vidéo a fait date, non seulement parce qu'elle a permis de sensibiliser le grand public aux problématiques informationnelles des *deepfakes*, mais aussi parce qu'elle illustre des progrès techniques frappants pour l'époque (Suwajanakorn et al., 2017).

La littérature actuelle concernant les *deepfakes* se concentre principalement sur la mise au point de solutions techniques visant à

---

<sup>2</sup> <https://www.youtube.com/watch?v=cQ54GDm1eL0>

identifier ces derniers. En comparaison, peu de travaux se sont intéressés aux aspects psychologiques de l'exposition à ces contenus. Vaccari et Chadwick (2020) ont examiné la crédibilité perçue de la vidéo que nous venons d'évoquer. Ils ont exposé des participants à trois contenus possibles : (a) un extrait trompeur court dans lequel « Obama » insultait explicitement Donald Trump, (b) un extrait trompeur long comprenant l'intégralité de l'intervention d'« Obama », mais pas la seconde partie révélant la tromperie ; et (c) la vidéo intégrale comportant l'allocution d'Obama et la révélation qu'il s'agit d'un extrait manipulé (Figure 1). A l'issue du visionnage, il était demandé aux participants si à leur avis Obama avait bien prononcé l'insulte contre Trump (oui / non / ne sait pas).

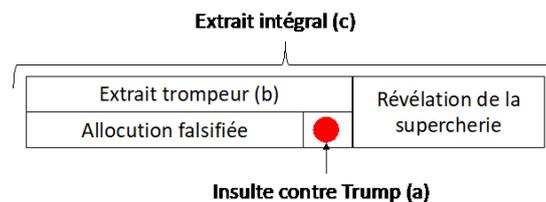


Figure 1 – Conditions expérimentales dans l'étude de Vaccari et Chadwick (2020)

Les résultats montrent que 49,2% des participants avaient été dupés par la vidéo. Ce chiffre intègre les participants ayant déclaré que l'insulte avait vraiment été prononcée (16,0%) et ceux ayant répondu qu'ils ne savaient pas si l'extrait était authentique ou non (33,2%). Si la proportion des participants ayant cru à l'authenticité de l'insulte ne différait pas significativement suivant la condition, le contenu de la vidéo a eu un effet significatif sur la proportion des participants ayant déclaré l'authenticité incertaine. Cette proportion était supérieure dans la condition « extrait long » (36,9%) et dans la condition « extrait court » (35,1%) par rapport à la condition « extrait intégral » (27,5%). Surtout, les auteurs ont montré que l'exposition à un média trompeur pouvait éroder la confiance dans les informations diffusées sur les médias sociaux, et que cet effet était causalement expliqué par l'incertitude ressentie quant à la véracité d'un contenu trompeur.

Moliner (2016) a proposé un programme de recherche sur le thème de la psychologie sociale de l'image, dans lequel il est supposé que la

transmission de l'information via l'image implique que l'émetteur de l'image et son destinataire partagent un ensemble de croyances collectives. La psychologie sociale invite à donner, dans cet ensemble, une place particulière aux catégorisations sociales et aux stéréotypes (Figure 2). Assurément, les *deepfakes* représentent une situation dans laquelle ce contrat n'est pas respecté, dans la mesure où l'émetteur est animé de l'intention de tromper le destinataire, c'est-à-dire de présenter comme véridique une information fautive. Nous avons vu que cette tromperie peut s'appuyer sur des croyances partagées : l'étude de ces dernières et de leur influence sur la croyance dans les *deepfakes* pourrait offrir de nouvelles perspectives de prévention de leurs effets néfastes.

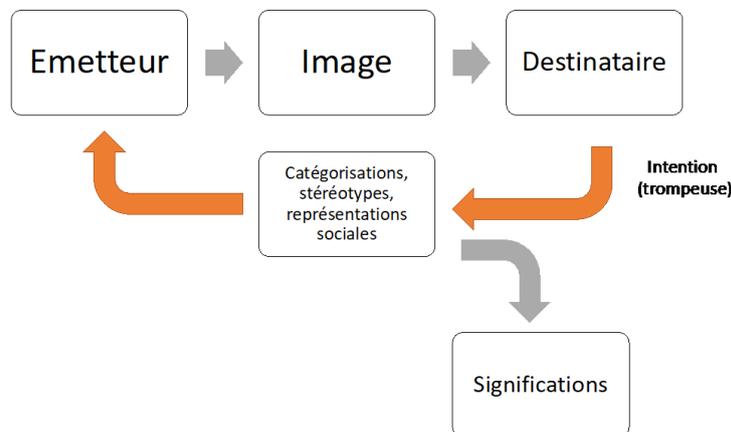


Figure 2 – Schéma d'un programme de recherche sur la psychologie sociale des *deepfakes*, adapté de Moliner (2016)

### 3. Conclusions

Le présent article s'est donné pour objectif de faire un état de quelques travaux actuels en psychologie sur le thème des *fake news*. Plus particulièrement, cette synthèse porte sur les travaux relevant d'une approche expérimentale. Ce champ de recherches constitue un prolongement naturel de travaux classiques en psychologie cognitive et sociale sur le raisonnement. Il montre que les *fake news* tirent parti de limites endogènes au raisonnement humain : l'exposition constante à des contenus trompeurs, notamment via les réseaux sociaux, tend à faciliter le traitement cognitif d'informations erronées. De plus, face à une *fake*

*news*, nous ne sommes pas un terrain neutre : nos convictions antérieures peuvent influencer sur notre vulnérabilité à ces informations erronées, tant dans la mesure où elles peuvent influencer sur nos croyances et nos actions, que dans la mesure où nous pouvons choisir de les propager.

Ces travaux de recherche permettent aussi d'évaluer l'efficacité d'interventions visant à prémunir le grand public des effets indésirables des fake news. Les évolutions de la technologie ouvrent des possibilités nouvelles pour la manipulation de contenus textuels et imagés de plus en plus vraisemblables, mais aussi pour de nouvelles stratégies de prévention.

### **Bibliographie**

- Brown, A. S., & Nix, L. A. (1996). Turning lies into truths: Referential validation of falsehoods. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(5), 1088–1100. <https://doi.org/10.1037/0278-7393.22.5.1088>
- DeStefano, F., & Shimabukuro, T. T. (2019). The MMR Vaccine and Autism. *Annual Review of Virology*, 6(1), 585–600. <https://doi.org/10.1146/annurev-virology-092818-015515>
- Fazio, L. K., Brashier, N. M., Payne, B. K., & Marsh, E. J. (2015). Knowledge does not protect against illusory truth. *Journal of Experimental Psychology: General*, 144(5), 993–1002. <https://doi.org/10.1037/xge0000098>
- Forgas, J. P., & Baumeister, R. F. (2019). Homo credulus: on the social psychology of credibility. In J. P. Forgas & R. F. Baumeister (Eds.), *The social psychology of gullibility: fake news, conspiracy theories and irrational beliefs* (pp. 1–17). Routledge.
- Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior*, 16(1), 107–112. [https://doi.org/10.1016/S0022-5371\(77\)80012-1](https://doi.org/10.1016/S0022-5371(77)80012-1)
- Kahan, D. M. (2013). Ideology, Motivated Reasoning, and Cognitive Reflection: An Experimental Study. *Judgement and Decision-Making*, 8(4), 407–424. <https://doi.org/10.2139/ssrn.2182588>
- Kahan, D. M., Jenkins-Smith, H., & Braman, D. (2011). Cultural cognition of scientific consensus. *Journal of Risk Research*, 14(2), 147–174. <https://doi.org/10.1080/13669877.2010.511246>
- Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgement under uncertainty: heuristics and biases*. Cambridge University Press.

- Kalsnes, B. (2018). Fake News. In Oxford Research Encyclopedia of Communication. Oxford University Press. <https://doi.org/10.1093/acrefore/9780190228613.013.809>
- Moliner, P. (2016). *Psychologie sociale de l'image*. Presses universitaires de Grenoble.
- Nickerson, R. S. (1998). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology*, 2(2), 175–220. <https://doi.org/10.1037/1089-2680.2.2.175>
- Nyhan, B., Reifler, J., Richey, S., & Freed, G. L. (2014). Effective Messages in Vaccine Promotion: A Randomized Trial. *Pediatrics*, 133(4), e835–e842. <https://doi.org/10.1542/peds.2013-2365>
- Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, 147(12), 1865–1880. <https://doi.org/10.1037/xge0000465>
- Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, 25(5), 388–402. <https://doi.org/10.1016/j.tics.2021.02.007>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Simon, H. A. (1957). *Models of Man*. Wiley.
- Suwajanakorn, S., Seitz, S. M., & Kemelmacher-Shlizerman, I. (2017). Synthesizing Obama. *ACM Transactions on Graphics*, 36(4), 1–13. <https://doi.org/10.1145/3072959.3073640>
- Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>
- Unkelbach, C., & Stahl, C. (2009). A multinomial modeling approach to dissociate different components of the truth effect. *Consciousness and Cognition*, 18(1), 22–38. <https://doi.org/10.1016/j.concog.2008.09.006>
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. *Social Media + Society*, 6(1), 205630512090340. <https://doi.org/10.1177/2056305120903408>
- Verdoliva, L. (2020). Media Forensics and DeepFakes: An Overview. *IEEE Journal of Selected Topics in Signal Processing*, 14(5), 910–932. <https://doi.org/10.1109/JSTSP.2020.3002101>