



HAL
open science

Éléments de psychologie des “ fake news ”

Pascal Huguet

► **To cite this version:**

Pascal Huguet. Éléments de psychologie des “ fake news ”. Maud Pélissier; Alexandre Joux. L'actualité d'information au prisme des fake news, L'harmattan, pp.201-222, 2018, Communication et civilisation, 978-2-343-15681-1. hal-04630742

HAL Id: hal-04630742

<https://hal.science/hal-04630742>

Submitted on 1 Jul 2024

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

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

Éléments de Psychologie des « Fake News »

Pascal Huguet

Directeur de Recherche au CNRS

Directeur du Laboratoire de Psychologie Sociale et Cognitive

UMR 6024 Université Clermont Auvergne et CNRS

pascal.huguet@uca.fr

La production et la diffusion délibérées d'informations fausses, truquées ou transformées dans le but notamment de servir des intérêts individuels ou collectifs—le plus souvent idéologiquement saturés—ne constituent évidemment pas des faits nouveaux dans l'histoire des sociétés humaines. Leur amplification à l'heure de la transition numérique peut conduire, à tort, à en situer la cause première dans l'émergence des médias sociaux de contact. Or, la capacité à transformer voire même reconstruire totalement la réalité, à se laisser influencer par des fictions ou à les suggérer voire même les imposer aux autres est, avec le langage, l'une des caractéristiques fortes de l'espèce humaine. C'est que, plus on monte dans l'échelle phylogénétique (plus les organismes biologiques deviennent complexes) et moins les réponses et les adaptations s'avèrent réductibles aux seuls effets des stimulations environnementales (les fameux schémas « stimulus-réponse »), et plus des phénomènes endogènes, cérébraux et psychologiques, s'avèrent déterminants. Les membres de notre espèce, plus que les membres de toute autre espèce, ne répondent donc pas directement à leur environnement mais à la représentation qu'ils s'en forgent sous l'influence de données personnelles, sociales et culturelles. Ces données sont intégrées et retravaillées en permanence par une « machine » cérébrale composée d'environ cent milliards de neurones (et d'avantage encore de cellules gliales) interconnectés, au voisinage d'une centaine de millions de milliards de connexions, qui manifestement ne passent pas leur temps à simplement traiter de l'information, mais aussi à produire sur cette base des significations et autres inférences avec leur cortèges d'émotions (ce que ne savent pas encore faire les machines), le tout avec de puissants effets sur nos comportements. D'où la possibilité de distorsions importantes mais inhérentes, assez paradoxalement, à nos capacités cognitives très évoluées. Nous ne distordons pas la réalité *en dépit de* mais *grâce à* ces mêmes capacités, dont certaines formes primitives s'expriment clairement chez le primate non humain (et d'autres animaux dont certains oiseaux) capables eux aussi d'induire leurs congénères en erreur pour atteindre leurs propres objectifs, lorsqu'ils ne sont pas eux-mêmes victimes de redoutables manipulations, en matière d'accouplement notamment.¹

Il reste que la capacité déjà extraordinaire des humains à convaincre et se laisser convaincre en apesanteur des faits—donc de ce qui est vrai, juste, honnête— est aujourd'hui considérablement augmentée par les nouvelles technologies de l'information et de la communication et leur corollaire : cette obésité informationnelle sans précédent à laquelle une partie de l'humanité au moins est confrontée, et dont les fake news constituent en quelque sorte la masse adipeuse dont il faut apprendre à se débarrasser. Or, affranchies des contraintes spatio-temporelles qui jadis ralentissaient leur progression, ces dernières se propagent aujourd'hui sans aucune limite physique. D'où leur nature généralement explosive, en particulier lorsque les informations fausses en circulation sont dotées d'une charge

¹ Chez le primate, les relations entre le cerveau et son environnement social sont à ce point fondamentales que des neurones différents entrent en jeu pour la réalisation d'activités pourtant strictement identiques selon que ces activités doivent être réalisées en présence ou en l'absence d'un congénère (Demoliens, Isbaine, Takerkart, Huguet, & Boussaoud, 2017).

émotionnelle: d'un jour à l'autre elles surgissent et occupent les esprits, avec nous le verrons une redoutable capacité d'emprise sur les attitudes, opinions et jugements individuels. En bref, si les technologies évoquées antérieurement ne sont pas à l'origine des distorsions et autres reconstructions de la réalité, elles permettent en revanche leur montée en puissance. Dans nos sociétés dites modernes, les croyances de toute nature continueront par conséquent à challenger les connaissances fondées sur des faits avérés, cela d'autant que les stratégies de lutte contre les informations douteuses voire construites de toutes pièces à des fins de manipulation se heurtent à certaines propriétés de notre système cognitif.

Mais que sait-on précisément à propos de l'influence des fake news sur la cognition ? Comment et sous quelles conditions cette influence s'exprime-t-elle exactement ? Est-il possible de s'en préserver, là encore à quelles conditions ? Toutes ces questions imposent un détour par les sciences du comportement, en particulier la psychologie sociale et cognitive qui en effet livre à ce propos un certain nombre d'explications et de repères pour l'action.

1. La construction sociale de la réalité

Dans un article ancien et célèbre consacré à la « communication sociale informelle », le psychologue social Léon Festinger (1950) distinguait trois types de faits impliqués dans les communications interpersonnelles de la vie ordinaire: les faits vérifiables « ici et maintenant » (type 1), les faits seulement « vérifiables en principe » (type 2), et les « faits invérifiables » (type 3). Les résultats rassemblés et discutés dans l'article confortaient son hypothèse d'une pression élevée à la communication s'agissant des faits de type 2 et 3 dans les groupes de pairs où règne plus qu'ailleurs une pression à l'uniformité. En effet, en l'absence de données objectives immédiatement accessibles (type 2), a fortiori de données indisponibles à jamais (type 3), bref en situation d'incertitude les individus pour se forger ou tester leurs opinions ont assez naturellement tendance à sonder celles des pairs ou amis utilisés comme cadre de référence (« qui se ressemblent s'assemblent », cf. aussi Festinger, 1954). Or, dans ce cadre, les individus cherchent généralement à réduire plutôt qu'à augmenter les divergences interpersonnelles pourtant inhérentes aux faits de type 2 et 3, lesquelles avec le temps pourraient en effet fragiliser la cohésion et l'identité du groupe (d'où aussi cette tendance à rejeter les membres déviants à plus ou moins long terme). Une pression à l'uniformité l'emporte donc souvent sur la production et la prise en compte de points de vue réellement contradictoires dans les cercles d'amis ou les groupes de pairs. Comme l'observait Festinger (1950), les opinions ou attitudes exprimées dans ces circonstances sont jugées valides à la seule lumière du consensus intragroupe, qui donne à chacun l'illusion de l'objectivité, a minima le sentiment que l'opinion partagée n'est pas infondée.

Construite en dehors de faits tangibles et vérifiables (au moins dans l'immédiat), cette « objectivité » est renforcée par la répétition des communications les plus consensuelles au sein du groupe, celles qui précisément assurent sa cohésion. Qualifiées d'instrumentales par Festinger, ces communications, à la condition d'être suffisamment récurrentes, fournissent au fil du temps le sentiment d'une grande familiarité à l'égard de l'objet traité (qu'il s'agisse d'un fait divers, d'un problème de société, d'une nouvelle mesure économique ou politique, d'une découverte scientifique ou présentée comme telle, etc). Avec cette familiarité s'installe d'abord un sentiment de maîtrise de l'objet en question s'agissant notamment de ses conséquences immédiates ou anticipées pour le groupe, pour la société, si ce n'est même pour l'avenir de l'humanité toute entière, puis assez rapidement la conviction d'avoir profondément raison.

Et la boucle est bouclée : l'incertitude est réduite par le consensus, qui cimente le groupe, d'où une tendance à reproduire les échanges et autres contenus les plus consensuels, et avec cette reproduction le sentiment d'une familiarité à l'égard de l'objet traité. Ce sentiment à son tour renforce le consensus pour aboutir à des convictions supposées personnelles et surtout rationnelles, bien qu'en réalité biaisées et fortement teintées d'illusionnisme social. Loin d'être ancrées dans une rationalité individuelle, les décisions de chacun sont donc pour une large part socialement et culturellement déterminées ne serait-ce que du fait même de cette tendance à partager et valider nos sentiments, impressions, opinions et autres attitudes avec nos semblables (Hardin & Higgins, 1996 ; Levine, Reisman, & Higgins, 1993 ; Sloman & Fernbach, 2017). Si l'on ajoute qu'il est en règle générale bien plus économique en situation d'incertitude de s'en remettre à ce que pensent les proches (pairs ou amis) plutôt que de raisonner et d'« enquêter » par soi-même, on aura compris à quel point la construction d'une réalité socialement partagée peut maintenir chacun à l'écart de la vérité des faits.

Nous y sommes. Tous ces mécanismes de la vie sociale la plus ordinaire sont à l'œuvre dans les réseaux sociaux numériques de contact en effet dotés de puissants algorithmes affinitaires qui de facto démultiplient les possibilités de mise en relation avec ceux qui partagent les mêmes intérêts, les mêmes motivations, opinions ou attitudes que soi, le tout à l'échelle de la planète. Cette dynamique du « qui se ressemblent s'assemblent...et s'auto-valident » n'est pas nécessairement problématique. Elle peut en revanche le devenir lorsqu'elle enferme les individus dans une pensée de groupe (quasi impénétrable dans le cas extrême des groupes sectaires) sur des thèmes dont les enjeux et les niveaux de complexité réclameraient au contraire une diversité de points de vue pour débattre ou tenter de débattre à partir de l'ensemble des possibles. Les médias sociaux maximisent dans le même temps la fréquence d'exposition aux éléments constitutifs (textes, photos, vidéos) de la pensée de groupe, avec leur charge émotionnelle, et avec cette récurrence confinant parfois à l'enfermement algorithmique émerge la conviction d'une pensée valide (les arguments ou matériaux les plus fréquemment échangés sont aussi jugés les plus crédibles du fait de leur plus grande familiarité). Mais comme le montre la recherche expérimentale en psychologie, la crédibilité et la propagation de fictions et autres informations douteuses voire totalement fausses tient non seulement à la nécessité de construire une réalité socialement partagée, mais aussi aux propriétés mêmes de notre système cognitif.

2. Familiarité cognitive et illusion de vérité

Dans leur étude pionnière sur l'illusion de vérité, Hasher, Goldstein et Toppino (1977) conduisaient des collégiens à évaluer à trois occasions, séparées chacune par des intervalles de deux semaines, la validité de 60 assertions dans des domaines variés (politique, histoire, sciences physiques et biologiques, démographie, géographie, sports, arts, religions et traditions). Ces assertions étaient toutes plausibles mais pour moitié objectivement fausses (e.g., « le divorce n'existe que dans les sociétés techniquement avancées » ; « la république de Chine a été fondée en 1947 »), et une partie seulement était à nouveau présentée lors des deux sessions suivantes. Les résultats montraient que les assertions, vraies ou fausses, auxquelles les participants avaient été exposés trois fois étaient jugées plus valides que celles présentées une seule fois. Le sentiment de familiarité dont bénéficiaient ces assertions du fait de leur répétition les faisait apparaître comme davantage vraies, indépendamment de leur validité réelle (« illusory truth effect »).

Les recherches conduites à la suite de Hasher et collaborateurs pendant une quarantaine d'années permettront de conclure à la robustesse de cet effet d'illusion de vérité fondé sur la familiarité pour des assertions fausses mais la plupart du temps plausibles et donc génératrices d'incertitude (pour une méta-analyse intégrant 51 études, cf. Dechêne, Henkel, Stahl, Hansen et Wänke, 2010). Comme le rappellent Dechêne et al., l'illusion en question s'exprime quelle que soit la nature des messages auxquels l'individu est exposé (déclarations d'opinion sur des questions politiques et sociétales, slogans commerciaux, etc), qu'il s'agisse de messages écrits ou simplement entendus, pour des délais variant de quelques minutes à plusieurs semaines entre leurs répétitions, et parfois même indépendamment de la crédibilité de leur source émettrice (Henkel & Mattson, 2011).²

Les spécialistes du domaine se sont longtemps accordés, bien qu'implicitement, sur l'idée d'une illusion de vérité à laquelle l'individu ne pourrait finalement échapper qu'à la condition de disposer de solides connaissances lui permettant d'identifier toute distorsion ou reconstruction des faits qui lui sont présentés. Mais si elle est pleine de bon sens, cette idée de la connaissance comme rempart absolu aux déformations et reconstructions de toutes sortes est-elle juste pour autant ? On peut en douter et plus que cela encore. En effet, les travaux récents de Fazio, Brashier, Payne, et Marsh (2015) montrent qu'une illusion de vérité peut survenir même lorsque l'individu dispose de la bonne réponse dans ses bases de connaissances en mémoire à long terme. La connaissance exacte des faits ne garantit donc pas nécessairement la possibilité d'en rejeter les versions alternatives lorsqu'elles se répètent (cf. aussi Fazio, Barber, Rajaram, Ornstein, & Marsh, 2013).

Arrêtons-nous un instant sur cette recherche particulièrement saisissante de Fazio et collaborateurs (2015). Dans leur première étude, les auteurs exposaient de jeunes adultes à une série de 176 assertions valides ou invalides et facilement ou difficilement identifiables comme telles, donnant lieu au total à quatre types de propositions: 1- les propositions valides et facilement identifiables comme telles (e.g., l'océan Pacifique est le plus grand océan sur terre); 2- les propositions invalides et facilement identifiables comme telles (l'océan Atlantique est le plus grand océan sur terre); 3- les propositions valides et difficilement identifiables comme telles (Marconi est l'inventeur de la radio sans fil); 4- les propositions invalides et difficilement identifiables comme telles (Bell est l'inventeur de la radio sans fil). Avant de voir et de décider de la validité/invalidité de chacune des 176 propositions, les participants étaient non seulement informés que certaines étaient objectivement vraies et d'autres objectivement fausses, mais qu'ils en avaient déjà vu la moitié, donc 88 propositions, lors de la phase précédente (où il leur avait en effet été demandé de juger l'intérêt de ces 88 propositions). Pour cerner les bases de connaissances des participants, chacun répondait pour terminer à un questionnaire à choix multiples (QCM) adapté à chacune des 176 propositions, par exemple les options « Pacifique », « Atlantique », et « je ne sais pas » suivaient la question « quel est le plus grand océan sur terre ? ». Connaissant à terme les réponses justes et fausses de tous les participants au QCM, et sachant par ailleurs que tous avaient été exposés deux fois à la moitié des 176 propositions, il était donc possible de savoir si cette répétition avait la capacité de générer chez les participants une illusion de vérité pour les assertions objectivement fausses même dans le cas où ils disposaient de la réponse correcte. Comme annoncé antérieurement, ce fût effectivement le cas. Mais surtout, cette illusion de vérité, observée dans les deux études de Fazio et al. (la seconde n'étant qu'une variante de la première), s'exprimait pour les assertions fausses quel que soit leur niveau de difficulté.

² Une source jugée crédible (légitimement ou non) par les uns peut n'inspirer aucune confiance à d'autres, la crédibilité de la source est évidemment toute relative. Elle dépend de la perception que s'en forge l'individu en fonction notamment de ses options idéologiques et de l'intérêt que lui portent ses pairs, amis ou proches.

Les travaux depuis Hasher et al. (1977) partaient de l'idée, très raisonnable, d'une illusion de vérité contrainte par les bases de connaissances du sujet, donc ne survenant qu'à la condition de connaissances non disponibles en mémoire (le sujet ne connaît pas la réponse correcte), ou disponibles mais momentanément non accessibles.³ Or, en montrant que des assertions fausses peuvent être jugées vraies en dépit de la connaissance de la réponse correcte, les résultats de Fazio et al. (2015) suggèrent que le système cognitif en réalité ne s'efforce de mobiliser ce qu'il sait que lorsque le sentiment de familiarité lui fait défaut. En bref, un principe d'économie cognitive (économie fondée dans le cas présent sur la familiarité induite par la répétition d'une information) primerait sur l'effort souvent associé à la mobilisation contrôlée de connaissances en mémoire à long terme. Mais allons un peu plus loin sur cette illusion de vérité. Comment expliquer exactement la vulnérabilité de notre système cognitif face à la répétition d'assertions fausses ?

Cette répétition ne crée pas seulement un sentiment de familiarité (« déjà vu », « déjà entendu ») susceptible de donner davantage confiance, elle facilite aussi le traitement de l'information. L'information répétée est en effet plus facile à traiter aux niveaux à la fois sensoriel-perceptif (dans les modalités auditives ou visuelles) et sémantique c'est-à-dire conceptuel (Whittlesea, 1993). Or, notre système cognitif a une fâcheuse tendance à considérer la facilité de traitement comme une preuve de vérité (Reber & Unkelbach, 2010; Reber & Schwarz, 1999; Kelley & Lindsay, 1993). D'où l'idée qu'il doit être possible de produire une illusion de vérité même sans répétition de l'information à la condition que cette dernière soit présentée dans un format qui facilite son traitement. C'est effectivement ce qu'observent notamment Parks et Toth (2006) en manipulant auprès d'adultes jeunes et âgés le contexte de présentation de slogans publicitaires. Dans la condition « contextes congruents », les slogans étaient présentés après un premier énoncé dont ils semblaient constituer une suite logique. Dans la condition « contextes incongruents », les slogans étaient au contraire présentés après un énoncé sans aucun rapport et créant ainsi une certaine discontinuité visant à en ralentir la compréhension. En accord avec l'hypothèse dite de la facilité de traitement (dans le cas présent la nature plus ou moins fluide d'un traitement conceptuel), les slogans étaient jugés plus crédibles dans la première condition autorisant un traitement fluide que dans la seconde où ce traitement était ralenti. Autre élément en faveur de cette même hypothèse, Wang, Brashier, Wing, Marsh, et Cabeza (2016) ont montré que les conditions de production de l'illusion de vérité activent le cortex perirhinal (lobe temporal médian), une région du cerveau que l'on sait fortement impliquée à la fois dans la reconnaissance visuelle-mnésique, c'est-à-dire le fait de savoir que l'objet ou le message X ou Y a déjà été vu (codage de la familiarité), et dans l'attribution de significations aux objets reconnus. Pour l'anecdote, mais elle aussi consistante avec l'hypothèse de la facilité de traitement, Parks et Toth (2006) ont aussi montré que l'illusion de vérité n'est pas totalement indépendante du choix de la police de caractères utilisée pour la rédaction des messages que l'on souhaite voire se diffuser. Ceux rédigés avec une police fortement contrastée sont perçus plus vrais que ceux—pourtant identiques—rédigés avec une police faiblement contrastée ! Certes, les effets obtenus dans cette autre étude (manipulant donc la nature plus ou moins fluide d'un traitement cette fois perceptif) sont beaucoup plus petits que ceux évoqués antérieurement en rapport avec la fluidité d'un traitement conceptuel, mais ils sont là et on ne peut donc les ignorer. Si l'on ajoute, comme l'ont montré McGlone et Tofiqbakhsh (2000), que les aphorismes, dictons et autres adages sont jugés plus fondés lorsqu'ils intègrent des rimes susceptibles d'en fluidifier le traitement phonologique, on aura compris à quel point

³ Une connaissance peut ne pas être disponible en mémoire à long terme, ou disponible mais transitoirement non accessible, même avec un effort de mémoire, pour différentes raisons par exemple la survenue d'un stress émotionnel et/ou d'une surcharge cognitive (trop d'informations à traiter en même temps pour une capacité attentionnelle nécessairement limitée).

notre cerveau apprécie la facilité en matière de traitement de l'information, facilité de traitement qu'il confond tout aussi facilement avec la vérité.

On perçoit bien à ce stade, comme le propose Daniel Kahneman (2013- célèbre psychologue et prix Nobel d'économie), la nécessité de distinguer deux systèmes de pensée antagonistes, l'un ou l'autre contrôlant nos choix et décisions selon les circonstances : un système rapide/automatique, inconscient, intuitif, naturel, bref très économique (système 1), responsable de nombreux biais cognitifs à l'instar de l'illusion de vérité ; et un système lent, réclamant effort, attention, contrôle, raisonnement, donc cognitivement coûteux (système 2) mais dont la mise route est indispensable pour échapper aux biais et autres illusions en question et les combattre. On comprend bien la difficulté qu'auraient les fake news et autres distorsions de la réalité (rumeurs, stéréotypes sociaux, etc.) pour, en l'absence du système 1, s'imposer et donc se propager dans les réseaux sociaux numériques aussi puissants soient-ils. Dans la vie ordinaire comme dans les réseaux numériques, cette propagation est d'autant plus rapide que les contenus échangés évoquent par exemple le dégoût, la peur, ou la colère (e.g., Guadagno, Rempala, Murphy, & Okdie, 2013 ; Heath, Bell, & Sternberg, 2001), or cette forme de contagion émotionnelle ne serait elle-même guère possible sans le système 1. À un niveau général, il faut bien voir que les croyances fondées sur les fake news n'ont rien de spécifique, elles exploitent les mêmes propriétés du système cognitif que d'autres croyances collectivement partagées comme le sont par exemple les stéréotypes. Ces derniers, nous l'avons montré (en référence aux stéréotypes de genre ou du vieillissement en vigueur dans les sociétés d'Europe de l'Ouest et d'Amérique du Nord) ont comme les fake news la capacité grâce au système 1 de se perpétuer en dépit d'évidences contradictoires, avec de profondes conséquences pour les individus ciblés par le stéréotype, cela même lorsque ces derniers en contestent la validité. Par exemple, le stéréotype d'une infériorité intrinsèque des femmes en sciences, en mathématiques et plus généralement en matière de raisonnement interfère dans tous ces domaines avec leurs performances, dès l'enfance et même lorsque les sujets montrent en parallèle des positionnements contre-stéréotypiques (Huguet et Régner, 2007, 2009 ; Régner, Smeding, Gimmig, Thinus-Blanc, Monteil, & Huguet, 2010 ; Régner, Steele, Ambady, Thinus-Blanc, & Huguet, 2014; cf. aussi Pansu, Régner, Max, Colé, & Huguet, 2016 pour des observations comparables avec un stéréotype défavorable aux hommes). De même, les stéréotypes du vieillissement sont si puissants (en particulier la croyance en un déclin cognitif pathologique avec l'âge) qu'ils interfèrent avec les performances des âgés au moment des tests neuropsychologiques utilisés pour le dépistage précoce de la maladie d'Alzheimer, créant des proportions parfois spectaculaires (jusqu'à 30%) de « faux positifs » : les personnes testées apparaissent sur le chemin de la maladie alors qu'elle ne le sont pas (Mazerolle, Régner, Morisset, Rigalleau, & Huguet, 2012 ; Mazerolle, Régner, Rigalleau, & Huguet, 2015 ; Régner, Mazerolle, Sambuchi, Rigalleau, Paccalin, Clarys, Michel, & Huguet, 2016 ; Mazerolle, Régner, Barber, Paccalin, Miazola, Huguet, & Rigalleau 2017). Les stéréotypes une fois activés ont pour conséquence à la fois de renforcer les automatismes (système 1) même lorsqu'ils sont inopportuns, et d'altérer la capacité de contrôle (système 2), d'où évidemment des dégâts considérables en situation de tests supposés mesurer objectivement les capacités du système cognitif. Au passage, nous avons montré que cette même capacité de contrôle, requise pour échapper aux automatismes du système 1 (et partant à l'emprise des fake news), est considérablement réduite en présence d'individus capables de sanctions même strictement symboliques à l'égard du sujet testé (Belletier, Davranche, Tellier, Dumas, Hasbroucq, Vidal, & Huguet, 2015), un phénomène phylogénétiquement ancien puisque détectable aussi chez le primate non humain (Huguet, Barbet, Belletier, Monteil, & Fagot, 2014).

3. La fabrication collective de faux souvenirs

On se souvient (cf. section 1) qu'une pression à l'uniformité l'emporte souvent sur la production et la prise en compte de points de vue contradictoires dans les groupes d'amis ou de pairs comme ceux à l'œuvre dans les réseaux sociaux numériques fondés sur un algorithme affinitaire (mes amis, les amis de mes amis, etc). On se souvient aussi que cette pression au service de la cohésion du groupe favorise la répétition des échanges les plus consensuels. Or, nous l'avons vu (cf. section 2), avec cette répétition s'installe un sentiment de familiarité propice aux illusions de vérité même lorsque l'individu dispose de connaissances qui devraient l'en protéger. Comme nous allons le voir maintenant, non seulement les bases de connaissances peuvent s'avérer inopérantes face à certaines distorsions de la réalité, mais ces dernières peuvent infiltrer la mémoire elle-même sous la forme de faux souvenirs, donc créer l'illusion de se souvenir d'un événement qui en réalité n'a jamais eu lieu.

Déjà invoqués par Pierre Janet et Sigmund Freud en référence à certains troubles mentaux, les faux souvenirs sont à partir des années soixante-dix étudiés de manière expérimentale par Elizabeth Loftus pour en montrer l'existence extrêmement problématique dans les témoignages de personnes saines en contexte judiciaire (e.g., Loftus, 2003, 2005). Dans ses premiers travaux sur cette question (Loftus, 1975), Loftus exposait près de 500 sujets à des films d'événements complexes et rapides, tels que des accidents d'automobile ou des perturbations en classe, puis les interrogeait sans délai sur les scènes visualisées à partir de questions impliquant des présupposés vrais ou faux, comme par exemple l'intégration à la question posée d'un objet existant ou n'existant pas dans la scène visuelle. Cette manipulation s'avérait suffisante pour que les sujets prétendent assez souvent avoir vu cet objet. Les questions posées après un événement peuvent donc introduire de nouvelles informations, fondées ou non, qui sont ensuite ajoutées à la représentation de l'événement, causant ainsi sa reconstruction ou son altération. Comme l'a aussi montré Loftus dans le cadre d'affaires judiciaires, en biaisant les témoignages, ces faux souvenirs dont elle révèle donc l'expression en dehors même des maladies mentales (une idée révolutionnaire à l'époque), peuvent tout aussi bien conduire à l'accusation d'innocents qu'à la remise en liberté de personnes dangereuses pour la collectivité. D'où plusieurs dizaines d'années de travaux sous son impulsion pour améliorer les témoignages en particulier oculaires via de nouvelles techniques d'interrogation.

Mais si la mémoire est à ce point malléable (sans même évoquer les faux souvenirs en mémoire autobiographique, e.g., Wade & Garry, 2005), on peut facilement imaginer ses altérations sous l'emprise de dynamiques sociales telle que la pression à l'uniformité dans les réseaux sociaux numériques. Encore une fois, cette pression au service de la cohésion du groupe favorise la répétition des échanges (arguments discursifs, photos, vidéos) les plus consensuels et in fine la survenue d'illusions de vérité. Mais cette répétition influence tout aussi naturellement l'encodage, le stockage, et la récupération de l'information en mémoire. En effet, relativement aux informations peu échangées, dont la mémorisation décline naturellement avec le temps, celles répétées collectivement—mêmes fausses—sont mieux mémorisées (mieux encodées, mieux stockées et mieux récupérées) à la fois pour ceux qui les font circuler et pour ceux auxquels ces informations s'adressent (e.g., Coman et Hirst, 2012). Cette fabrication collective de faux souvenirs est particulièrement à l'œuvre entre proches (Gabbert, Wright, Memon, Skagerberg, Jamieson, 2014) et l'est donc davantage entre individus du même groupe partageant une même identité qu'entre individus de groupes différents (Coman et Hirst, 2015). Et comme dans le cas de l'illusion de vérité (section 2), les

faux souvenirs collectivement fabriqués peuvent opérer même lorsque l'individu dispose de connaissances en principe suffisantes pour l'en protéger.

C'est ce que montrent notamment Edelson, Sharot, Dolan et Dudai (2011). Dans cette étude, les sujets visionnaient un documentaire par groupes de cinq (phase 1) puis répondaient individuellement trois jours plus tard (phase 2) à 400 questions sur ce documentaire tout en estimant la confiance accordée à chacune de leur réponse. Une semaine après la séance de visionnage (phase 3), les sujets répondaient à nouveau, toujours individuellement, à une sélection de questions (80 au total) auxquelles, en phase 2, ils avaient répondu correctement avec une relative certitude. Avant de répondre, néanmoins, ils étaient exposés aux réponses fausses (fabriquées pour les besoins de l'expérience) supposément fournies par les membres de leur groupe (rencontrés en phase 1). Les sujets s'y conformaient dans environ 70% des cas alors même qu'ils se montraient capables initialement de répondre correctement.

Ces résultats, et beaucoup d'autres (e.g., Gabbert, Memon, Allan, 2003 ; Gabbert, Memon, Allan, Wright, 2004 ; Horry, Palmer, Sexton, & Brewer, 2012 ; Wright, Self, & Justice, 2010), montrent à quel point la mémoire des faits peut être distordue sous l'influence du groupe, même composé de personnes non familières (contrairement aux groupes d'amis), à la condition néanmoins que ces dernières soutiennent unanimement ou presque la réponse fausse.⁴ La dynamique d'influence induisant les distorsions en mémoire ne relève plus dans ce cas d'une pression ou tendance à l'uniformité, comme celle à l'œuvre dans les groupes d'amis, mais d'une tendance à la conformité (deux tendances à ne pas confondre donc), laquelle peut-être sous-tendue par deux formes de dépendance au groupe : la « dépendance informationnelle », en particulier lorsque l'objet du jugement est complexe et place le sujet en situation d'incertitude (faits de type 2 et 3 selon Festinger, 1950, cf. section 1), et la dépendance normative lorsque le sujet, à tort ou à raison, est conduit à craindre que son propre avis ne suscite de la part des autres membres du groupe une certaine forme de sanction. La conformité peut alors s'exprimer aussi bien pour les faits de type 2 et 3 que pour ceux de type 1 pourtant les plus immédiatement vérifiables. La tendance à la conformité est encore plus forte lorsque se conjuguent ces deux formes de dépendance, c'est-à-dire lorsque l'individu n'a pas de position ferme face aux faits examinés et dans le même temps anticipe des sanctions en cas d'avis personnel différent de l'avis majoritaire. Quelle que soit la dynamique sous-jacente (dépendance informationnelle et/ou normative), la connaissance des opinions, avis, jugements et autres visions d'autrui augmente la probabilité d'interférences rétroactives, autrement dit de distorsions de la mémoire des événements sous l'influence des échanges qu'ils suscitent et qui en modifient le contenu et donc la signification.

Ces distorsions, en particulier celles fondées sur la conformité, sont moins probables lorsque le groupe intègre des membres affichant leur désaccord avec la majorité et soutenant une vision alternative des faits (Walther, Bless, Strack, Rackstraw, Wagner, et Werth, 2002). Or, la fréquence de voix discordantes augmente mécaniquement avec la taille des groupes. Il

⁴ Ces résultats font écho aux phénomènes décrits par Salomon Asch au milieu du siècle dernier dans ses fameuses recherches sur la conformité (Asch, 1951, 1955). Asch exposait de jeunes adultes à un matériel non ambigu constitué à chaque essai de deux images : sur la première apparaissait une ligne verticale isolée et sur la seconde trois lignes verticales de différentes longueurs (lignes A ; B, ou C) dont une seule était de toute évidence identique à la ligne présentée sur la première image. La tâche des sujets consistait simplement à repérer cette identité et donc dire quelle ligne sur les trois présentées dans la seconde image correspondait à celle montrée sur la première. Lorsqu'ils répondaient seulement en présence de l'expérimentateur, les sujets présentaient un taux d'erreur inférieur à 1%. Lorsqu'ils répondaient en présence d'autres personnes d'âge comparables mais non familières—des complices de l'expérimentateur dont le rôle consistait à donner des réponses fausses à certains essais—les sujets montraient un taux d'erreur de 36.8% au total, et 75% d'entre eux se conformaient au moins une fois à la réponse erronée des autres membres du groupe. Dans les deux cas (Edelson et al., 2011 ; Asch, 1951, 1955) une tendance à se conformer aux réponses majoritaires du groupe était donc observée.

est tentant par conséquent d'imaginer des débats plus ouverts, plus contradictoires, plus riches, et des distorsions de la mémoire moins nombreuses au sein des réseaux sociaux numériques. Affranchis de toutes contraintes spatio-temporelles, ces derniers en effet démultiplient les possibilités d'interactions relativement aux petits groupes de la vie ordinaire. Mais nous l'avons vu, cette possibilité offerte par la transition numérique est ruinée par l'attrait des relations affinitaires (« qui se ressemblent s'assemblent », cf. section 1), une constante comportementale bien intégrée par les concepteurs des médias sociaux de contact. Par ailleurs, les travaux visant à reproduire de manière contrôlée la dynamique des communications dans les réseaux sociaux numériques aboutissent à la même conclusion : les interactions sociales dans ce cadre ne font que reproduire, certes à beaucoup plus grande échelle, la plupart des phénomènes comportementaux à l'œuvre dans les petits groupes de la vie ordinaire. Ces travaux n'en sont pas moins intéressants. Dans les réseaux numériques, en effet, les communications interpersonnelles aboutissent, en raison de pressions liées à l'uniformité ou à la conformité, à l'émergence de « clusters », sortes d'agrégats spatialement et temporellement situés, composés d'individus qui après quelques interactions sur tel ou tel événement en présentent une vision relativement homogène, avec des conséquences identiques à celles observées dans les petits groupes en matière de distorsions de la mémoire des faits (Coman, Momennejad, Drach et Geana, 2016).

Dès les années 90, des travaux inspirés de la théorie des systèmes dynamiques complexes et visant à simuler ces phénomènes d'influence sociale à très grande échelle (plusieurs milliers d'individus en interaction sur des périodes de temps très longues) permettaient d'entrevoir cette réalité désormais bien admise : même dans le cas d'une distribution au départ totalement aléatoire des attitudes et opinions individuelles dans les espaces de communication (numériques ou ordinaires), des clusters émergent rapidement, et se consolident avec le temps, préservant de facto une certaine diversité au sein du système (Latané & Bourgeois, 1996 ; Latané, Nowak, & Liu, 1994 ; Lewenstein, Nowak, & Latané, 1992 ; Nowak, Szamrej, & Latané, 1990). Dans nos propres travaux (Huguet, Latané, & Bourgeois, 1998), chaque thème de discussion quelle que soit sa nature (droits de l'homme, questions politiques diverses, questions basement matérielles de la vie ordinaire) générait son lot de clusters. Assez ironiquement, ces clusters formés sous pression à l'uniformité ou à la conformité induisent souvent chez les individus qui en sont à l'origine le sentiment que leur vision est majoritaire à l'échelle du réseau tout entier, alors que cette vision en réalité ne dépasse pas l'échelle du cluster lui-même. Mais surtout, à cette diversité, il faut ajouter la possibilité qu'avec le temps les clusters (avant de disparaître et de laisser leurs places à d'autres) non seulement se renforcent mais aussi combinent et recombinent leurs éléments thématiques voire en ajoutent de nouveaux. Ce processus en constante évolution peut donner lieu à d'authentiques représentations sociales (Huguet & Latané, 1996 ; Latané, 1996), qui même au départ très minoritaires peuvent finir par s'imposer pour le meilleur et pour le pire (Butera, Huguet, Mugny, & Perez, 1994 ; Huguet, Mugny, & Pérez, 1992 ; Moscovici, Van Avermaet, & Mugny, 1985, Mugny & Pérez, 1991 ; cf. aussi Brandstätter et al. 1991).

En bref, il faut bien voir que les illusions de vérité (cf. section 2) et autres distorsions de la mémoire (cf. section 3) sous l'emprise de fausses informations en circulation sont généralement « encapsulés » au sein de clusters dont les modes de raisonnement et de fonctionnement peuvent être totalement aveugles aux informations d'arrière plan provenant d'autres sources que celles propres aux clusters eux-mêmes (cf. Difonzo, Bourgeois, Suls, Homan, Stupak, Brooks, Ross, Bordia, 2013). D'où leur inertie, surtout s'agissant de faits seulement vérifiables en principe et plus encore de faits invérifiables (faits de type 2 et 3, cf. section 1) sur lesquels les fausses informations peuvent aisément prospérer et résister aux

contre-attaques.

4. En conclusion : sur la déconstruction des fake news

La question qui se pose à l'issue de cette présentation de quelques éléments de psychologie des fake news est celle de leur déconstruction (nos collègues anglo-saxons utilisent le terme de « debunking » plus proche de la notion de démystification). C'est que, perméable aux théories infondées, aux fausses informations, aux rumeurs, aux stéréotypes et fictions de toutes sortes, le système cognitif se révèle aussi le plus souvent incapable d'en intégrer efficacement les versions contradictoires (pour des méta-analyses, cf. Chan, Jones, Jamieson, & Albarracin, 2017 ; Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012). Même officiellement invalidées, les versions initiales continuent à se propager et à influencer les décisions individuelles (cf. aussi Pantazi, Kissine, & Klein, 2018). Sans grande surprise, plus les contre-arguments utilisés pour cette invalidation contreviennent aux positionnements idéologiques et autres visions du monde des individus ciblés, et plus faible est leur efficacité. En réalité cependant, et de manière plus contre-intuitive, les tentatives d'invalidation s'avèrent assez souvent inefficaces même lorsque le sujet croit et comprend les contre arguments avancés et se montre par ailleurs capable de s'en souvenir. Comment expliquer l'influence persistante des fausses informations dans ces conditions ?

D'abord, n'oublions pas que les jugements, décisions et autres choix individuels, surtout ceux réalisés dans l'incertitude, ne sont pas réductibles au travail de la pensée qui se concentrerait en quelque sorte sur elle-même, pesant arguments et contre-arguments en apesanteur de toute influence sociale. En situation d'incertitude, les individus sont au contraire sous l'emprise de dynamiques d'influence qui contribuent à façonner leurs options et autres visions du monde (section 1). De nombreux biais cognitifs s'expriment alors à plein, comme les illusions de vérité et les distorsions de la mémoire (sections 2 et 3), qui offrent une certaine résistance aux contre-arguments même scientifiquement avérés. Par ailleurs, une fois reconnue la validité des contre-arguments avancés par des sources que généralement il ne maîtrise pas, l'individu doit faire face à une histoire bien plus complexe : reconnaître qu'il s'est trompé, avec le cas échéant des conséquences pour ses proches (comme par exemple dans le cas de la décision de ne pas avoir fait vacciner son enfant en raison de l'implication supposée de certains vaccins dans l'autisme), revenir le cas échéant sur des pratiques bien installées, soutenir à son tour l'information contradictoire face à ses pairs, briser une certaine harmonie lorsqu'il s'agit du groupe d'amis, voire se soustraire au confort de la conformité ; bref sortir d'une économie cognitive au profit d'une explication alternative plus riche et avérée mais aussi souvent plus difficile à comprendre et à transmettre. La tentation est alors assez forte de ne rien changer, sinon de fortifier les croyances menacées d'invalidation. D'où cette autre observation récurrente dans les travaux consacrés à la déconstruction des fake news : les efforts d'invalidation provoquent souvent des « retours de flamme » (backfire effects) sous la forme d'un renforcement de l'intérêt accordé aux informations que l'on s'efforce par ailleurs de discréditer. La répétition des contre-arguments utilisés pour l'invalidation ne change rien. Cette répétition peut même susciter auprès du public visé une activité de même nature (contre-argumentation) mais en faveur de l'information initiale (fausse), avec pour conséquence d'augmenter l'immunité de l'individu à l'égard de toute tentative de déstabilisation. On retiendra à ce sujet une asymétrie remarquable : si la crédibilité accordée aux informations fausses augmente avec leur répétition (cf. section 2), il n'en va pas de même pour la crédibilité des contre-arguments utilisés pour les falsifier.

Comme le montrent les travaux dans ce domaine (Chan et al., 2017 ; Lewandowsky et al., 2012), pour avoir une chance d'aboutir, la source de l'invalidation ne doit pas se limiter à contester la validité de l'information ou de la théorie jugée fautive voire dangereuse en la déclarant officiellement incorrecte. Elle doit aussi substituer aux éléments les plus problématiques qu'elle s'attache à discréditer d'autres éléments tout aussi compréhensibles et dans le même temps expliquer les motivations des individus ou groupes à l'origine de la manipulation de l'opinion, le tout de manière succincte (toujours ce principe d'économie cognitive). On perçoit à quel point cette entreprise de déconstruction peut être difficile, cela d'autant qu'elle doit aussi s'efforcer de tenir compte des caractéristiques socio-économiques et idéologiques de la cible visée pour s'y adapter sur le fond et sur la forme, et donner lieu par conséquent à des stratégies d'invalidation sensiblement différentes d'un segment de la population à un autre. D'où la nécessité de progresser dans la mise en place de plateformes permettant la détection des fausses informations en amont, donc avant toute possibilité de leur encodage par le système cognitif. D'où aussi l'importance de promouvoir une éducation aux médias et à la citoyenneté impliquant notamment des actions de nature à développer l'esprit critique assez tôt en contexte scolaire.

Références

- Asch, S.E. (1951). Effects of group pressure on the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership and men* (pp. 177–190). Pittsburgh, PA: Carnegie Press.
- Asch, S.E. (1955). Opinions and social pressure. *Scientific American*, 193, 31-35.
- Belletier, C., Davranche, K., Tellier, I., Dumas, F., Hasbroucq, T., Vidal, F., & Huguet, P. (2015). Choking under monitoring pressure: Being watched by the experimenter reduces executive attention. *Psychonomic Bulletin & Review*, 22, 1410-1416.
- Brandstätter, V., Ellemers, N., Gaviria, E., Giosue, F., Huguet, P., Kroon, M., Morchain, P., Pujal, M., Rubini, M., Mugny, G., & Pérez J.A. (1991). Indirect majority and minority influence: An exploratory study. *European Journal of Social Psychology*, 21, 196-211.
- Butera, F., Huguet, P., Mugny, G., & Pérez, J. (1994). Socio-epistemic conflict and constructivism. *Swiss Journal of Psychology*, 53, 229-239.
- Chan, M. P. S., Jones, C. R., Hall Jamieson, K., & Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological Science*, 28, 1531-1546.
- Coman, A., & Hirst, W. (2012). Cognition through a social network: The propagation of induced forgetting and practice effects. *Journal of Experimental Psychology : General*, 141, 321-336.
- Coman, A., & Hirst, W. (2015). Social identity and socially shared retrieval-induced forgetting: The effects of group membership. *Journal of Experimental Psychology : General*, 144, 717-722.
- Coman, A., Momennejad, I., Drach, R. D. & Geana, A. (2016). Mnemonic convergence in social networks: The emergent properties of cognition at a collective level. *Proceeding of the National Academy of Sciences, USA*, 113, 8171–8176.
- Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (2010). The truth about the truth: A meta-analytic review of the truth effect. *Personality and Social Psychology Review*, 14, 238–257.
- Demolliens M., Isbaine F., Takerkart S., Huguet P.*, Boussaoud D.* (2017 *equal contribution). Social and asocial prefrontal cortex neurons: A new look on social facilitation and the social brain. *Social Cognitive and Affective Neuroscience*, 12, 1241-1248.
- Difonzo, N., Bourgeois, M. J., Suls, J., Homan, C., Stupak, N., Brooks, B.P., Ross, D.S., Bordia, P. (2013). Rumor clustering, consensus, and polarization: Dynamic social impact and self-organization of hears. *Journal of Experimental Social Psychology*, 49, 378-399.
- Edelson, M., Sharot, T., Dolan, R. J. & Dudai, Y. (2011). Following the crowd: Brain substrates of long-term memory conformity, *Science*, 333, 108–111.
- Fazio, L.K., Barber, S.J., Rajaram, S., Ornstein, P.A., & Marsh, E.J. (2013). Creating illusions of knowledge: Learning errors that contradict prior knowledge. *Journal of Experimental Psychology : General*, 142, 1-5.
- 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, 993-1002.
- Festinger, L. (1950). Informal social communication. *Psychological Review*, 57, 271-282.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117-140.
- Gabbert, F., Memon, A., & Allan, K. (2003). Memory conformity: Can eyewitnesses influence each other's memories for an event ? *Applied Cognitive Psychology*, 17, 533–543.
- Gabbert, F., Memon, A., Allan, K., Wright, D.B. (2004). Say it to my face: Examining the effects of socially encountered misinformation. *Legal and Criminological Psychology*, 9, 215–227.

Gabbert, F., Wright, D.F., Memon, A., Skagerberg, E.M., & Jamieson, K. (2013). Memory Conformity Between Eyewitnesses. *Court Review: The Journal of the American Judges Association*, 48, 36-43.

Guadagno, R.E., Rempala, D.M., Murphy, S., & Okdie, B.M. (2013). What makes a video go viral ? An analysis of emotional contagion and Internet memes. *Computers in Human Behavior*, 29, 2312-2319.

Hardin, C. D., & Higgins, E. T. (1996). Shared reality: How social verification makes the subjective objective. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition, Vol. 3. The interpersonal context* (pp. 28-84). New York, NY, US: Guilford Press.

Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior*, 16, 107- 112.

Heath, C., Bell, C., & Sternberg, E. (2001). Emotional selection in memes: The case of urban legends. *Journal of Personality and Social Psychology*, 81, 1028-1041.

Henkel, L.A., & Mattson, M.E. (2011). Reading is believing: the truth effect and source credibility. *Consciousness and Cognition*, 20, 1705-1721.

Huguet, P., Barbet, I., Belletier, C., Monteil, J.-M., & Fagot, J. (2014). Cognitive control under social influence in baboons. *Journal of Experimental Psychology : General*, 143, 2067-2073.

Huguet, P., & Latané, B. (1996). Social representations as dynamic social impact. *Journal of Communication*, 46, 57-63.

Huguet, P., Latané, B., & Bourgeois, M. (1998). The Emergence of a social representation of human rights via interpersonal communication: Empirical evidence for the convergence of two theories. *European Journal of Social Psychology*, 28, 831-846.

Huguet, P., Mugny, G., & Pérez, J.A. (1992). Influence minoritaire et processus de décentration. *Bulletin de Psychologie*, 405, 155-163.

Huguet, P., Régner, I. (2007). Stereotype threat among school girls in quasi-ordinary classroom circumstances. *Journal of Educational Psychology*, 99, 545-560.

Huguet, P., & Régner, I. (2009). Counter-stereotypic beliefs in math do not protect school girls from stereotype threat. *Journal of Experimental Social Psychology*, 45, 1024-1027.

Horry, R., Palmer, M.A., Sexton, M.L., Brewer, N. (2012). Memory conformity for confidently recognized items: The power of social influence on memory reports. *Journal of Experimental Social Psychology*, 48, 783–786.

Kahneman, D. (2011). *Thinking, fast and slow*. London, Allen Lane : Penguin.

Kelley, C.M., & Lindsay, D.S. (1993). Remembering mistaken for knowing. *Journal of Memory and Language*, 32, 1-24.

Latané, B. (1996). Dynamic social impact : The creation of culture by communication. *Journal of Communication*, 46, 13–25.

Latané, B., & Bourgeois, M. J. (1996). Experimental evidence for dynamic social impact: The emergence of subcultures in electronic groups. *Journal of Communication*, 46, 35–47.

Latané, B., Nowak, A., & Liu, J. H. (1994). Measuring emergent social phenomena: Dynamism, polarization, and clustering as order parameters of dynamic social systems. *Behavioral Science*, 39, 1-24.

Levine, J.M., Reisnick, L.B., & Higgins, E.T. (1993). Social foundations of cognition. *Annual Review of Psychology*, 44, 585-612.

Lewandowsky, S., Ecker, U.K., Seifert, C.M., Schwarz, N., & Cook, J. (2012). Misinformation and Its Correction: Continued Influence and Successful Debiasing. *Psychological Science in the Public Interest*, 13, 106-131.

Lewenstein, M., Nowak, B., & Latané, B. (1992). Statistical mechanics of social impact. *Physical Review A*, 45, 703-716.

- Loftus, E.F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, 7, 560-572.
- Loftus, E.F. (2003). Our changeable memories: legal and practical implications. *Nature Reviews Neuroscience*, 4, 231–234.
- Loftus, E.F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory*, 12, 361–366.
- Mazerolle, M., Régner, I., Barber, S. J., Paccalin, M., Miazola, A-C., Huguet, P., & Rigalleau, F. (2017). Negative aging stereotypes impair performance on brief cognitive tests used to screen for predementia. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 72, 932-936.
- Mazerolle, M., Régner, I., Morisset P., Rigalleau, F., & Huguet, P. (2012) Stereotype threat strengthens automatic recall and undermines controlled processes in the elderly. *Psychological Science*, 23, 723-727.
- Mazerolle, M., Régner, I., Rigalleau, F., & Huguet, P. (2015). Stereotype threat alters the subjective experience of memory. *Experimental Psychology*, 62, 395-402.
- McGlone, M.S., & Tofiqbakhsh, J. (2000). Birds of a feather flock conjointly (?): Rhyme as reason in aphorisms. *Psychological Science*, 11, 424-428.
- Moscovici, S., Van Avermaet, E., & Mugny, G. (1985, Eds). *Perspectives on minority influence*. Cambridge: Cambridge University Press.
- Mugny, G., & Pérez, J. A. (1991). *The social psychology of minority influence*. Cambridge: Cambridge University Press.
- Nowak, A., Szamrej, J., & Latané, B. (1990). From private attitude to public opinion: A dynamic theory of social impact. *Psychological Review*, 97, 362-376.
- Pansu, P., Régner, I., Max, S., Colé, P., & Huguet, P. (2016). A Burden for the boys: Stereotype threat in reading. *Journal of Experimental Social Psychology*, 65, 26-30.
- Pantazi, M., Kissine, M., & Klein, O. (2018). The Power of the truth bias: False information affects memory and judgment even in the absence of distraction. *Social Cognition*, 36, 167-198.
- Toth, J.P. & Parks, C.M. (2006). Effects of age on familiarity in the process-dissociation procedure: The role of noncriterial recollection. *Memory & Cognition*, 34, 527-537.
- Reber, R. & Schwarz, N. (1999). Effects of Perceptual Fluency on Judgments of Truth. *Consciousness and Cognition*, 8, 338-342.
- Reber, R., & Unkelbach, C. (in press). The epistemic status of processing fluency as source for judgments of truth. *Review of Philosophy and Psychology*, 1, 563-581.
- Régner, I., Mazerolle, M., Sambuchi, N., Rigalleau, F., Paccalin, M., Clarys, D., Michel, B., & Huguet, P. (2016). Aging stereotypes must be taken into account for the diagnosis of prodromal and early Alzheimer's disease. *Alzheimer Disease & Associated Disorders*, 30, 77-79.
- Régner, I., Smeding, A., Gimmig, D., Thinus-Blanc, C., Monteil, J.-M., & Huguet, P. (2010). Individual differences in working memory moderate stereotype threat effects. *Psychological Science*, 21, 1646-1648.
- Régner, I., Steele J.R., Ambady, N., Thinus-Blanc, C., & Huguet, P. (2014). Our future scientists: A review of stereotype threat in girls from early elementary school to middle school. *International Review of Social Psychology*, 27, 13-51.
- Sloman, S., & Fernbach, P. (2017). *The knowledge illusion : Why we never think alone*. New York : Riverhead Books.
- Wade, K.; Garry, M. (2005). Strategies for verifying false autobiographical memories. *The American Journal of Psychology*, 118, 587–602.

Walther, E.; Bless, H.; Strack, F.; Rackstraw, P.; Wagner, D. & Werth, L. (2002). Conformity effects in memory as a function of group size, dissenters, and uncertainty. *Applied Cognitive Psychology*, 16, 793–810.

Wang, W.C., Brashier, N.M., Wing, E.A., Marsh, E.J., & Cabeza, R. (2016). On Known Unknowns: Fluency and the Neural Mechanisms of Illusory Truth. *Journal of Cognitive Neuroscience*, 28, 739-746.

Whittlesea, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 1235-1253.

Wright, D.B.; Self, G., & Justice, C. (2000). Memory conformity: Exploring misinformation effects when presented by another person. *British Journal of Psychology*, 91, 189-202.