

## Pour une éducation aux hypertrucages malveillants et un développement de l'agentivité dans les contextes numériques

## Education on malicious deepfakes and developing agency in digital environments

## Por una educación a los hipertrucajes malintencionados y un desarrollo de la agencidad en los contextos digitales

Nadia Naffi, Ann-Louise Davidson, Sylvie Barma, Marie-Claude Bernard, Nicolas Brault, François Berger and Aude Gagnon-Tremblay

Volume 49, Number 2, Fall 2021

La complexité de l'éducation à la citoyenneté numérique : des enjeux sociétaux, éducatifs et politiques

URI: <https://id.erudit.org/iderudit/1085307ar>

DOI: <https://doi.org/10.7202/1085307ar>

[See table of contents](#)

Publisher(s)

Association canadienne d'éducation de langue française

ISSN

1916-8659 (digital)

[Explore this journal](#)

Cite this article

Naffi, N., Davidson, A.-L., Barma, S., Bernard, M.-C., Brault, N., Berger, F. & Gagnon-Tremblay, A. (2021). Pour une éducation aux hypertrucages malveillants et un développement de l'agentivité dans les contextes numériques. *Éducation et francophonie*, 49(2). <https://doi.org/10.7202/1085307ar>

Article abstract

Online disinformation is not on the decline. The term *infocalypse* was actually coined by Schick (2020) to describe the seriousness of disinformation in the form of malicious deepfakes, which obscure what was previously considered authentic. This reflective article explores the complexity of the deepfake phenomenon to emphasize the urgent need for research that directly addresses educators and decision-makers. Under the lens of citizenship education, we argue that teachers, parents and educational actors must support students in developing critical thinking, agency in the digital environment and the ability to recognize deepfakes so that they can protect themselves and counteract this phenomenon. To answer the question "Why educate young people on malicious deepfakes?" we started by exploring this phenomenon and its dangers in the context of disinformation. First, this article exposes how deepfakes destabilize the perceptual foundations of human knowledge and discusses their threat to our societies. Secondly, it presents avenues considered and draws attention to the urgency of alerting young people to these dangers and involving them in counteracting this form of disinformation. Lastly, the article offers avenues for research in digital citizenship.

# Pour une éducation aux hypertrucages malveillants et un développement de l'agentivité dans les contextes numériques

*Education on malicious deepfakes and developing agency in digital environments*

*Por una educación a los hipertrucajes malintencionados y un desarrollo de la agenceidad en los contextos digitales*

**Nadia NAFFI**

Université Laval, Québec, Canada

**Ann-Louise DAVIDSON**

Université Concordia, Québec, Canada

**Sylvie BARMA**

Université Laval, Québec, Canada

**Marie-Claude BERNARD**

Université Laval, Québec, Canada

**Nicolas BRAULT**

Université Laval, Québec, Canada

**François BERGER**

Université Laval, Québec, Canada

**Aude GAGNON-TREMBLAY**

Université Laval, Québec, Canada

---

**Notes éditoriales**

**Note des auteurs et auteurs**

La revue de la littérature présentée dans cet article de réflexion a été financée par le CRSH.

**Authors' note**

The literature review presented in this reflective article was funded by SSHRC.

**Nota de los autores**

La revisión literaria efectuada en este artículo reflexivo ha sido financiada por SSHRC.

**Résumé**

La désinformation en ligne n'est pas en perte de vitesse. Le terme *infocalypse* a d'ailleurs été inventé par Schick (2020) pour décrire la gravité de la désinformation sous forme d'hypertrucages malveillants (*deepfakes*) qui obscurcissent ce qui était auparavant considéré comme authentique. Cet article de réflexion circonscrit la complexité du phénomène des *deepfakes* afin de mettre en relief le besoin urgent de mener des recherches qui interpellent directement les acteurs et les actrices du monde de l'éducation. Sous la lentille de l'éducation à la citoyenneté, nous argumentons que le personnel enseignant, les parents ainsi que les actrices et les acteurs éducatifs doivent accompagner les jeunes dans le développement de l'esprit critique, de l'agentivité dans le contexte numérique et du discernement à l'égard des hypertrucages malveillants afin qu'ils puissent se protéger et contrer ce phénomène. Pour répondre à la question « Pourquoi éduquer les jeunes aux hypertrucages malveillants? », nous avons commencé par circonscrire ce phénomène et ses dangers dans un contexte de désinformation. Dans un premier temps, cet article expose la façon dont les *deepfakes* ébranlent les bases perceptuelles liées à la connaissance humaine et discute de leur menace pour nos sociétés. Deuxièmement, il présente des pistes envisagées et attire l'attention sur l'urgence de sensibiliser les jeunes à ces dangers et de les impliquer pour contrer cette forme de désinformation. Finalement, l'article propose des pistes de recherche à explorer en citoyenneté numérique.

**Abstract**

Online disinformation is not on the decline. The term *infocalypse* was actually coined by Schick (2020) to describe the seriousness of disinformation in the form of malicious deepfakes, which obscure what was previously considered authentic. This reflective article explores the complexity of the deepfake phenomenon to emphasize the urgent need for research that directly addresses educators and decision-makers. Under the lens of citizenship education, we argue that teachers, parents and educational actors must support students in developing critical thinking, agency in the digital environment and the ability to recognize deepfakes so that they can protect themselves and counteract this phenomenon. To answer the question "Why educate young people on malicious deepfakes?" we started by exploring this phenomenon and its dangers in the context of disinformation. First, this article exposes how deepfakes destabilize the perceptual foundations of human knowledge and discusses their threat to our societies. Secondly, it presents avenues considered and draws attention to the urgency of alerting young people to these dangers and involving them in counteracting this form of disinformation. Lastly, the article offers avenues for research in digital citizenship.

**Resumen**

La desinformación en línea no ha disminuido. El término info-calipsis fue inventado por Shick (2020) para describir la gravedad de la desinformación bajo la forma de hiper-argucias malintencionadas (*deepfakes*) que oscurecen lo que antes era considerado como auténtico. La presente reflexión circunscribe la complejidad del fenómeno de *deepfakes* con el fin de subrayar la urgente necesidad de realizar investigaciones que interpelen directamente a los actores y actrices del mundo de la educación. Desde la óptica de la educación a la ciudadanía, proponemos que el personal docente, los padres de familia y los actores y actrices del magisterio deben acompañar a los jóvenes en el desarrollo de un espíritu crítico, de la agentividad en el contexto numérico y del discernimiento frente a los hiper-argucias malintencionadas para que puedan protegerse y hacer frente a este fenómeno. Para dar respuesta a la cuestión «¿Por qué educar a los jóvenes a los hiper-argucias malintencionadas?» comenzamos por circunscribir dicho fenómeno y sus peligros en un contexto de desinformación. Por principio, este artículo expone la manera en que los *deepfakes* socavan las bases perceptivas ligadas al conocimiento humano y examina los peligros que corren nuestras sociedades. Después, presenta las opciones exploradas y señala a la atención lo urgente que es sensibilizar a los jóvenes a estas amenazas e implicarlos para hacer frente a esta forma de

desinformación. Finalmente, el artículo propone líneas de investigación por explorar en el terreno de la ciudadanía digital.

## INTRODUCTION

Avec les progrès et l'essor de la démocratisation de l'intelligence artificielle, nous assistons à l'émergence du phénomène des *deepfakes*<sup>1</sup> dont la technologie se limitait jusqu'à très récemment aux films issus des grandes productions cinématographiques (Glick, 2020; Paris et Donovan, 2019;). Grâce à l'apprentissage profond (*deep learning*), cette technologie permet de générer du contenu audiovisuel de manière ultraréaliste (Office québécois de la langue française, 2019a). Le degré de sophistication a atteint un point tel qu'il est désormais possible de reproduire numériquement un individu et lui faire dire ce que l'on souhaite ou le faire agir, sans son consentement, de façon à tromper la vigilance humaine (Westerlund, 2019).

L'objectif de cet article de réflexion est de circonscrire le phénomène des *deepfakes* et de tracer les contours des dangers qu'il soulève pour nos sociétés quand il est utilisé pour disséminer de la désinformation afin de mettre en relief le besoin urgent d'une éducation à la citoyenneté visant cette problématique. Nous présentons d'abord la méthodologie utilisée pour la revue de littérature. Nous explicitons ensuite la menace des *deepfakes* pour nos sociétés ainsi que les solutions envisagées, et nous expliquons l'urgence d'impliquer davantage l'éducation, notamment l'éducation à la littératie médiatique et l'éducation à la citoyenneté numérique, comme l'indiquent Hwang et al. (2021), pour développer l'esprit critique des jeunes, leur agentivité<sup>2</sup> en contexte numérique ainsi que leur discernement à l'égard des *deepfakes* afin qu'ils puissent se protéger et contrer ce phénomène. Ces apports convoquent des perspectives issues de différents domaines qui sont rassemblés à la fois pour mettre en lumière la complexité du phénomène et pour l'urgence d'agir qui dépasse le fait de simplement offrir de l'information pour sensibiliser les jeunes à la menace. Nous concluons l'article avec des pistes de recherche à explorer en citoyenneté numérique.

## MÉTHODOLOGIE

Afin de répondre à la question « Pourquoi éduquer les jeunes aux hypertrucages malveillants ? », nous avons commencé par circonscrire le phénomène des *deepfakes* et ses dangers dans un contexte de désinformation. Nous avons d'abord effectué une recherche dans les trois bases de données scientifiques ERIC, Education Source (EBSCO) et PsycINFO avec les mots-clés « *deepfake* » et « *désinformation* », puis avec « hypertrucage » et « désinformation », ce qui a généré des résultats très limités (Tableau 1). Ces trois bases de données couvrent le secteur de l'éducation et les ordres du milieu d'enseignement. Il convient de noter que nous n'avons appliqué aucun critère d'exclusion.

Tableau 1

### Résultats de la recherche dans trois bases de données

Bases de données	ERIC	Education Source (EBSCO)	PsycInfo
KW <i>deepfake</i> AND KW désinformation	0	0	1
KW ( <i>deepfake</i> ou <i>deep fake</i> or <i>face replacement</i> ) AND KW (désinformation ou <i>fake news</i> or <i>misinformation</i> )	0	2	5
KW <i>hypertrucage</i> AND KW désinformation	0	0	0
TI <i>hypertrucage</i> AND TI désinformation	0	0	0

Cela pourrait s'expliquer par le fait que le terme *deepfake* aurait été créé très récemment, en 2017, sur la plateforme Reddit, par un utilisateur qui se nommait « *deepfakes* » et qui utilisait la technologie de l'hypertrucage des domaines du génie logiciel et de la filmographie pour créer des vidéos pornographiques. Étant donné l'émergence du phénomène, nous avons amorcé une revue exploratoire du contenu partagé en ligne depuis 2017 produit par des spécialistes de ce phénomène, y compris des équipes de recherche.

## LA MENACE DE LA DÉSINFORMATION SOUS FORME DE DEEPPFAKE

La désinformation sous forme de *deepfake* risque fort de miner nos perceptions de ce qu'est un fait véridique et de menacer la démocratie dans un avenir rapproché (Citron, 2019; TAP Staff Blogger, 2019). Les experts prédisent que la technologie d'hypertrucage sera utilisée pour la cyberintimidation, la destruction de réputations, le chantage, la diffusion de discours haineux, l'incitation à la violence et la perturbation des processus démocratiques (Maras et Alexandrou, 2019), ainsi que pour la cybercriminalité et les fraudes (Strupp, 2019), pour ne citer que quelques exemples. La technologie évolue si rapidement vers l'impossibilité de distinguer le vrai du faux (Knight, 2019) qu'elle présente des risques alarmants pour la légitimité des informations en ligne (Agence France-Presse, 2019).

En éducation, il y a une urgence d'agir parce que les jeunes sont de plus en plus exposés à cette nouvelle forme de désinformation en ligne. Le Centre canadien d'éducation aux médias et de littératie numérique (HabiloMédias) a créé de nombreuses ressources pédagogiques pour authentifier l'information, y compris des textes, des jeux, des questionnaires, des ateliers, des plans de leçons ainsi que des affiches. Ces ressources visent à développer le sens critique des élèves basé sur l'authentification des faits vérifiables sur Internet. Avec le phénomène des *deepfakes* les établissements d'enseignement doivent d'emblée éduquer les jeunes aux mécanismes sous-jacents à la production des médias et les préparer à être critiques. Ils doivent également être capables d'agir dans l'environnement médiatique pour concevoir des programmes efficaces afin de préparer leurs élèves à faire face à d'ingénieux propagandistes numériques (Gordon, 2018; LSE Commission on Truth Trust and Technology, 2019; Valtonen et al. 2019). Selon Pérez-Escoda et al. (2021), les jeunes manquent de confiance à l'égard des politiciens, des médias et des journalistes. Un tel manque de confiance pourrait réduire leur participation à notre société démocratique ou ils

pourraient se fier à celle-ci aveuglément, mais, pire encore, des jeunes qui ne savent pas discerner ce qui est faux risquent de construire leurs connaissances d'un domaine ou leur opinion sur un sujet à partir de faussetés indétectables.

Pour savoir comment discerner ce qui est faux de ce qui est vrai, il faut d'abord faire la distinction entre la mésinformation, la malinformation et la désinformation. Alors que la mésinformation désigne une « information transmise au moyen des médias de masse ou des médias sociaux, qui est considérée comme véridique par l'émetteur, mais qui, en réalité, déforme les faits ou est erronée » (Office québécois de la langue française, 2019b) et que la malinformation est une information exacte, diffusée par des individus de manière trompeuse dans le but de nuire (Brin et al., 2021), la désinformation, quant à elle, est un contenu faux, manipulé ou fabriqué dans l'intention délibérée de nuire aux individus, aux organisations, aux sociétés, aux économies (Baines et Elliott, 2020). Quand la désinformation est associée à la technologie de l'hypertrucage, sa menace devient encore plus néfaste (Hwang et al., 2021). Dans les sections qui suivent, nous abordons différentes facettes du risque dont, notamment, la perte de confiance et la modification du rapport à l'information, l'incrédulité face à des vidéos authentiques et la possibilité de créer de faux souvenirs.

### **Le réalisme des *deepfakes* et la perte de confiance**

La diffusion sur Internet d'informations erronées ou déformées par certains médias contribue à l'émergence de théories du complot manipulant les populations à des fins politiques. Ce fait revêt une importance mondiale (Schroepfer, 2019). De nombreux experts nous mettent sans cesse en garde contre les *deepfakes* (Lee et al., 2021) qui ont vu le jour sous la forme de vidéos pornographiques superposant les visages de célébrités sur les corps de vedettes du porno (Maras et Alexandrou, 2019). Aujourd'hui, les hypertrucages ont le pouvoir d'être utilisés pour faire basculer des élections (Lalonde, 2019; O'Brien, 2019) et affaiblir la confiance des gens dans ces dernières (Charlet et Citron, 2019). L'introduction de l'intelligence artificielle n'a fait que faciliter la création de ces vidéos hyperréalistes (Maras et Alexandrou, 2019).

Les possibilités créatives de la technologie de l'hypertrucage semblent infinies; en voici deux exemples dans le monde du divertissement et de la culture dont les enjeux convoquent peu de débats. Après sa mort, l'acteur Paul Walker, de la franchise cinématographique *Rapides et dangereux*, est apparu dans le film *Dangereux 7*. Ce tour de force a été possible grâce au travail de Hao Li, pionnier de l'hypertrucage et professeur associé à l'Université de Californie du Sud (Howell O'Neill, 2019). Le Dalí Museum a fait revivre synthétiquement Salvador Dalí afin que les visiteurs de son exposition puissent être accueillis par l'artiste même, pour ensuite être invités à prendre un égoportrait en sa compagnie (The Dalí Museum, 2019).

Cependant, le potentiel d'armement des hypertrucages augmente de façon alarmante et incontrôlée. Auparavant, aucune technologie accessible au grand public n'aurait pu créer synthétiquement un média avec un réalisme comparable (Ajder et al., 2019). Récemment, une mère est allée aussi loin que de créer des *deepfakes* mettant en scène des meneuses de claqué rivales de sa fille afin de les discréditer (Slate, 2021). Ce changement de créneau a fait émerger de multiples dangers. Parmi les dangers à long terme, certains experts mentionnent le fait que les *deepfakes* accentuent la perte de confiance (Wahl-Jorgensen et Carlon 2021) déjà amorcée à l'égard des médias (Langlois et al., 2020), ce qui soulève le doute sur l'authenticité de l'information ainsi que la possibilité pour les jeunes d'être eux-mêmes et elles-mêmes la cible de *deepfake*, ce qui pourrait ruiner leur réputation ou les rendre victimes de cyberintimidation.

## L'incrédulité par défaut et le dividende du menteur

À l'heure actuelle, les possibilités de manipulations malveillantes de matériel audiovisuel font craindre l'émergence d'une attitude que l'on qualifie « d'incrédulité par défaut » (traduction libre de *disbelief by default*), qui consiste à remettre en question la véracité de vidéos authentiques afin de saper la crédibilité et de jeter le doute sur celles-ci (Weghe, 2019).

Selon son intention, le *deepfake* peut prendre pour cible des organisations et des entreprises, des personnalités publiques et politiques, des segments de la population et des personnes vulnérables. À long terme, il peut engendrer une perte de confiance généralisée touchant la société dans son ensemble, passant alors des microdangers (qui ne sont pas moins alarmants pour autant) aux macrodangers susceptibles de modifier radicalement notre rapport à la vérité.

Un monde dans lequel chaque image, vidéo ou enregistrement audio peut être manipulé engendrerait une attitude qui supposerait (par défaut) l'inauthenticité de tout matériel audiovisuel. Ainsi, « si tout peut être faux, alors rien ne doit être réel et n'importe qui peut commodément écarter les faits gênants comme étant des médias synthétiques » (Sonnemaker, 2021, traduction libre).

Dans le contexte où une telle attitude commune à l'égard de l'information existe, le discours démocratique perd son fondement dans des faits communément admis. Tout politicien, toute personne pourrait simplement nier les preuves d'actes répréhensibles et dire qu'il s'agit de manipulation (Goldzweig et Brady, 2020). Cette conséquence, inhérente au *deepfake*, est connue sous le nom « dividende du menteur » (traduction libre de *liar's dividend*) (Chesney et al., 2020). Terme formulé par les professeurs en droit Danielle Citron et Robert Chesney, le dividende du menteur est défini comme étant le risque que des menteurs puissent invoquer l'hypertrucage afin d'échapper aux responsabilités de leurs actes répréhensibles (TAP Staff Blogger, 2019).

Le rôle de l'éducation aux médias et à la citoyenneté numérique est crucial non seulement pour pouvoir reconnaître ou rejeter la désinformation, mais aussi pour pouvoir reconnaître la mésinformation et la malinformation, aussi en surface (vérifiables factuellement) ou en profondeur (intégrées faussement dans les médias, ce qui peut les rendre plus difficiles à vérifier factuellement parce qu'elles sont plus difficilement détectables) soient-elles, ce que nous expliquons plus en détail ci-dessous.

## Le pouvoir de la communication visuelle pour créer de faux souvenirs

Le pouvoir de la communication visuelle a déjà été démontré par de nombreux chercheurs (Block et Keller, 1997; Grabe et Bucy, 2009; Graber, 1990; Lazard et Atkinson, 2015; Messaris, 1997). Plusieurs chercheurs (Frenda et al., 2013; Prior, 2013; Stenberg, 2006; Sundar, 2008; Witten et Knudsen, 2005) soutiennent que l'information visuelle est intégrée plus efficacement que d'autres types de données sensorielles en raison de sa correspondance avec des représentations de la réalité. Ce constat trouve d'ailleurs écho chez Purnell (2020) qui, s'interrogeant sur la facilité avec laquelle les *deepfakes* peuvent tromper la vigilance du public, affirme que la réponse se trouve en partie dans ce qui est connu comme l'« oculocentrisme » et qui, comme son nom l'indique, est « un biais perceptif et épistémologique privilégiant la vision aux autres sens » (Oxford Reference, 2021, traduction libre). De plus, si l'image est familière, cette « familiarité suscite un effet de vérité – un sentiment de fluidité qui rend l'information plus facile à assimiler et donc plus crédible » (Vaccari et Chadwick, 2020, p. 2, traduction libre). L'image et la familiarité avec celle-ci jouent donc un rôle considérable dans la mémorisation d'une information.

Dans la foulée des avancées scientifiques permettant d'expliquer la relation qu'entretiennent image, information et mémoire, Liv et Greenbaum (2020) soulignent que « non seulement les *deepfakes* peuvent créer des représentations crédibles de la réalité, mais ils peuvent également être

utilisés pour créer de faux souvenirs » (p. 96, traduction libre). Selon ces auteurs, à la manière d'un cheval de Troie, ces faux souvenirs profitent de nos biais cognitifs qui tendent à favoriser les souvenirs évoquant notre propre conception du monde (*weltanschauung*). En ce sens, pour Purnell (2020), le principal danger ne vient pas directement de la manipulation visuelle présente dans les *deepfakes*, qualifiés de « désinformation sur les stéroïdes » (Chesney et Citron, 2018, traduction libre), mais plutôt dans la prédisposition des individus à croire instinctivement ce qu'ils voient. Lorsque les technologies qui permettent de médier ce rapport à l'image via les plateformes numériques accentuent le pouvoir de tromper la vision par l'hypertrucage, il est envisageable de voir ce danger s'accroître dans le temps.

Comme l'explique Moukheiber (2020), les réseaux sociaux font en sorte que nous ne regardons que les actualités qui confirment ou renforcent ce que l'on croit déjà. Le biais cognitif à la base de ce comportement, à savoir le biais de confirmation, joue un rôle clé dans la propagation de fausses nouvelles (Vicario et al., 2019). Le risque que les citoyennes et les citoyens construisent leurs connaissances et renforcent leurs convictions à partir d'informations en ligne délibérément fausses est alarmant.

Pour préparer les citoyennes et les citoyens du futur, il faut aller au-delà de la formation axée sur la maîtrise des contenus, du développement de la littératie numérique et de la formation à la pensée critique. Les *deepfakes* ouvrent une porte sur une formation plus englobante qui inclut ce qui a été fait dans le passé en passant par la législation et en ouvrant sur le travail des constructions psychologiques, aussi appelées perceptions ou représentations personnelles selon la tradition de recherche privilégiée. Dans la prochaine section, nous offrons quelques pistes à adopter pour sensibiliser les jeunes et les aider à contrer ces *deepfakes*.

## LES TENTATIVES DE SOLUTIONS POUR CONTRER LES HYPERTRUCAGES ET L'URGENCE DE SENSIBILISER LES JEUNES

Différents acteurs et actrices de la société se joignent dans un effort commun afin de faire face au phénomène des *deepfakes*. Jaiman (2020) explique que les solutions pour contrer les *deepfakes* peuvent se résumer à quatre grandes catégories d'interventions : 1) les réglementations et les actions législatives; 2) les politiques et la gouvernance des plateformes; 3) les outils technologiques; 4) l'éducation aux médias et à la citoyenneté numérique. En effet, pour protéger la vérité, garantir la liberté d'expression et faire face à l'hypertrucage malveillant, une approche hybride impliquant les différentes parties prenantes est primordiale (Goldzweig et Brady, 2020; Jaiman, 2020). Goldzweig et Brady (2020) précisent ceci :

Certains aspects de la solution sont liés aux actions globales nécessaires pour lutter contre la désinformation en tant que problème sociétal, tandis que d'autres sont plus spécifiques aux menaces liées à l'utilisation de l'intelligence artificielle dans la manipulation de contenus

p. 28, traduction libre

En 2019, soutenu par Microsoft, un consortium sur l'intelligence artificielle et un fonds de 10 millions de dollars (USD), Facebook a lancé le *Deepfake Detection Challenge* (Ferrer et al., 2020). Cette initiative visait à contrer la désinformation par la conception de nouveaux outils qui détectent les contenus médiatiques manipulés (Schroepfer, 2019). Récemment, Facebook a annoncé une solution prometteuse qui utilisera l'intelligence artificielle pour déconstruire la fabrication des

hypertrucages malveillants et repérer des imperfections infiltrées au produit (Le Figaro et Agence France-Presse, 2021; ZDNet, 2021). La Defense Advanced Research Projects Agency (2019), fer de lance du programme *Semantic Forensics* (*SemaFor*) qui évalue l'intégrité des médias visuels numériques grâce à l'utilisation de types d'algorithmes spécifiques, s'efforce de détecter, d'attribuer et de caractériser automatiquement les actifs médiatiques falsifiés. En parallèle, certains chercheurs discutent de l'utilisation de réseaux neuronaux pour détecter les incohérences entre les multiples images des *deepfakes* (Güera et Delp, 2018). D'autres proposent des algorithmes pour détecter des visages complètement générés en utilisant des données et des méthodes spécifiques pour construire des descripteurs pour la détection automatisée (Matern et al., 2019). Hasan et Salah (2019) proposent d'utiliser la technologie de la chaîne de blocs (*blockchain*) pour déterminer la source d'origine et l'historique du contenu numérique. Farid et Agarwal construisent un logiciel qui utilise les caractéristiques subtiles de la façon dont une personne parle pour distinguer cette personne de la fausse version d'elle-même (Manke, 2019).

Cependant, les solutions technologiques de détection des *deepfakes*, aussi performantes soient-elles, n'empêchent pas tous ceux-ci de circuler (Villasenor, 2019). En parallèle, les recours juridiques, quelle que soit leur efficacité, sont généralement appliqués après coup et ont une utilité limitée pour faire face aux dommages potentiels que peuvent causer les *deepfakes*, en particulier compte tenu des délais courts qui caractérisent la création, la distribution et la consommation des médias numériques (Villasenor, 2019). Par conséquent, il est de la plus haute urgence de trouver d'autres solutions ancrées dans l'application sociétale pour résister aux *deepfakes*. La convocation de l'éducation pour faire face aux menaces exposées ci-dessus devient inéluctable. Elle peut non seulement préparer les nouvelles générations de citoyens et de citoyennes à comprendre les dangers liés aux *deepfakes* et leur fonctionnement technique, mais elle peut aussi leur permettre de développer leur pensée critique et favoriser leur agentivité.

Le concept d'agentivité humaine émerge du travail du psychologue américain Albert Bandura (2006) pour expliquer notre capacité d'influencer le cours des événements ou l'environnement par l'entremise de nos actions. Toujours selon Bandura (2006), l'agentivité humaine s'exerce à partir de quatre fonctions : 1) l'intentionnalité (formuler une intention en planifiant comment nos actions se dérouleront); 2) la capacité de se projeter dans la temporalité par l'anticipation (planifier ses actions et savoir comment les exécuter pour atteindre ses objectifs); 3) l'autoréactivité (revenir sur le résultat de ses actions et corriger le cours, si nécessaire); 4) l'autoréflexivité (réfléchir sur ses actions). Dans un contexte numérique, l'agentivité pourrait impliquer de passer de l'observation passive des partages *des deepfakes* sur les réseaux sociaux à l'adoption active de stratégies en ligne pour dénoncer ces tromperies et les contrer par des faits (Naffi et al., 2021).

### **L'urgence de sensibiliser les jeunes aux dangers des hypertrucages malveillants**

À l'ère du numérique, chaque réponse à une question est censée être trouvée par une recherche Google, une vidéo YouTube ou une conférence TED, pour ne citer que quelques exemples. En fait, la majorité des jeunes Canadiens âgés de 15 à 24 ans utilise Internet quotidiennement, la plupart d'entre eux pour suivre les nouvelles et l'actualité grâce aux plateformes de médias sociaux comme Facebook, Twitter et Instagram (Mitchell et al., 2018; Statistique Canada, 2018). En 2020, 90 % des Canadiens âgés de 18 à 24 ans utilisait YouTube (Statista, 2021a). La demande croissante des consommateurs pour les vidéos en ligne a entraîné une augmentation de 40 % des vidéos téléchargées entre 2014 et 2020, et « à partir de février 2020, plus de 500 heures de vidéo étaient téléchargées sur YouTube chaque minute » (Ceci, 2021). Selon Statista, l'application TikTok<sup>3</sup> a réussi à attirer 65,9 millions d'utilisateurs en 2020 avec une projection de 73,7 millions pour l'année 2021, allant jusqu'à 88,7 millions en 2024, ne parlant là que des États-Unis (Statista, 2021b).

Dans le contexte de la littératie visuelle, Kearney (2020) explique que « le sens n'est pas situé uniquement dans l'image, mais est construit en partie par le spectateur en utilisant son propre système de signification précédemment intériorisé » (p. 171, traduction libre). C'est d'ailleurs la compréhension de ce phénomène que l'auteur nomme la « conscience sémiotique » (traduction libre de *semiotic awareness*). Dans la même ligne de pensée, Kelly (1955/1991) postule que les interprétations et les actions d'une personne fonctionnent de manière structurée et sont déterminées par ses prédictions de l'avenir. Ces prédictions sont générées par des systèmes de construits personnels que la personne développe à partir de ses propres expériences des événements. À travers le prisme de la psychologie des construits personnels (Kelly, 1955/1991), une théorie qui met l'humain au centre de ses expériences, les jeunes sont des sujets conscients qui, malgré la vulnérabilité de plusieurs d'entre eux, pourraient contrôler la façon dont ils interprètent les informations diffusées en ligne. Cette théorie a été utilisée avec succès dans plusieurs études sur des sujets critiques et délicats pour lesquels l'expérience subjective de la personne est remise en question, notamment la perception que les sociétés d'accueil se font des réfugiés syriens (Naffi et Davidson, 2016, 2017) ou dans un contexte de résolution de problème avec des communautés qui ont des divergences d'opinions (Davidson et al., 2016).

Le postulat de la théorie des construits personnels appliqué aux *deepfake*, pourrait se manifester de deux manières diamétralement opposées :

- a. Si les jeunes considèrent le phénomène du *deepfake* comme inoffensif et uniquement divertissant, ce groupe pourrait baisser la garde lorsqu'il est confronté à une fausse vidéo et interpréter les informations partagées comme étant fiables.
- b. En revanche, si les jeunes considèrent le phénomène comme une menace sérieuse, ce groupe risquerait de douter de chaque information partagée en ligne sur les plateformes de partage de vidéos et, par conséquent, de perdre des occasions d'apprentissage illimitées.

Moukheiber (2020) avance d'ailleurs que « l'humain a tendance à faire une confiance aveugle à sa perception, au point de la considérer comme partagée par tout le monde » (p. 20-21), que « nous ne voyons pas le monde tel qu'il est, mais tel que nous sommes » (p. 25) et que lors d'une ambiguïté, notre cerveau comble les vides pour retrouver la certitude. Devant le phénomène des *deepfakes*, il est crucial d'insister sur l'éducation aux médias pour aider les jeunes à développer leur esprit critique et pour utiliser une méthode favorisant aussi leur réflexivité quant à leurs propres interprétations des *deepfakes*, et, par la suite, leur agentivité pour les contrer.

### **Les *deepfakes* interpellent les acteurs et les actrices de l'éducation et d'autres disciplines**

Afin de développer la résilience face à la désinformation, notamment sous forme de *deepfake*, plusieurs initiatives de formation et de sensibilisation grand public ont vu le jour. Financé par Patrimoine Canada, l'organisme Journalists for Human Rights a lancé son programme de neuf mois intitulé *Fighting Disinformation Through Strengthened Media and Citizen Preparedness in Canada* (Journalists for Human Rights, 2019).

Visant la sensibilisation grand public, Radio-Canada a, de son côté, lancé l'émission *Les décrypteurs*. À travers l'usage d'un robot conversationnel, Alexis De Lancer, animateur de l'émission, initie les visiteurs aux *deepfakes* et à la façon de « combattre la désinformation » (Les décrypteurs, 2021). En parallèle, l'agence de presse Reuters (2021) offre une formation sur les contenus médiatiques manipulés. La formation est offerte en ligne en 16 langues. Le site Web du News Literacy Project (2021b) offre la plateforme d'apprentissage en ligne Checkology, laquelle partage des programmes et des ressources afin d'aider le personnel de l'éducation à enseigner la littératie médiatique aux élèves et aider les adultes à déterminer la crédibilité d'une information, ainsi que le *Sift*, un

bulletin d'information hebdomadaire qui fournit des exemples concrets de désinformation. Le Stanford History Education Group (s. d.) partage des ressources avec des groupes éducateurs et étudiants à travers leur site Web Civic Online Reasoning pour notamment évaluer la fiabilité des sources en ligne. CrashCourse, Mediawise et le Stanford History Education Group (2021) s'associent afin de présenter la série *Navigating Digital Information*, des vidéos éducatives pour développer des compétences pratiques dans l'évaluation des informations en ligne. Un autre exemple est celui du projet artistique immersif *Moon Disaster*, qui invite ses visiteurs à entrer dans une histoire alternative en leur demandant de réfléchir à la façon dont les nouvelles technologies peuvent déformer, réorienter et obscurcir la vérité qui nous entoure (Burgund et Panetta, 2019).

En particulier, la formation virtuelle *InfoZones*, créée par le News Literacy Project (2021a), discute de la technique du « zonage de l'information » (traduction libre). Cette technique consiste à amener le grand public à s'interroger sur l'objectif principal de l'information, soit divertir, vendre, persuader, provoquer, documenter ou informer, et ensuite à « zoner l'information » dans l'une des six catégories correspondantes : nouvelles, opinion, divertissement, publicité, propagande ou information brute. Cela permet de saisir l'information dans toute sa complexité et, par l'analyse de cas spécifiques, d'exposer les modalités d'un jugement critique. En ce sens, « zoner l'information » permet de faire ressortir l'objectif principal qui peut être dissimulé sous diverses apparences. Ces formations accessibles à tout public peuvent être employées dans les milieux éducatifs au service du personnel enseignant et des apprenants et apprenantes. Bien que ces formations contribuent à la compréhension de l'information, elles seules ne suffisent pas pour développer l'agentivité permettant de contrer la désinformation.

Dans les pays dont l'éducation est progressiste, l'éducation aux médias débute depuis le très jeune âge. Par exemple, la Finlande et le Danemark affirment avoir développé la résilience de leurs jeunes à l'égard de la désinformation grâce à l'éducation (Bellanger, 2021; Charlton, 2019). Leurs jeunes apprennent à analyser les médias d'une façon critique de même qu'à repérer les mensonges et la tromperie.

Le Canada fait figure de proue dans l'éducation aux médias avec de nombreuses initiatives menées par des organismes à but non lucratif ou des conseils scolaires d'un océan à l'autre, par exemple HabiloMédias (MediaSmarts en anglais), un organisme qui a « pour objectif de veiller à ce que les enfants, les adolescents et les adultes de confiance développent une pensée critique qui leur permet d'utiliser les médias à titre de citoyens numériques actifs et éclairés » (HabiloMédias, 2021). De nombreux organismes à but non lucratif tels que CIVIX (<https://educationauxmedias.ca>), Printemps numérique (<https://www.printempsnumerique.ca>), PAFEME, la plateforme d'actions et de formation en éducation aux médias pour enfants (<https://www.pafeme.ca>) (<https://www.pafeme.ca>), Le Curieux (<https://lecurieux.info>) (<https://lecurieux.info>) et AlphaNumérique (<https://alphanumerique.ca/espace-public>) (<https://alphanumerique.ca/espace-public/>), pour n'en nommer que quelques-uns, offrent des formations pour soutenir les jeunes dans la littératie numérique et l'éducation aux médias. Au Québec, RÉCIT (<https://recit.qc.ca>) (<https://recit.qc.ca/>) et CADRE21 (<https://www.cadre21.org>) offrent des formations au personnel enseignant et aux jeunes pour l'éducation aux médias et le développement de la compétence numérique. En Ontario, plusieurs conseils scolaires ont créé des ressources pédagogiques pour que les élèves soient informés, proactifs et responsables en ligne, et ils les ont mises à la disposition du public (Queen's Printer for Ontario, 2020). Par le truchement de ces multiples initiatives, le Canada possède un éventail de ressources pour favoriser le développement de la pensée critique des enfants, y compris l'authentification de l'information, la cyberintimidation, la cybersécurité et l'utilisation excessive d'Internet.

De fait, la pensée critique est une compétence principale du cadre de référence de la compétence numérique issu du plan d'action numérique du ministère de l'Éducation et de l'Enseignement supérieur du Québec (2019). La citoyenneté numérique, quant à elle, se trouve au cœur des

programmes d'études dans un grand éventail d'établissements d'enseignement du primaire au collégial. Elle comprend, entre autres, la littératie numérique, l'étiquette en ligne et la sécurité en ligne. Similairement, le *Cadre de référence de la compétence numérique* (ministère de l'Éducation et de l'Enseignement supérieur, 2019) définit la citoyenneté numérique comme ceci :

Fait de posséder des équipements et des compétences liés aux TIC et permettant de participer à la société numérique, par exemple d'accéder à des informations gouvernementales en ligne, d'utiliser des sites de réseaux sociaux et de faire usage d'un téléphone mobile

p. 28

Richard Culatta, directeur général de la Société internationale pour la technologie dans l'éducation, explique que la citoyenneté numérique consiste également à utiliser la technologie pour améliorer sa communauté, à participer à des débats respectueux, à façonner les politiques publiques et à reconnaître la validité des sources en ligne (Société internationale pour la technologie dans l'éducation, 2018).

Aujourd'hui, les appareils numériques des jeunes sont inondés d'informations. Qu'ils obtiennent leurs nouvelles en recherchant activement l'information ou simplement en tombant dessus lors d'interactions sociales en ligne avec leurs amis et leur famille, les jeunes doivent filtrer cette information et décider de sa crédibilité, surtout à l'ère des fausses nouvelles (*fake news*) et des *deepfakes*. Selon Turan (2021) la citoyenneté numérique pourrait avoir des effets en ce qui a trait à la protection des jeunes contre les dangers et les menaces des *deepfakes*. Pour que les jeunes puissent développer leur citoyenneté numérique et jouer un rôle actif dans la lutte contre la désinformation, notamment celle utilisant les médias synthétiques, ils doivent comprendre les préoccupations, les risques et la gravité de ce phénomène dans leur vie quotidienne et pour l'avenir de leur société. C'est pourquoi il est impératif de mener des recherches pour étudier l'expérience subjective des jeunes à l'égard des hypertrucages malveillants. Ces recherches permettraient de comprendre les mécanismes qui sont actifs lorsque les jeunes décodent les *deepfakes* et la façon dont ils peuvent s'y prendre pour freiner la propagation de ce type de média.

Parmi les utilisateurs actifs et les utilisatrices actives des médias sociaux âgés de 18 à 24 ans, nombreux sont ceux et celles qui font preuve d'esprit critique à l'égard du contenu des médias (Canadian Commission for UNESCO, 2018; Naffi, 2018). Bien que les jeunes soient qualifiés par certains chercheurs de « natifs du numérique » (Prensky, 2001) ou de « digital naïfs » (Frau-Meigs, 2015), leur choix d'information et de nouvelles qu'ils et elles lisent en ligne est loin d'être aléatoire (Thom, 2016). Un grand nombre d'entre eux reconnaissent les intentions et les algorithmes qui se cachent derrière les messages qui figurent sur leurs murs. Les jeunes ont soif d'avoir une voix influente qui pourrait perturber les discours sur les questions qui affectent leur vie (Naffi, 2018).

L'énorme succès de la grève mondiale pour le climat menée par Greta Thunberg en 2019 est un exemple de la résilience des jeunes face aux attaques contre la jeune militante et ses partisans sur les médias sociaux. Des mouvements tels que DREAMers, #BlackLivesMatter, Printemps Érable et #IfTheyGunnedMeDown, ainsi que les nombreuses veillées en réponse à l'attentat et à la fusillade à la mosquée de Québec en 2017 montrent que les jeunes sont capables de s'engager dans des actes politiques participatifs et d'utiliser les médias sociaux pour induire des changements politiques et sociaux. Cependant, si de nombreux jeunes adhèrent à la culture du partage et profitent des possibilités offertes par les médias sociaux pour s'engager dans la politique participative, des portions substantielles choisissent de rester des spectateurs passifs (Andreadakis, 2020), d'où l'urgence de convoquer le monde de l'éducation. Ce milieu peut réactualiser les programmes éducatifs qui abordent l'éducation aux médias et la citoyenneté numérique pour y inclure le développement de l'agentivité en contexte numérique pour contrer les *deepfakes* (Naffi et al., 2021).

## CONCLUSION

Actuellement, la désinformation est plus trompeuse et manipulatrice que jamais. Le potentiel d'armement des *deepfakes* augmente de manière alarmante, et les solutions technologiques de détection, quelle que soit leur qualité, s'avèrent encore insuffisantes pour prévenir et empêcher la propagation de cette nouvelle forme de fausses nouvelles. La préparation d'une génération de citoyennes et de citoyens éduqués face aux médias, qui fait preuve d'agentivité dans un contexte numérique, et qui est responsable, alerte et équipée pour contrer la désinformation est urgente pour assurer une démocratie pour l'avenir. Dans un contexte où les jeunes seront de plus en plus exposés à la désinformation sous forme de *deepfakes* le personnel de l'éducation doit s'efforcer de trouver des moyens pour les aider à développer un oeil aiguisé pour la détecter, à développer leur raisonnement civique en ligne, c'est-à-dire leur capacité à juger la crédibilité des informations qu'ils trouvent sur différentes plateformes (Stanford History Education Group, 2016), à développer leur pensée critique face à l'utilisation du numérique et à les aider à exercer une citoyenneté éthique dans les contextes numériques (ministère de l'Éducation et de l'Enseignement supérieur, 2019). Il importe aussi de les préparer à différentes formes d'engagement dans la vie civique et politique à l'ère numérique, à s'opposer à la désinformation et à déterminer les moyens de tirer parti de la puissance des médias sociaux pour avoir « davantage de contrôle, de voix et d'influence sur les questions qui comptent le plus dans leur vie » (Kahne et al., 2016, p. 38, traduction libre). Aujourd'hui plus que jamais, il est urgent d'explorer l'agentivité en contexte numérique des jeunes comme une stratégie pour contrebalancer les effets délétères des *deepfakes*. Plusieurs pistes de recherche peuvent être tracées, en commençant par examiner la façon dont les jeunes interprètent le phénomène de l'hypertrucage malveillant et reconnaissent la désinformation qui utilise les médias synthétiques. Il importe d'examiner comment les jeunes perçoivent leur rôle et leur responsabilité dans la lutte contre la propagation des *deepfakes*, ainsi que les stratégies numériques potentielles qu'ils et elles utiliseraient pour contrebalancer la désinformation sous forme d'hypertrucages malveillants. Les résultats de telles études guideront le personnel de l'éducation, les spécialistes de la conception de programmes éducatifs ainsi que les décideurs et décideuses dans leur quête pour promouvoir la pensée critique et l'agentivité en contexte numérique des jeunes, ainsi que celles des citoyens et des citoyennes dans le contexte d'une désinformation alimentée par l'utilisation non éthique des médias numériques, des données massives et de l'intelligence artificielle, et diffusée sur les médias sociaux.

## Notes

[1] Le terme *deepfake* utilisé dans cet article désigne l'utilisation malveillante et non-éthique de la technologie de l'hypertrucage.

[2] L'agentivité se manifeste lorsque des individus formulent des intentions et exécutent des actions volontaires qui vont au-delà des habitudes acceptées et des conditions normalement rencontrées. Ces actions vont au-delà des pratiques établies et sont éventuellement transformatives au sens où elles prennent de nouvelles formes d'action pour implanter une nouvelle culture (Barma et al., 2017).

[3] TikTok est une plateforme mobile et un phénomène viral et mondial qui connaît un grand succès auprès des jeunes en leur permettant de créer, de modifier et de partager des clips vidéo de courte durée agrémentés de filtres et accompagnés des dernières tendances musicales.

## Bibliographie

- Agence France-Presse. (2019, 6 septembre). Les géants de la technologie lancent le « deepfake challenge » pour contrer la désinformation. *Radio-Canada*. (<https://ici.radio-canada.ca/nouvelle/1288876/deepfake-challenge-lutte-facebook-desinformation>)
- Ajder, H., Patrini, G., Cavalli, F. et Cullen, L. (2019). *The state of deepfakes: Landscape, threats, and impact*. Deeptrace. ([https://regmedia.co.uk/2019/10/08/deepfake\\_report.pdf](https://regmedia.co.uk/2019/10/08/deepfake_report.pdf))
- Andreadakis, Z. (2020). Deep fakes and intelligence in the digital landscape: Preliminary systematic review findings. SSRN. ([https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3516344](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3516344)) doi : <https://doi.org/10.2139/ssrn.3516344>
- Baines, D.L. et Elliott, R.J.R. (2020, 2 février). Defining misinformation, disinformation and malinformation: An urgent need for clarity during the COVID-19 infodemic. *IDEAS*. (<https://ideas.repec.org/p/bir/birmec/20-06.html>)
- Bandura, A. (2006). Toward a psychology of human agency. *Perspectives on Psychological Science*, 1(2), 164-180. (<https://doi.org/10.1111/j.1745-6916.2006.00011.x>) doi : <https://doi.org/10.1111/j.1745-6916.2006.00011.x>
- Barma, S., Laferrière, T., Lemieux, B., Massé-Morneau, J. et Vincent, M.C. (2017). Early stages in building hybrid activity between school and work: The case of PénArt. *Journal of Education and Work*, 30(6), 669-687. doi : <https://doi.org/10.1080/13639080.2017.1294247>
- Bellanger, A. (2021, 17 février). Comment la Finlande a-t-elle vaincu les « fake news »? *France Inter*. (<https://www.franceinter.fr/emissions/les-histoires-du-monde/les-histoires-du-monde-17-fevrier-2021>)
- Block, L.G. et Keller, P.A. (1997). Effects of self-efficacy and vividness on the persuasiveness of health communications. *Journal of Consumer Psychology*, 6(1), 31-54. doi : [https://doi.org/10.1207/s15327663jcp0601\\_02](https://doi.org/10.1207/s15327663jcp0601_02)
- Brin, C., Durand, F., Gramaccia, J. et Thiboutot, C. (2021, avril). *Portrait d'une infodémie. Retour sur la première vague de COVID-19*. Observatoire international sur les impacts sociétaux de l'IA et du numérique. (<https://www.docdroid.com/NkXvSWt/portrait-dune-infodemie-retour-sur-la-premiere-vague-de-la-covid-19-pdf>)
- Burgund, H. et Panetta, F. (2019). *In Event of Moon Disaster - Home*. (<https://moondisaster.org/>) doi : <https://doi.org/10.1145/3414687.3434109>
- Canadian Commission for UNESCO. (2018, 5 novembre). *Fighting “fake news”: How youth are navigating modern misinformation and propaganda online*. (<https://en.ccunesco.ca/blog/2018/11/fighting-fake-news>)
- Ceci, L. (2021, 14 septembre). *Hours of video uploaded to YouTube every minute 2007–2020*. Statista. (<https://www.statista.com/statistics/259477/hours-of-video-uploaded-to-youtube-every-minute/>)
- Charlet, K. et Citron, D. (2019, 5 septembre). Campaigns must prepare for deepfakes: This is what their plan should look like. *Carnegie Endowment for International Peace*. (<https://carnegieendowment.org/2019/09/05/campaigns-must-prepare-for-deepfakes-this-is-what-their-plan-should-look-like-pub-79792>)
- Charlton, E. (2019, 21 mai). How Finland is fighting fake news in the classroom. *World Economic Forum*. (<https://www.weforum.org/agenda/2019/05/how-finland-is-fighting-fake-news-in-the-classroom/>)
- Chesney, R. et Citron, D. (2018, 16 octobre). Disinformation on steroids: The treat of deep fakes. *Council on Foreign Relations*. (<https://perma.cc/V9VF-J4DF>)
- Chesney, R., Citron, D. et Farid, H. (2020, 11 mai). All's clear for deepfakes: Think again. *Lawfare*. (<https://www.lawfareblog.com/all-clear-deepfakes-think-again>)

- Citron, D. (2019, 11 septembre). *How deepfakes undermine truth and threaten democracy* [vidéo]. Conférences TED. ([https://www.ted.com/talks/danielle\\_citron\\_how\\_deepfakes\\_undermine\\_truth\\_and\\_threaten\\_democracy](https://www.ted.com/talks/danielle_citron_how_deepfakes_undermine_truth_and_threaten_democracy))
- CrashCourse, Mediawise et Stanford History Education Group. (2021). *Navigating Digital Information*. YouTube. (<https://www.youtube.com/playlist?list=PL8dPuuaLjXtN07XYqqWSKpPrtNDiCHTZU>)
- Davidson, A.-L., Naffi, N. et Raby, C. (2017). A PCP approach to conflict resolution in learning communities. *Personal Construct Theory & Practice*,14, 61-72. (<http://www.pcp-net.org/journal/pctp17/davidson17.pdf>)
- Defense Advanced Research Projects Agency. (2019, 3 septembre). Uncovering the who, why, and how behind manipulated media. *DARPA*. (<https://www.darpa.mil/news-events/2019-09-03a>)
- Ferrer, C. C., Dolhansky, B., Pflaum, B., Bitton, J., Pan, J., & Lu, J. (2020, 12 juin). *Deepfake detection challenge results : An open initiative to advance AI*. FacebookAI. (<https://ai.facebook.com/datasets/dfdc/?fbclid=IwAR3wzAS4I75X4dC9GWJ3p1HwXxvd982rYF7N3nWbtuVZLZ2TMtywzCiSn74>)
- Frau-Meigs, D. (2015). *De l'identité à la présence en ligne : enjeux de translittératie et de socialisation* [vidéo]. Sam Network. (<https://www.sam-network.org/video/conference-prof-frau-meigs-universitat-sorbonne-paris-fr>)
- Frenda, S.J., Knowles, E.D., Saletan, W. et Loftus, E.F. (2013). False memories of fabricated political events. *Journal of Experimental Psychology*,49(2), 280-286.doi : <https://doi.org/10.1016/j.jesp.2012.10.013>
- Glick, J. (2020). *Deepfakes 101*. (<https://moondisaster.org/deepfakes-101>)
- Goldzweig, R. et Brady, M. (2020). *Deepfakes: How prepared are we?* Democracy Reporting International. (<https://democracy-reporting.org/uploads/publication/14512/document/2020-11-deepfakes-publication-no-2-6086d42418da4.pdf>)
- Gordon, M. (2018). Lying in politics: Fake news, alternative facts, and the challenges for deliberative civics education. *Educational Theory*,68(1), 49-64.doi : <https://doi.org/10.1111/edth.12288>
- Grabe, M.E. et Bucy, E.P. (2009). *Image bite politics: News and the visual framing of elections*. Oxford University Press.doi : <https://doi.org/10.1093/acprof:oso/9780195372076.001.0001>
- Graber, D.A. (1990). Seeing is remembering: How visuals contribute to learning from television news. *Journal of Communication*,40(3), 134-156.doi : <https://doi.org/10.1111/j.1460-2466.1990.tb02275.x>
- Güera, D. et Delp, E.J. (2018). *Deepfake video detection using recurrent neural networks*. IEEE International Conference on Advanced Video and Signal-based Surveillance (AVSS). <https://engineering.purdue.edu/~dgueraco/content/deepfake.pdf> (<https://engineering.purdue.edu/%7Edgueraco/content/deepfake.pdf>) doi : <https://doi.org/10.1109/AVSS.2018.8639163>
- HabiloMédias. (2021). *Notre mission et notre philosophie*. (<https://habilomedias.ca/qui-nous-sommes/notre-mission-et-notre-philosophie>)
- Hasan, H.R. et Salah, K. (2019). Combating deepfake videos using blockchain and smart contracts. *IEEE Access*,7, 41596-41606. (<https://doi.org/10.1109/ACCESS.2019.2905689>) doi : <https://doi.org/10.1109/ACCESS.2019.2905689>
- Howell O'Neill, P. (2019, 19 septembre). The world's top deepfake artist: 'Wow, this is developing more rapidly than I thought.'. *MIT Technology Review*. (<https://www.technologyreview.com/s/614343/the-worlds-top-deepfake-artist-wow-this-is-developing-more-rapidly-than-i-thought/>)
- Hwang, Y., Ryu, J.Y. et Jeong, S.-H. (2021). Effects of disinformation using deepfake: The protective effect of media literacy education. *Cyberpsychology, Behavior, and Social Networking*, 24(3), 188-193. (<https://doi.org/10.1089/cyber.2020.0174>) doi : <https://doi.org/10.1089/cyber.2020.0174>
- Jaiman, A. (2020, 27 août). Technical countermeasures to deepfakes. *Towards Data Science*. (<https://towardsdatascience.com/technical-countermeasures-to-deepfakes-564429a642d3>)

- Journalists for Human Rights. (2019, 27 septembre). Launching JHR's program on "Fighting Disinformation through Strengthened Media and Citizen Preparedness in Canada." *Newswire*. (<https://www.newswire.ca/news-releases/launching-jhr-s-program-on-fighting-disinformation-through-strengthened-media-and-citizen-preparedness-in-canada-899686785.html>)
- Kahne, J., Hodgin, E. et Eidman-Aadahl, E. (2016). Redesigning civic education for the digital age: Participatory politics and the pursuit of democratic engagement. *Theory & Research in Social Education*, 44(1), 1-35. ([https://ypp.dmlcentral.net/sites/default/files/publications/Redesigning%20Civic%20Ed\\_Kahne%2C%20Hodgin%2C%20Eidmain-Aadahl.pdf](https://ypp.dmlcentral.net/sites/default/files/publications/Redesigning%20Civic%20Ed_Kahne%2C%20Hodgin%2C%20Eidmain-Aadahl.pdf)) doi : <https://doi.org/10.1080/00933104.2015.1132646>
- Kearney, R. (2020). *Postdigital visual literacy: A semiotic perspective*. Auckland University of Technology.
- Kelly, G. (1955/1991). *The psychology of personal constructs*. W. W. Norton/Routledge.
- Knight, W. (2019, 19 août). The world's top deepfake artist is wrestling with the monster he created. *MIT Technology Review*. (<https://www.technologyreview.com/s/614083/the-worlds-top-deepfake-artist-is-wrestling-with-the-monster-he-created/>)
- Lalonde, D. (2019, 11 septembre). The election's on: Now Canadians should watch out for dumbfakes and deepfakes. *The Conversation*. (<https://theconversation.com/the-elections-on-now-canadians-should-watch-out-for-dumbfakes-and-deepfakes-122927>)
- Langlois, S., Proulx, S. et Sauvageau, F. (2020). *La confiance envers les médias d'information et les médias sociaux au Québec*. Centre d'études sur les médias. <https://www.cem.ulaval.ca/wp-content/uploads/2020/02/cem-confiance-langlois-proulx-sauvageau.pdf> (<https://www.cem.ulaval.ca/wp-content/uploads/2020/02/cem-confiance-langlois-proulx-sauvageau.pdf>)
- Lazard, A. et Atkinson L. (2015). Putting environmental infographics center stage: The role of visuals at the elaboration likelihood model's critical point of persuasion. *Science Communication*, 37(1), 6-33. doi : <https://doi.org/10.1177/1075547014555997>
- Le Figaro et Agence France-Presse. (2021, 17 juin). Facebook dit progresser dans la détection des images manipulées ou « deepfake ». *Le Figaro*. (<https://www.lefigaro.fr/secteur/high-tech/facebook-dit-progresser-dans-la-detection-des-images-manipulees-ou-deepfake-20210617>)
- Lee, Y., Huang, K.T., Blom, R., Schriener, R. et Ciccarelli, C.A. (2021). To believe or not to believe: Framing analysis of content and audience response of top 10 deepfake videos on YouTube. *Cyberpsychology, Behavior and Social Networking*, 24(3), 153-158. (<https://doi.org/10.1089/cyber.2020.0176>) doi : <https://doi.org/10.1089/cyber.2020.0176>
- Les décrypteurs. (2021). Comment combattre la désinformation. *Radio-Canada*. (<https://ici.radio-canada.ca/info/dcrypteurs/robot-conversationnel-combattre-desinformation/initiation-hypertrucages/#top>)
- Liv, N. et Greenbaum, D. (2020). Deep fake and memory malleability: False memories in the service of fake news. *AJOB Neuroscience*, 11(2), 96-104. doi : <https://doi.org/10.1080/21507740.2020.1740351>
- LSE Commission on Truth Trust and Technology. (2019). *Tackling the information crisis: A policy framework for media system resilience*. (<http://www.lse.ac.uk/media-and-communications/assets/documents/research/T3-Report-Tackling-the-Information-Crisis-v6.pdf>)
- Manke, K. (2019, 18 juin). Researchers use facial quirks to unmask 'deepfakes.' *Berkley News*. (<https://news.berkeley.edu/2019/06/18/researchers-use-facial-quirks-to-unmask-deepfakes/>)
- Maras, M.H. et Alexandrou, A. (2019). Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos. *The International Journal of Evidence & Proof*, 23(3), 255-262. doi : <https://doi.org/10.1177/1365712718807226>
- Matern, F., Riess, C. et Stamminger, M. (2019). *Exploiting visual artifacts to expose deepfakes and face manipulations*. 2019 IEEE Winter Applications of Computer Vision Workshops (WACVW). (<https://ieeexplore.ieee.org/abstract/document/8638330/metrics#metrics>) doi : <https://doi.org/10.1109/WACVW.2019.00020>
- Messaris, P. (1997). *Visual persuasion: The role of images in advertising*. Sage.

- Ministère de l'Éducation et de l'Enseignement supérieur. (2019). *Cadre de référence de la compétence numérique*. ([http://www.education.gouv.qc.ca/fileadmin/site\\_web/documents/ministere/Cadre-reference-competece-num.pdf](http://www.education.gouv.qc.ca/fileadmin/site_web/documents/ministere/Cadre-reference-competece-num.pdf))
- Mitchell, A., Simmons, K., Matsa, K.E. et Silver, L. (2018, 9 janvier). Young people much more likely than older to get news daily via social media. *Pew Research Center*. ([https://www.pewresearch.org/global/2018/01/11/people-in-poorer-countries-just-as-likely-to-use-social-media-for-news-as-those-in-wealthier-countries/pg\\_2018-01-11\\_global-media-habits\\_3-02/](https://www.pewresearch.org/global/2018/01/11/people-in-poorer-countries-just-as-likely-to-use-social-media-for-news-as-those-in-wealthier-countries/pg_2018-01-11_global-media-habits_3-02/))
- Moukheiber, A. (2020). *Votre cerveau vous joue des tours*. J'ai Lu.
- Naffi, N. (2018). *Learning about oneself: An essential process to confront social media propaganda against the resettlement of syrian refugees* [Thèse de doctorat, Université Concordia]. Spectrum. (<https://spectrum.library.concordia.ca/983399/>)
- Naffi, N. et Davidson, A.-L. (2016). Examining the integration and inclusion of Syrian refugees through the lens of personal construct psychology. *Personal Construct Theory & Practice*, 13, 200-209. (<https://tinyurl.com/5vr6wdyt>)
- Naffi, N. et Davidson, A.-L. (2017). Engaging host society youth in exploring how they construe the influence of social media on the resettlement of Syrian refugees. *Personal Construct Theory & Practice*, 14, 116-128. (<https://tinyurl.com/7zrapavz>)
- Naffi, N., Davidson, A.-L. et Berger, F. (2021, 20 avril). Infocalypse : la propagation des hypertrucages menace la société. *The Conversation*. (<https://theconversation.com/infocalypse-la-propagation-des-hypertrucages-menace-la-societe-158335>)
- News Literacy Project. (2021a). *InfoZones*. (<https://get.checkology.org/lesson/infozones/>)
- News Literacy Project. (2021b). *What is Checkology?* (<https://get.checkology.org/what-is-checkology/>)
- O'Brien, M. (2019, 12 juin). Hany Farid on the threat of deepfakes during the 2020 election. *I School Berkley*. (<https://www.ischool.berkeley.edu/news/2019/hany-farid-threat-deepfakes-during-2020-election>)
- Office québécois de la langue française. (2019a). Hypertrucage. Dans *Grand dictionnaire terminologique*. ([http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id\\_Fiche=26552557](http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=26552557))
- Office québécois de la langue française. (2019b). Mésinformation. Dans *Grand dictionnaire terminologique*. ([http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id\\_Fiche=26556735](http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=26556735))
- Oxford Reference. (2021). *Ocularcentrism*. (<https://www.oxfordreference.com/view/10.1093/oi/authority.20110803100245338>) .
- Paris, B. et Donovan, J. (2019). Deepfakes and cheap fakes: The manipulation of audio and visual evidence. *Data & Society*. ([https://datasociety.net/wp-content/uploads/2019/09/DS\\_Deepfakes\\_Cheap\\_FakesFinal-1-1.pdf](https://datasociety.net/wp-content/uploads/2019/09/DS_Deepfakes_Cheap_FakesFinal-1-1.pdf))
- Pérez-Escoda, A., Pedrero-Esteban, L.M., Rubio-Romero, J. et Jiménez-Narros, C. (2021). Fake news reaching young people on social networks: Distrust challenging media literacy. *Publications*, 9(2), 24. (<https://doi.org/10.3390/publications9020024>) doi : <https://doi.org/10.3390/publications9020024>
- Prensky, M. (2001). Digital natives, digital immigrants. Part I. *On the Horizon*, 9(5), 1-6. doi : <https://doi.org/10.1108/10748120110424816>
- Prior, M. (2013). Visual political knowledge: A different road to competence? *Journal of Politics*, 76(1), 41-57. doi : <https://doi.org/10.1017/S0022381613001096>
- Purnell, C. (2020). Solutions deepfakes. *Psychology Today*, 30-31.
- Queen's Printer for Ontario. (2020). *Digital literacy*. (<https://www.dcp.edu.gov.on.ca/en/program-planning/transferable-skills/digital-literacy>)
- Reuters. (2021). *Identifier et lutter contre les contenus médiatiques manipulés*. (<https://www.reuters.com/manipulatedmedia/fr/>)
- Schick, N. (2020). *Deepfakes: The coming infocalypse*. Grand Central Publishing.

- Schroepfer, M. (2019, 5 septembre). *Creating a data set and a challenge for deepfakes* [page Facebook]. (<https://ai.facebook.com/blog/deepfake-detection-challenge/>)
- Slate. (2021, 14 mars). Une mère utilise le deepfake pour discréditer des pom-pom girls rivales de sa fille. *Slate*. (<http://www.slate.fr/story/205547/une-mere-utilise-le-deepfake-pour-discrediter-des-pom-pom-girls-rivales-de-sa-fille>)
- Société internationale pour la technologie dans l'éducation. (2018, 11 octobre). *Rethinking digital citizenship*. Youtube. (<https://www.youtube.com/watch?v=iwKTYHBG5kk>)
- Sonnemaker, T. (2021, 13 avril). "Liar's dividend": The more we learn about deepfakes, the more dangerous they become. *Business Insider*. (<https://www.businessinsider.com/deepfakes-liars-dividend-explained-future-misinformation-social-media-fake-news-2021-4>)
- Stanford History Education Group. (s. d.). *Curriculum*. Civic Online Reasoning. (<https://cor.stanford.edu/curriculum/>)
- Stanford History Education Group. (2016). *Evaluating information: The cornerstone of civic online reasoning*. (<https://sheg.stanford.edu/upload/V3LessonPlans/Executive%20Summary%2011.21.16.pdf>)
- Statista. (2021a, 5 mai). *YouTube user reach among internet users in Canada as of April 2020, by age group*. (<https://www.statista.com/statistics/484416/canada-youtube-penetration-by-age/>)
- Statista. (2021b, 19 avril). *Number of TikTok users in the United States from 2019 to 2024*. (<https://www.statista.com/statistics/1100836/number-of-us-tiktok-users/>)
- Statistique Canada. (2018, 7 février). *A portrait of Canadian youth*. (<https://www150.statcan.gc.ca/n1/pub/11-631-x/11-631-x2018001-eng.htm>)
- Stenberg, S. (2006). Conceptual and perceptual factors in the picture superiority effect. *European Journal of Cognitive Psychology*, 18(6), 813-847. doi : <https://doi.org/10.1080/09541440500412361>
- Strupp, C. (2019, 30 août). Fraudsters used AI to mimic CEO's voice in unusual cybercrime case. *The Wall Street Journal*. (<https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402>)
- Sundar, S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. Dans M. Metzger et A. Flanagin (dir.), *Digital media, youth, and credibility* (p. 73-100). MIT Press.
- TAP Staff Blogger. (2019, 12 septembre). Danielle Citron discusses how deepfakes undermine truth and threaten democracy. *Technology, Academics, Policy (TAP)*. (<https://www.techpolicy.com/Blog/September-2019/Danielle-Citron-Discusses-How-Deepfakes-Undermine.aspx>)
- The Dalí Museum. (2019, 8 mai). *Behind the scenes: Dali Lives* [vidéo]. YouTube. (<https://www.youtube.com/watch?v=BIDaxl4xqJ4>)
- Thom, J. (2016). *Believing the news: Exploring how young Canadians make decisions about their news consumption* [Thèse non publiée, University of Western Ontario]. Electronic Thesis and Dissertation Repository. (<http://ir.lib.uwo.ca/cgi/viewcontent.cgi?article=5898&context=etd>)
- Turan, S.G. (2021). Deepfake and digital citizenship: A long-term protection method for children and youth. Dans R.J. Blankenship (dir.), *Deep fakes, fake news, and misinformation in online teaching and learning technologies* (p. 124-142). IGI Global. (<https://doi.org/10.4018/978-1-7998-6474-5.ch006>) doi : <https://doi.org/10.4018/978-1-7998-6474-5.ch006>
- Vaccari, C. et 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), 1-13. doi : <https://doi.org/10.1177/2056305120903408>
- Valtonen, T., Tedre, M., Mäkitalo, K. et Vartiainen, H. (2019). Media literacy education in the age of machine learning. *Journal of Media Literacy Education*, 11(2), 20-36. doi : <https://doi.org/10.23860/JMLE-2019-11-2-2>

- Vicario, M.D., Quattrociochi, W., Scala, A. et Zollo, F. (2019). Polarization and fake news: Early warning of potential misinformation targets. *ACM Journals*, 13(2), 1-22. doi : <https://doi.org/10.1145/3316809>
- Villasenor, J. (2019, 14 février). Artificial intelligence, deepfakes, and the uncertain future of truth. *Brookings*. (<https://www.brookings.edu/blog/techtank/2019/02/14/artificial-intelligence-deepfakes-and-the-uncertain-future-of-truth/>)
- Wahl-Jorgensen, K. et Carlson, M. (2021). Conjecturing fearful futures: Journalistic discourses on deepfakes. *Journalism Practice*, 15(6), 803-820. (<https://www.tandfonline.com/doi/full/10.1080/17512786.2021.1908838>) doi : <https://doi.org/10.1080/17512786.2021.1908838>
- Weghe, TV. (2019, 29 mai). Six lessons form my deepfakes research at Stanford. *Medium*. (<https://medium.com/jsk-class-of-2019/six-lessons-from-my-deepfake-research-at-stanford-1666594a8e50>)
- Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review*, 9(11), 39-52. doi : <https://doi.org/10.22215/timreview/1282>
- Witten, I.B. et Knudsen, E.I. (2005). Why seeing is believing: Merging auditory and visual worlds. *Neuron*, 48(3), 489-496. doi : <https://doi.org/10.1016/j.neuron.2005.10.020>
- ZDNet. (2021, 18 juin). Vidéo : dans la lutte contre les deepfakes, Facebook a une solution à base d'IA. *ZDNet*. (<https://www.zdnet.fr/actualites/video-lutte-contre-les-deepfake-facebook-a-une-solution-a-base-d-ia-39924757.htm>)