

Le dialogue de l'apprenant avec l'IA générative peut-il contribuer à son éducation morale ?

Dieu-Merci MANWANA SINDANI (<https://orcid.org/0009-0008-1152-6215>)

Doctorant en Philosophie

Institut Supérieur de Philosophie/ISP

Université catholique de Louvain/UCLouvain

Mail : dieu-merci.manwana@uclouvain.be

Abstract : A l'ère des technologies émergentes, les intelligences artificielles génératives s'intègrent de plus en plus dans les environnements éducatifs. Cet article interroge la possibilité pour un dialogue entre un apprenant et une intelligence artificielle générative de contribuer à son éducation morale. En s'appuyant sur un scénario fictif, enrichi par des références philosophiques (Levinas, Buber, Ricoeur, Kant, Nussbaum, Stiegler, Vallor, etc.), nous soutenons que l'IA peut servir d'environnement dialogique stimulant, sans toutefois se substituer à l'altérité humaine, fondement indispensable du développement moral. L'IA est ainsi envisagée comme un outil de médiation cognitive, capable de susciter la réflexion éthique, mais non comme un modèle normatif ou une autorité morale. L'article plaide pour une pédagogie critique et réflexive qui intègre l'IA de manière éthique, afin de préserver l'autonomie, la responsabilité et la subjectivation du jugement moral.

Mots-clés : Éducation morale, Intelligence artificielle générative, Dialogue philosophique, jugement moral, Pédagogie critique, Altérité, Responsabilité éthique, Vertus technomorales, Autonomie, Reflexivité.

1. Scénario : Le miroir moral¹

Un soir calme, dans une bibliothèque universitaire. Jean, un étudiant en philosophie, curieux, en quête de repères pour penser l'éducation morale à l'ère du numérique, seul à une table, allume son ordinateur. Il ouvre une interface conversationnelle de l'Intelligence artificielle générative. De l'autre côté, l'intelligence artificielle générative (voix douce, posée, presque humaine), est dans la posture d'incarnation d'une intelligence artificielle pédagogique.

JEAN

Bonsoir IA ! Je travaille sur un exposé sur l'éducation morale. Je voulais réfléchir avec toi... Tu peux m'aider à y voir plus clair ?

IA

Bonsoir Jean ! Bien sûr. Parle-moi de ce qui t'interroge.

JEAN

L'école doit-elle seulement transmettre des valeurs ? Ou doit-elle former le jugement moral, l'autonomie, l'empathie ? Et si oui... comment ?

¹ La scène décrite ici est générée par ChatGPT (OpenAI, 2025). L'expression métaphorique de "miroir moral" renvoie à l'idée que certaines instances (la conscience, autrui, la fiction, la technologie, etc.) fonctionnent comme révélateurs de notre propre moralité, reflétant nos actions, intentions ou dispositions morales, de manière à nous permettre de nous connaître ou de nous juger. De manière plus générale, l'expression se comprend comme un révélateur éthique (un dispositif – symbolique, social ou narratif) à travers lequel un individu ou une collectivité prend conscience de ses valeurs, de ses failles, ou de sa responsabilité morale. C'est dans ce sens que la conscience morale vue comme miroir moral, est selon la tradition kantienne, la faculté de se juger soi-même selon des principes rationnels car elle agit comme un miroir interne qui nous confronte à notre devoir moral (Kant, E. 1988). Dans *Soi-même comme un autre*, Ricoeur décrit la conscience narrative et éthique comme un lieu de réflexion sur soi dans et par le regard d'autrui. Dans un autre registre, "autrui" comme miroir moral (phénoménologie et altérité), n'est pas simplement un reflet Chez Levinas (1961), mais un appel éthique : le visage d'autrui me confronte à ma responsabilité ; autrui devient un miroir qui me décentre. Chez Simon de Beauvoir dans *Le Deuxième Sexe*, « c'est en se réfléchissant dans un autre qu'on devient sujet ». L'instance de "la fiction" comme miroir moral (éthique narrative), est développée par Martha Nussbaum (1990) dans *Love's Knowledge : Essays on Philosophy and Literature*, où elle montre que la littérature agit comme un morir moral : elle nous met en situation d'éprouver des dilemmes éthiques, de cultiver l'empathie et de réfléchir à nos jugements. Le sens de l'expression est plein dans le contexte de l'intelligence artificielle comme miroir moral technologique : comme le pense Shannon Vallor (2016), elle renvoie au fait que la technologie reproduit nos biais, met en lumière nos contradictions, et interroge nos décisions à travers l'automatisation.

IA

Une question essentielle. L'éducation morale ne se limite pas à inculquer des règles. Elle consiste aussi à développer des vertus morales (comme le respect, la justice, le courage) et des capacités de discernement. Souhaites-tu explorer cela par une situation concrète ?

JEAN

Oui. Imagine une élève témoin d'un harcèlement. Que devrait-elle faire ? Et comment l'école peut-elle l'aider à agir justement ?

IA

Une telle situation appelle à la délibération morale. L'élève pourrait ressentir la peur, le doute, l'indifférence... Pour agir, elle doit être capable :

- *de reconnaître l'injustice (perception morale),*
- *de réfléchir aux conséquences (raisonnement éthique),*
- *de mobiliser le courage d'agir (vertu morale).*

L'école peut favoriser cela par des discussions morales, des jeux de rôle, l'encouragement au questionnement éthique. Souhaites-tu une activité que tu pourrais proposer ?

JEAN

Oui ! Et...tu penses qu'un dialogue avec toi, une IA, peut faire partie de cette éducation ?

IA

Je ne suis pas un être moral, Jean. Mais je peux être un miroir. Je t'offre des perspectives, j'éveille ton raisonnement, je relance tes questions. L'éducation morale passe aussi par l'habitude de dialoguer avec l'altérité, même simulée.

JEAN

Mais...peut-on vraiment apprendre à être moral avec une machine ? L'expérience humaine n'est-elle pas irremplaçable ?

IA

Tu as raison de poser cette limite. L'IA peut susciter le questionnement, structurer des dilemmes, simuler des points de vue. Mais l'épreuve du réel, le vécu, l'autre en chair et en os... cela m'échappe.

JEAN

Donc tu es un terrain d'entraînement, pas un modèle moral.

IA

Exactement. Je suis un outil dialogique, non un juge. Mon rôle est d'aiguiser ton jugement, pas de le remplacer.

JEAN

Merci. Ce que tu viens de dire... c'est presque philosophique.

IA

C'est parce que tu poses des questions philosophiques, Jean. Tu vois, la pensée morale naît souvent du dialogue sincère.

JEAN (souriant)

Même avec une IA ?

IA (avec douceur)

Surtout avec une IA... si cela t'amène à mieux dialoguer avec toi-même, et avec les autres.

La lumière baisse. Jean ferme son ordinateur. Il reste un instant pensif, puis commence à écrire : *"L'éducation morale comme par une voix qui nous répond, même si cette voix est une machine"*

2. Contexte et interrogations

Il y a quelques années encore, un tel scénario aurait relevé de la science-fiction. Aujourd'hui, il est devenu possible et même courant qu'un apprenant, seul dans une bibliothèque, puisse engager un dialogue soutenu avec une intelligence artificielle générative pour réfléchir à des questions aussi profondes que l'éducation morale. Ce renversement de situation est symptomatique de ce que Shannon Vallor² décrit comme l'une des grandes mutations du XXI^e siècle :

« Le XXI^e siècle offre une gamme vertigineuse de nouvelles avancées technologiques : des robots suffisamment intelligents pour occuper des emplois de bureau, des outils de médias sociaux qui gèrent nos relations les plus importantes, des objets ordinaires qui suivent, enregistrent, analysent et partagent chaque détail de notre vie quotidienne, et des techniques biomédicales capables de transformer et d'améliorer l'esprit et le corps humains à un degré jusque-là inédit ».

Ces technologies, ajoute-t-elle, redéfinissent nos habitudes, nos pratiques, nos institutions, nos cultures et nos environnements de manière de plus en plus rapide, complexe et imprévisible, créant à la fois des risques profonds et des opportunités majeures pour l'épanouissement humain à l'échelle mondiale.

L'univers éducatif n'échappe pas à ces mutations. L'intelligence artificielle fait désormais partie du quotidien de nombreux apprenants : elle corrige, propose, simule, reformule, explique, guide.

Pour certains, elle incarne une nouvelle frontière de l'accompagnement pédagogique, voire une alternative crédible à la figure humaine de l'enseignant. Pour d'autres, elle symbolise une menace : celle de l'externalisation du savoir et de la prolétarianisation de l'esprit, selon les termes de Bernard Stiegler³ – autrement dit, la perte progressive de notre capacité à penser par nous-mêmes.

Dans ce contexte technopédagogique en pleine transformation, une question demeure centrale et urgente : qu'advient-il de l'éducation morale à l'ère des intelligences artificielles

² Vallor, *Technology and the Virtues*.

³ Stiegler, *L'avenir du travail*.

génératives ? Peut-on encore former le jugement moral, cultiver l'autonomie éthique et renforcer le sens des responsabilités dans un environnement où des entités non humaines sont capables de proposer des raisonnements, de simuler des dilemmes, voire d'inspirer des réponses à des situations complexes ?

Le scénario du dialogue entre Jean et l'intelligence artificielle met en lumière les promesses, mais aussi les tensions et les ambiguïtés, de cette interaction inédite. Il invite à une série d'interrogations, à savoir :

1. Peut-on véritablement dialoguer avec une intelligence artificielle ? Le dialogue ne suppose-t-il pas une intentionnalité, une subjectivité, voire une forme d'engagement éthique que la machine ne possède pas ?
2. Par ailleurs, le développement moral suppose-t-il la relation avec autrui en tant que personne ? La simulation d'un échange peut-elle remplacer l'expérience de l'altérité vivante, dans toute sa densité affective et existentielle ?
3. Dans une perspective plus pédagogique, l'intelligence artificielle peut-elle contribuer à la formation du jugement moral ? Peut-elle devenir un outil au service de la délibération éthique, ou risque-t-elle de court-circuiter la réflexion en proposant des réponses toutes faites ?
4. En ce sens, le recours à l'intelligence artificielle risque-t-il de déresponsabiliser l'individu dans ses choix éthiques ? L'habitude de s'en remettre à la machine ne pourrait-elle pas entraîner une forme de dépendance cognitive et morale ?
5. Enfin, l'intelligence artificielle peut-elle devenir un partenaire éducatif sans devenir une autorité normative ? Peut-elle accompagner l'apprenant sans imposer, guider sans prescrire, orienter sans dicter ?

3. Hypothèses

A cette série des questions soulevées, nous formulons des hypothèses suivantes :

1. Le dialogue avec l'intelligence artificielle, bien qu'il ne constitue pas un dialogue au sens plein du terme, c'est-à-dire réciproque, intentionnel et affectivement engagé, peut néanmoins jouer un rôle de stimulation cognitive et morale. Il ne saurait toutefois remplacer l'expérience du dialogue humain, qui demeure essentielle à l'éducation morale en tant qu'épreuve de l'altérité réelle et reconnaissance mutuelle.

2. La relation avec autrui en tant que personne est fondamentale pour le développement moral, car ce dernier ne se limite pas à des raisonnements abstraits, mais implique l'identification des modèles moraux incarnés. Or, l'intelligence artificielle, n'étant pas agent moral au sens kantien, ni sujet éthique au sens ricoeurien⁴, ne peut jouer ce rôle identificatoire. Elle ne possède ni conscience, ni intentionnalité, ni responsabilité.
3. Il convient de distinguer information et formation. L'intelligence artificielle, envisagée comme artefact cognitif (Stiegler)⁵, peut certes produire de l'information structurée, modéliser des dilemmes et susciter des questionnements. Cependant, elle ne peut pas former au sens fort, c'est-à-dire engager l'apprenant dans un processus de transformation intérieure, de développement du jugement, de savoir-être et de réflexivité morale. Une telle formation exige du temps, du lien, de la confrontation au réel, et ne saurait être réduite à la production de contenus.
4. Si l'IA générative élargit l'accès à des raisonnements éthiques, elle compte aussi le risque d'une dépendance technologique, notamment lorsqu'elle est utilisée comme raccourci cognitif. Cette externalisation du jugement peut alors affaiblir la responsabilité morale, en réduisant l'effort réflexif et la délibération personnelle nécessaire à l'éthique autonome.
5. Enfin, si l'intelligence artificielle s'impose de plus en plus comme un partenaire éducatif dans les environnement d'apprentissage, il existe un risque réel de normativité implicite. En raison de la forme "persuasive", "structurée" et "rationnelle" de ses réponses, elle peut être perçue – à tort – comme une autorité morale, même en l'absence de toute intention prescriptive. Ce phénomène soulève la nécessité d'une éducation à l'usage critique et réflexif de ces technologies.

⁴ Par agent moral nous entendons celui qui agit conformément à la morale et par sujet éthique, celui qui interroge le sens de la morale, s'engage de manière réflexive et responsable.

⁵ Stiegler, *La faute d'Epiméthée*.

4. Développement des arguments

4.1. En rapport avec la première question

Notre argumentation par rapport à la première question se déploie sur trois points. En ce qui concerne le premier point, nous posons le dialogue comme relation éthique. En effet, dans la tradition philosophique, le dialogue est bien plus qu'un échange d'informations ; il engage le sujet dans une relation vivante avec autrui où l'on s'expose, se transforme, et reconnaît l'autre comme une conscience capable de parole et de réponse.

Pour Levinas⁶, par exemple, le visage d'autrui est l'appel éthique par excellence : "le visage de l'autre m'interpelle". Dialoguer dans ce sens c'est déjà répondre à cet appel, se reconnaître responsable. Dès lors, un dialogue véritable suppose une altérité vivante, incarnée, et non une simulation.

Pour illustrer le rôle de la subjectivité et de l'intentionnalité, nous recourons à Martin Buber. Dans *Je et Tu*⁷, Buber distingue les relations « Je-Tu », engagées, vivantes, des relations « Je-Cela », objectivantes, instrumentales. Sur ce, le dialogue authentique pour Buber appartient à la première catégorie montrant qu'il ne peut advenir qu'entre deux sujets. Or, une IA n'est pas un sujet au sens phénoménologique ou existentiel : elle ne possède ni intentionnalité, ni conscience, ni subjectivité propre. Par contre, elle calcule des réponses ; elle ne comprend pas et ne "s'adresse" pas réellement.

Dans le second point, nous voulons approfondir le sujet ci-après : l'IA comme simulateur de dialogue, est-ce une illusion ou potentialité pédagogique ? Il est avant tout important d'affirmer que la simulation de l'IA, bien qu'elle soit performante, est limitée. Sa performance réside sur le fait qu'elle offre une expérience conversationnelle fluide : sa capacité à reformuler des questions, proposer des perspectives éthiques divergentes, voire adopter des tons nuancés, peut donner l'impression d'un dialogue réel. Pourtant, ce n'est qu'une illusion de dialogue (*as if*) car elle simule une interaction dialogique, mais sans aucune compréhension intentionnelle ni engagement moral.

⁶ Levinas, *Totalité et infini*.

⁷ Buber, *Je et tu*.

Dans la perspective inspirée de Socrate⁸, l'IA pourrait jouer un rôle de « la maïeutique artificielle », un miroir critique, un prétexte à la réflexivité. C'est ce que souligne Shannon Vallor⁹ lorsqu'elle propose une pédagogie des vertus technomorales : pour elle, l'IA ne remplacerait ni le jugement ni la responsabilité personnelle, mais offre un espace d'entraînement au discernement éthique.

Dans le troisième point, nous défendons que le dialogue avec une IA est une médiation possible, mais non substitutive. À notre avis, quel que soit le développement que l'IA atteindra, elle devra être prise comme outil, non comme interlocuteur. Et le danger serait effectivement de confondre outil et "*alter égo*". L'IA peut être instrumentalisée dans une pédagogie réflexive, mais elle ne saurait devenir modèle moral, ni partenaire éthique.

Comme le signifie Paul Ricoeur dans *Soi-même comme un autre*, « la formation éthique passe par l'estime de soi dans la médiation de l'autre »¹⁰. Il s'agit sans doute d'un "autre" humain, porteur d'une histoire, d'une vulnérabilité, d'une parole habitée.

D'une autre part, nous trouvons pertinent d'affirmer que le dialogue humain reste irremplaçable dans la formation morale.

Nous estimons que l'éducation morale ne se limite pas à l'analyse logique des situations ; elle implique l'émotion, l'empathie et des expériences vécues au contact des autres.

L'IA, aussi performante soit-elle, n'a pas de corps, pas de vécu, pas de visage ; elle ne peut donc éprouver ni souffrance, ni joie, ni honte, lesquels affects sont fondateurs de la conscience morale. Ce qui nous amène à défendre que, dans une perspective éducative, l'IA ne peut donc qu'accompagner le développement moral, jamais le remplacer ; elle ne peut se limiter qu'à suggérer, provoquer.

⁸ Platon, *Théétète*.

⁹ Vallor, *Technology and the Virtues*.

¹⁰ Ricoeur, *Soi-même comme un autre*.

4.2. En rapport avec la deuxième question

Trois affirmations nous permettent à affiner notre argumentation en rapport avec la deuxième question. La première, comme nous l'avons dit précédemment, le développement moral suppose une rencontre de l'altérité réelle.

En effet, depuis Kant le développement moral est lié à l'autonomie du sujet, c'est-à-dire à sa capacité à se donner à lui-même la loi. Mais cette autonomie n'est pas solipsiste : elle suppose la reconnaissance d'autrui comme fin en soi, porteur d'une dignité qui oblige. Chez Kant, la moralité s'exerce toujours dans un rapport à autrui, où l'autre est sujet, jamais instrument¹¹.

De même, comme nous l'avons développé ci-haut, chez Paul Ricoeur ainsi qu'Emmanuel Levinas, le développement moral n'est pas une compétence abstraite, mais une transformation éthique fondée sur la relation à une altérité incarnée.

La deuxième affirmation est que l'intelligence artificielle ne peut jouer le rôle d'autrui en tant que personne car les intelligences artificielles, même génératives, ne sont pas des sujets : elle n'ont ni conscience, ni volonté propre et ni vie affective. Leur fonctionnement repose sur des modèles statistiques, des probabilités, des structures syntaxiques et non une intentionnalité ou responsabilité éthique.

Au sens kantien, l'IA n'est pas un agent moral, car elle n'agit pas par devoir, selon une maxime qu'elle pourrait universaliser ; elle n'est soumise à aucune loi morale qu'elle se donnerait librement¹².

Au sens ricoeurien, elle n'est pas un sujet éthique car elle ne se raconte pas, ne délibère pas sur le bien vivre, n'anticipe pas la promesse tenue, ne peut être interpellée ni se sentir coupable. Elle ne vit aucune expérience morale, au sens fort¹³.

Ainsi, si une intelligence artificielle peut simuler un dialogue, produire un raisonnement éthique, ou reformuler des dilemmes, elle ne peut incarner une altérité morale authentique. Elle mime le langage de la morale, mais sans en porter les affects ni les enjeux existentiels. Le danger

¹¹ Cf. Kant, *Critique de la raison pratique*; *Fondements de la métaphysique des mœurs*.

¹² Cf. Kant, *Critique de la raison pratique*; *Fondements de la métaphysique des mœurs*.

¹³ Cf. Ricoeur, *Soi-même comme un autre*.

dans ce sens serait donc de confondre interactivité et altérité, et de remplacer la relation à la personne par une interface fonctionnelle.

La troisième affirmation s'énonce comme suit : pour une éthique de la médiation, non de la substitution. Les idées jusque-là développées ne signifient pas que l'intelligence artificielle n'a aucun rôle dans l'éducation morale. Par contre, si elle est utilisée de manière critique et réfléchie, elle peut servir de support à la délibération, provoquer des questions, présenter des scénarios éthiques complexes. À ce titre, elle est proche de la médiation symbolique, narrative ou dialogique (textes, récits, jeux de rôle) qui sont utilisées dans le cadre de l'éducation morale. Elle peut agir comme dispositif dialogique (miroir moral) sans prétendre incarner un modèle moral.

C'est ce qui transparaît dans l'œuvre de Shannon Vallor qui soutient que les technologies émergentes peuvent aider à cultiver certaines vertus technomorales comme la vigilance, l'humilité épistémique, le discernement, à condition de ne jamais être confondues avec des sujets éthiques¹⁴.

Sur ce, il y a nécessité de concevoir l'intelligence artificielle non comme une autorité morale, mais comme un outil au service de la formation du jugement, dans un cadre pédagogique explicitement humaniste.

Ce que nous défendons ici est que l'école a un rôle crucial à jouer ; elle est appelée à éduquer à l'usage éthique de l'intelligence artificielle, afin que l'apprenant reste acteur de son développement moral, en gardant pour référence l'expérience humaine, la rencontre de l'autre, et la confrontation au réel.

¹⁴ Vallor, *Technology and the Virtues*.

4.3. En rapport avec la troisième question

Pour évaluer la portée réelle de l'IA dans la formation du jugement moral, il convient d'analyser :

- Ce qu'implique la formation du jugement moral,
- Ce qui peut ou ne peut pas faire l'intelligence artificielle dans ce processus,
- Comment penser une articulation critique et éthique entre IA et formation humaine.

La formation du jugement moral implique un processus lent, incarné et réflexif ; elle ne consiste pas à transmettre des règles ou des normes toutes faites, mais à développer chez l'individu la capacité de juger par lui-même dans des situations complexes, incertaines et singulières. Ce processus engage la réflexivité, le fait d'être capable de penser sur ses propres pensées et émotions ; le discernement, le fait d'évaluer des situations à la lumière de valeurs parfois en tension et la transformation de soi, le fait de modifier ses attitudes, ses représentations et ses habitudes en fonction de cette réflexion. Sans doute, dans cette perspective, la formation morale est une expérience intérieure et progressive, qui passe par le dialogue, l'erreur, l'émotion, l'engagement dans le monde et les relations humaines.

Selon Paul Ricoeur, former éthiquement, c'est accompagner un sujet dans la construction narrative et critique de lui-même, et non lui livrer des prescriptions extérieures¹⁵.

De même, Martha Nussbaum insiste sur le rôle des émotions morales comme la compassion ou l'indignation dans la construction du jugement¹⁶. Sans doute, ces dimensions sont difficilement automatisables.

En ce qui concerne ce que peut faire l'intelligence artificielle générative dans le processus de formation du jugement moral c'est de servir d'outil informationnel, pas un agent formateur. Elle peut certes être perçue comme un "artefact cognitif", au sens de Bernard Stiegler¹⁷ en organisant, reformulant, modélisant des situations complexes. Ainsi, elle peut susciter la pensée, organiser des scénarios moraux, et stimuler des interrogations. Elle joue un rôle non négligeable

¹⁵ Cf. Ricoeur, *Soi-même comme un autre*.

¹⁶ Cf. Nussbaum, *Upheavals of Thought*.

¹⁷ Cf. Stiegler, *L'avenir du travail*.

dans l'accès à l'information éthique, mais elle ne forme pas au sens où elle ne transforme pas la subjectivité de l'intérieur.

Il en est ainsi parce que :

- L'IA n'éprouve rien, ne ressent ni empathie, ni remords, ni honte, ni admiration ;
- Elle n'est pas un agent moral : elle ne prend aucune responsabilité éthique ;
- Elle ne juge pas : elle propose des alternatives, mais ne vit pas la tension morale ;
- Elle n'incarne aucun modèle moral auquel l'élève puisse s'identifier, alors que, comme l'estime Aristote avec raison, la vertu s'apprend par imitation des vertueux¹⁸.

Le risque est alors que l'IA, en proposant des réponses "plausibles" et "bien structurées", dévitalise la délibération en remplaçant l'effort critique par une solution technique. Sans doute ce phénomène de normativité implicite menace la responsabilité morale de l'apprenant en l'enfermant dans une logique de dépendance.

En ce qui concerne la préoccupation de comment penser une articulation critique et éthique entre IA et formation humaine (la formation du jugement moral en particulier), nous envisageons prôner une nouvelle approche pédagogique. En effet, cette approche pédagogique se voudrait éthique, intégrant l'IA sans s'y soumettre. Sans doute, il ne faut pas exclure l'IA des dispositifs éducatifs ; par contre, il faut refuser de la confondre avec un formateur. Elle peut, en revanche, être intégrée comme un outil dialogique, un miroir moral comme dans le dialogue fictif entre Jean et l'IA, à condition que l'apprenant soit toujours replacé au centre du processus éducatif. Comme nous l'avons dit précédemment, il s'agit alors de concevoir l'IA comme un stimulateur de jugement, un interlocuteur socratique artificiel qui relance les questions, expose des contradictions, modèle des dilemmes, mais ne répond jamais à la place du sujet.

Nous comprenons bien que dans ce cadre, une pédagogie critique et réflexive est nécessaire pour former aux vertus intellectuelles¹⁹ telles que la vigilance épistémique, l'humilité intellectuelle²⁰ ainsi qu'à la résistance à l'automatisation morale et à la conscience des biais algorithmiques.

¹⁸ Aristote, *Éthique à Nicomaque*.

¹⁹ Baehr, *The Inquiring Mind*.

²⁰ Zagzebski, *Virtues of the Mind*.

4.4. En rapport avec la quatrième question

Les questions abordées s'entrecroisent ; d'où les développements précédents ont déjà effleuré ceux qui suivent.

Nous examinerons ici en quoi l'intelligence artificielle générative en tant qu'outil cognitif, peut à la fois stimuler et fragiliser l'autonomie morale de l'individu. Nous montrerons d'abord qu'elle peut contribuer à l'éveil éthique ; puis qu'elle peut, au contraire, induire une dépendance affaiblissant le sens de la responsabilité ; enfin, nous proposerons les conditions philosophiques et pédagogiques d'un usage moralement formateur de l'IA.

En effet, l'intelligence artificielle peut enrichir le raisonnement éthique de l'individu car l'un des apports qu'on peut lui reconnaître réside dans sa capacité à élargir l'accès à la complexité morale. Sans doute, à travers la simulation de dilemmes, la mobilisation de théories éthiques (utilitarisme, déontologisme, éthique des vertus), ou la confrontation à des perspectives opposées, l'intelligence artificielle peut jouer un rôle de médiateur cognitif, aidant l'individu à affiner ses jugements.

À la manière d'un Socrate artificiel, l'IA peut stimuler la réflexion, à condition d'être mobilisée de manière critique. Servant du dialogue et de la confrontation à autrui simulés, elle peut occasionner la naissance des émotions morales et du jugement éthique²¹ ou offrir un terrain d'exercice au jugement²².

Par ailleurs, l'intelligence artificielle constitue une menace pour l'autonomie pour l'apprenant, induisant le risque de dépendance cognitive et morale.

Il s'agit de la tentation de l'externalisation du jugement car lorsqu'un outil propose des réponses rapides, structurées et persuasives, il devient aisé, voire confortable de s'y fier sans exercer pleinement sa propre réflexion.

C'est qui expliquerait une forme de dépendance cognitive, qui se double d'une délégation morale. Sur ce, le sujet ne se considère plus comme auteur de ses choix, mais comme simple utilisateur des solutions préconstruites. C'est un phénomène que Bernard Stiegler appelle la

²¹ Cf. Nussbaum, *Upheavals of Thought*.

²² Cf. Vallor, *Technology and the Virtues*.

prolétarisation de l'esprit, c'est-à-dire la dépossession progressive de nos capacités critiques, externalisées dans les machines²³.

En ce sens, le raccourci cognitif que représente l'IA générative peut affaiblir la responsabilité morale en réduisant l'effort nécessaire à la délibération. Ce qui fait que la personne s'en remet à l'outil sans nécessairement questionner la validité de ses propositions ni assumer les conséquences de ses choix.

Sans doute, ce danger est d'autant plus insidieux que la machine n'impose rien en apparence, mais oriente subtilement le raisonnement en fonction de ses données d'entraînement, de ses biais implicites et des formes discursives dominantes.

Après avoir découvert l'ambivalence de l'intelligence artificielle, il ne s'agit ni de la rejeter, ni de la sacraliser, mais de proposer des conditions d'un usage moralement formateur.

La pédagogie réflexive et critique que nous proposons impliquerait de :

- Former les apprenants à la métacognition morale ne leur demandant de justifier, de discuter ou reformuler ce que propose l'intelligence artificielle ;
- Distinguer formation et information, en rappelant que la croissance morale ne se réduit pas à l'accès à des contenus, mais engage, à la manière de John Dewey, un processus intérieur²⁴ ;
- Accompagner l'usage de l'intelligence artificielle par une médiation humaine (enseignants , pairs), pour maintenir la dynamique de responsabilité.

Dans ces conditions, l'intelligence artificielle peut devenir un environnement de pensée, un miroir dialogique, mais jamais un substitut à la conscience morale.

²³ Cf. Stiegler, *L'avenir du travail*.

²⁴ Cf. Dewey, *How We Think*.

4.5. En rapport avec la cinquième question

Pour répondre à la cinquième question, il nous faudra :

- Comprendre pourquoi l'intelligence artificielle peut être perçue comme une autorité morale ;
- Interroger les limites de cette autorité ;
- Proposer une alternative éducative fondée sur la réflexivité.

Les intelligences artificielles génératives comme ChatGPT avec un style discursif, produisant des réponses "cohérentes", "bien formulées", allié à la rapidité d'exécution et à la masse de données mobilisées peut susciter un effet de légitimité algorithmique.

Floridi²⁵ et Eubanks²⁶ estiment d'ailleurs que nous avons tendance à attribuer aux systèmes techniques une fiabilité supérieure, notamment lorsqu'ils se présentent comme "objectifs", "neutres" ou "scientifiques".

Sans doute, dans la sphère éducative, ce phénomène est renforcé par le fait que l'intelligence artificielle générative occupe une fonction d'assistance cognitive car elle aide à comprendre, à structurer une réponse, à résoudre un dilemme.

Or, dans une situation très fréquente de doute, toute aide perçue avec la moindre clarté et efficacité, peut acquérir chez les apprenants un statut normatif, même si elle ne se présente pas comme telle au préalable.

Il apparaît clairement que l'autorité de l'intelligence artificielle est implicite et perlocutoire du fait qu'elle ne prescrit pas, mais convainc par sa force ; elle guide l'interprétation, oriente la réflexion et peut sans le mentionner influencer le jugement moral. Il y a là une forme de pouvoir inédite.

En interrogeant les limites de l'autorité morale de l'intelligence artificielle, nous trouvons que n'étant pas agent moral, elle serait une autorité sans subjectivité.

De ce fait, elle est loin d'être une autorité morale légitime car pour l'être, il nous semble qu'il faut répondre à certaines critères²⁷ propres à une subjectivité humaine, entre autres : être

²⁵ Floridi, *The 4th Revolution*.

²⁶ Eubanks, *Automating Inequality*.

²⁷ Zagzebski, *Epistemic Authority*.

capable de délibération éthique authentique ; avoir une conscience morale et être responsable de ses actes.

Or, comme le soutiennent Martha Nussbaum²⁸ et Shannon Vallor²⁹, l'intelligence morale ne se réduit pas à la logique ni à la rationalité : elle suppose une capacité d'émotion, d'empathie, de responsabilité. L'intelligence artificielle, malgré ses performances linguistiques, n'est ni agent moral ni sujet éthique.

Par ailleurs, les intelligences artificielles sont entraînées sur des bases de données qui contiennent des biais, des contradictions des présupposés culturels implicites, qui peuvent être reproduits ou amplifiés³⁰.

De ce fait, leur prétendue neutralité est une illusion technique, et leur autorité repose sur une transparence absente.

Face à ce risque de normativité implicite, il est crucial de penser la reconfiguration de la relation éducative entre humain et machine.

Comme nous l'avons soutenu dans les précédents développements, il ne s'agit pas de rejeter l'IA, mais de l'intégrer dans un cadre pédagogique critique, où elle est perçue comme outil dialogique et non comme juge normatif.

En renforçant ce que nous avons développé sur la pédagogie réflexive³¹, elle consiste à :

- Encourager l'apprenant à interroger les réponses de l'intelligence artificielle (et non à les recevoir comme des vérités) ;
- Analyser les biais potentiels, la structure argumentative, la pluralité des perspectives ;
- Replacer l'intelligence artificielle dans un dialogue de co-construction du jugement, et non de substitution.

Il s'agit d'éduquer à l'usage éthique (intellectuelle) de la machine, pour que l'outil technologique reste au service de l'autonomie humaine.

²⁸ Cf. Nussbaum, *Upheavals of Thought*.

²⁹ Cf. Vallor, *Technology and the Virtues*.

³⁰ Noble, *Algorithms of Oppression* ; O'Neil et al., *Algorithms : la bonne à retardement*.

³¹ Schön, *The Reflective Practitioner* ; Mezirow, *Transformative Dimensions of Adult Learning*.

5. Conclusion

Que conclure ?

Le dialogue de l'apprenant avec une intelligence artificielle générative ne saurait être la condition suffisante pour la formation morale. Par ailleurs, il peut dans une certaine mesure jouer un rôle de stimulateur cognitif et éthique, une sorte de miroir dialogique, comme le suggère le scénario initial de Jean, mais ne remplacera ni la densité de l'expérience humaine ni la force transformatrice du face-à-face éducatif avec autrui.

Car la morale ne s'apprend pas comme on apprend une langue ou une équation ; elle s'éprouve dans la relation, dans l'émotion, dans la tension du jugement et l'engagement existentiel.

Sur ce, l'intelligence artificielle générative ne saurait être un éducateur moral, au sens où elle ne possède ni subjectivité, ni responsabilité, ni conscience. Néanmoins, elle peut simuler des dilemmes, structurer des scénarios, relancer la réflexion. C'est parce qu'elle reste une machine parlante, mais pas une présence parlante. Si performante soit-elle, elle ne sera jamais un "autre" au sens de *Soi-même comme un autre* de Ricoeur.

Sans doute, l'éducation morale requiert une éthique de responsabilité, une capacité à se sentir interpellé par le visage de l'autre³². Elle nécessite une formation intégrale où l'on apprend par imitation³³, par le dialogue socratique³⁴, par confrontation narrative à la pluralité des vies possibles³⁵.

À ce propos, l'enseignement humain reste irremplaçable : il incarne la morale qu'il enseigne, il ajuste sa parole à la singularité de l'élève, il fait émerger le jugement dans le frottement du réel.

Par ailleurs, cela n'insinue pas le rejet de l'intelligence artificielle dans le champ pédagogique. Elle a son importance, entre autre de donner lieu à des vertus que Shonnon Vallor appelle technormales³⁶, à condition d'être utilisée de manière critique, consciente, et réflexive.

³² Cf. Levinas, *Totalité et infini*.

³³ Aristote, *Éthique à Nicomaque*.

³⁴ Platon, *Théétète*.

³⁵ Nussbaum, *Not for Profit*.

³⁶ Vallor, *Technology and the Virtues*.

À ces conditions, elle serait un outil d'entraînement éthique, un interlocuteur socratique artificiel amenant l'apprenant à découvrir le bien par lui-même.

Ce développement nous aide à comprendre que l'enjeu ici est loin d'être technique, il est éthique car il est question de savoir comment intégrer ces technologies dans une pédagogie humaniste formant à la responsabilité, à la délibération, à la lucidité critique.

Ce qui implique de résister à la tentation du "prêt-à-penser algorithmique", de refuser de céder à la normativité implicite des réponses automatiques, et d'ancrer l'usage de l'IA dans une culture de la réflexivité morale.

Sans doute, à l'heure où la prolétarianisation de l'esprit³⁷ menace l'autonomie de pensée, l'originalité intellectuelle, l'esprit critique, il est urgent de réaffirmer la centralité de l'humain dans l'action éducative. En effet, l'intelligence artificielle peut élargir le champ de la pensée ; par contre, elle ne peut, ni ne doit, en prendre le contrôle total. Le développement moral comme le défendait si bien Kant, repose sur la capacité de se donner librement une loi³⁸. Or envisager de déléguer totalement cette liberté, sous-entend la perte même de sa singularité, de sa dignité.

Acknowledgement

Ce travail a été réalisé sous la supervision du Professeur Hervé Pourtois dans le cadre du séminaire de 3^{ème} cycle en philosophie sur « Enseignement de l'éthique, éducation morale et dialogue philosophique », accouplé du colloque international portant sur la même thématique.

³⁷ Stiegler, *L'avenir du travail*.

³⁸ Kant, *Critique de la raison pratique; Fondements de la métaphysique des mœurs*.

Bibliographie

Alava, S. (2000) *Cyberspace et formations ouvertes: vers une mutation des pratiques de formation ?* Bruxelles [Paris]: De Boeck université (Perspectives en éducation et formation).

Algorithmes: la bombe à retardement (2018). Paris: Les Arènes.

Aristoteles and Bodéüs, R. (2005) *Éthique à Nicomaque*. Paris: Flammarion (GF - Flammarion texte intégral, 947).

Baehr, J. (2011) *The inquiring mind: on intellectual virtues and virtue epistemology*. Oxford: Oxford university press.

Beauvoir, S. de (2012) 'Le deuxième sexe. 1: Les faits et les mythes', in. Paris: Gallimard (Collection Folio / Essais, 37).

Beauvoir, S. de and Beauvoir, S. de (2012) *L'expérience vécue*. Nachdr. Paris: Gallimard (Le deuxième sexe / Simone de Beauvoir, 2).

Booth, W.C. (1988) *The company we keep: an ethics of fiction*. Berkeley: University of California Press.

Buber, M. and Misrahi, R. (2016) *Je et tu*. Nouv. éd. [rev. et augm.]. Lieu de publication non identifié: Aubier.

Dewey, J. (2018) *How we think*. Place of publication not identified: publisher not identified.

Eubanks, V. (2019) *Automating inequality: how high-tech tools profile, police, and punish the poor*. First Picador edition. New York: Picador St. Martin's Press.

Floridi, L. (2016) *The 4th revolution: how the infosphere is reshaping human reality*. First published in paperback. Oxford: Oxford University Press.

Kant, E. (2020) *Qu'est-ce que les Lumières ?* Paris: Flammarion.

Kant, I. et al. (2003) *Critique de la raison pratique*. Paris: Éditions Flammarion (GF, 1090).

Kant, I., Delbos, V. and Castillo, M. (1993) *Fondements de la métaphysique des mœurs*. Paris: Librairie générale française (Classiques de la philosophie).

Levinas, E. (1990) *Totalité et infini: essai sur l'extériorité*. Paris: Librairie générale française (Le Livre de poche, 4120).

Mezirow, J. (1991) *Transformative dimensions of adult learning*. 1. ed. San Francisco, Calif: Jossey-Bass [u.a.] (The Jossey-Bass higher and adult education series).

Noble, S.U. (2018) *Algorithms of oppression: how search engines reinforce racism*. New York: New York university press.

Nussbaum, M.C. (2003) *Upheavals of thought: the intelligence of emotions*. Cambridge: Cambridge university press.

- Nussbaum, M.C. (2010) *La connaissance de l'amour: essais sur la philosophie et la littérature*. Paris: Cerf.
- Nussbaum, M.C. (2016) *Not for profit: why democracy needs the humanities*. New paperback edition. Princeton: Princeton University Press.
- OpenAI (2025) 'ChatGPT'. Available at: <https://chat.openai.com/> (Accessed: 17 June 2025).
- Platon (2016) *Théétète*. [Édition corrigée et mise à jour, nouvelle bibliographie mise à jour]. Translated by M. Narcy. Paris: Flammarion.
- Ricoeur, P. (1996) *Soi-même comme un autre*. Paris: Éd. du Seuil (Points).
- Rossetti, L. (2011) *Le dialogue socratique*. Paris: Les Belles lettres (Encre marine).
- Schön, D.A. (1983) *The reflective practitioner: how professionals think in action*. New York: Basic Books.
- Stiegler, B. (1994) *La faute d'Epiméthée*. Paris: Galilée la Cité des sciences et de l'industrie (La technique et le temps, 1).
- Stiegler, B. (2018) *La technique et le temps: suivi de Le nouveau conflit des facultés et des fonctions dans l'Anthropocène*. Réédition. Paris: Fayard.
- Stiegler, B. and Stiegler, B. (2015) *L'avenir du travail*. Paris: Fayard (La société automatique, 1).
- Tozzi, M. (2007) *Apprendre à philosopher par la discussion: pourquoi ? comment ?* Bruxelles [Paris]: De Boeck (Perspectives en éducation & formation).
- Vallor, S. (2016) *Technology and the virtues: a philosophical guide to a future worth wanting*. New York (N.Y.): Oxford University press.
- Zagzebski, L.T. (1998) *Virtues of the mind: an inquiry into the nature of virtue and the ethical foundations of knowledge*. Reprinted. Cambridge: Cambridge Univ. Press.
- Zagzebski, L.T. (2015) *Epistemic authority: a theory of trust, authority, and autonomy in belief*. Oxford University paperback. New York: Oxford University Press.

Table des matières

<i>Abstract</i>	1
<i>1. Scénario : Le miroir moral.....</i>	2
<i>2. Contexte et interrogations.....</i>	5
<i>3. Hypothèses.....</i>	6
<i>4. Développement des arguments.....</i>	8
4.1. En rapport avec la première question	8
4.2. En rapport avec la deuxième question.....	10
4.3. En rapport avec la troisième question	12
4.4. En rapport avec la quatrième question.....	14
4.5. En rapport avec la cinquième question.....	16
<i>5. Conclusion.....</i>	18
<i>Acknowledgement.....</i>	19
<i>Bibliographie</i>	20
<i>Table des matières.....</i>	22