

Évaluer à l'ère de l'intelligence artificielle générative : pourquoi nos examens vacillent, et comment repenser ce qu'ils mesurent

Julie Mulet – Ingénieure pédagogique (LHUB, ENAC)

Un étudiant remet un devoir particulièrement bien rédigé. Les idées sont claires, la structure solide, le raisonnement convaincant. Pourtant, une incertitude s'installe : a-t-il réellement produit ce travail ? Et si oui, qu'a-t-il mobilisé pour y parvenir ?

Depuis l'arrivée des intelligences artificielles génératives (IAG), cette scène est devenue familière dans l'enseignement supérieur. Très vite, le débat s'est cristallisé autour de questions de triche, de contrôle et de détection. Mais à y regarder de plus près, l'IAG ne fait que rendre visibles des fragilités bien plus anciennes. Elle agit comme un révélateur d'un problème de validité : nos évaluations mesurent-elles réellement ce que nous pensons mesurer ? Explorer cette problématique amène à se poser une question bien plus fondamentale : que souhaite-t-on évaluer – et donc enseigner – dans l'enseignement supérieur ?

Cet article défend une idée simple : l'IAG ne crée pas une crise de l'évaluation, elle met en crise un modèle d'enseignement déjà fragilisé. Pour le comprendre, il faut revenir aux raisons d'être de l'évaluation traditionnelle, analyser ses effets sur les apprentissages, puis examiner ce que l'IAG met brutalement en lumière. A partir de ce diagnostic, nous verrons que l'approche par compétences apparaît non comme une mode, mais comme une réponse cohérente aux défis actuels. Pour autant, il ne s'agit pas de se reposer sur de simples ajustements superficiels des gestes enseignants : cette problématique implique d'interroger une représentation globale de l'enseignement supérieur et de l'évaluation, et d'appeler à une évolution systémique.

1. Un état des lieux critique : des évaluations traditionnelles déjà sous tension

Des formats historiquement dominants, aujourd'hui fragilisés

Dans l'enseignement supérieur, l'évaluation prend le plus souvent la forme d'examens écrits, de QCM, de questions de cours, d'exercices d'application ou de productions normées. Ces formats, bien connus des enseignants comme des étudiants, ont été privilégiés pour répondre à des contraintes fortes : gérer des effectifs importants, assurer une comparabilité des résultats, garantir une certaine équité de traitement. Ils s'inscrivent dans une logique majoritairement sommative : il s'agit d'évaluer les apprentissages, une fois ceux-ci supposément réalisés, et d'en certifier le niveau (Birenbaum et al., 2006). Dans cette perspective, l'évaluation intervient souvent comme un moment distinct, en fin de parcours,

venant sanctionner un apprentissage pensé comme linéaire : l'enseignant transmet, l'étudiant assimile, puis restitue (Rege-Colet & Berthiaume, 2013).

L'arrivée des IAG a très clairement mis ces formats en tension. Lorsqu'un étudiant peut produire en quelques secondes une réponse structurée, une synthèse cohérente ou un raisonnement apparemment solide, la question de la fraude, et des moyens de la détecter avec assurance, se pose. Evacuons directement ce dernier point : les outils de détection d'IAG ne sont pas assez fiables pour être une réponse soutenable (Weber-Wulff et al., 2023). Mais surtout, la question de la fraude masque une problématique plus profonde : que cherche-t-on réellement à évaluer, et comment s'assurer de la validité de nos formats d'évaluation ?

Des effets contre-productifs sur les stratégies d'apprentissage des étudiants

Indépendamment de l'IAG, de nombreux travaux ont montré que les modalités d'évaluation influencent fortement la manière dont les étudiants apprennent. Lorsque les tâches évaluatives privilégient la restitution, la conformité à des attentes implicites ou la reproduction de procédures, les étudiants adoptent des stratégies rationnelles d'optimisation : apprendre « juste ce qu'il faut », s'entraîner sur des formats types, « bachoter » en visant la bonne note plutôt que la compréhension (Dweck, 2013). Ces stratégies sont souvent également renforcées par une absence de sens : les étudiants peinent à comprendre l'utilité des savoirs mobilisés, et notamment leur lien avec des situations professionnelles futures. L'évaluation devient alors, au mieux, un « jeu scolaire », dans lequel l'objectif principal devient la conformité aux attentes institutionnelles plutôt que le développement de compétences professionnelles (Wiggins, 1990). Dans le pire des cas, elle devient une source d'anxiété, voire de découragement, plutôt qu'un levier d'apprentissage (Dweck, 2013).

Dans ce contexte, l'IAG peut apparaître comme une solution pragmatique pour les étudiants : un moyen de réduire l'effort requis pour satisfaire les exigences formelles de la tâche et passer un seuil certificatif. Là encore, l'outil agit comme un révélateur : si une tâche peut être déléguée sans perte perçue de sens, c'est peut-être qu'elle n'engage pas les apprentissages les plus essentiels du point de vue des apprenants. Ces éléments d'analyse fragilisent encore davantage l'illusion de validité de nos formats traditionnels d'évaluation.

Un problème central : la validité des évaluations

Pour comprendre ces fragilités, le modèle du triangle de l'évaluation (adapté de Pellegrino et al., 2001) constitue un outil particulièrement éclairant. Ce cadre invite à analyser et concevoir toute situation évaluative à travers trois pôles interconnectés :

1. **La cognition**, c'est-à-dire les processus mentaux que l'on cherche à évaluer chez les étudiants ;
2. **La tâche** proposée à l'étudiant pour estimer s'il sait mobiliser ces processus mentaux ;

3. **L'interprétation**, c'est-à-dire le jugement que l'on peut porter sur la base des observable (ce qui est produit par l'élève).

Une évaluation valide suppose un alignement solide entre ces trois dimensions (Messick, 1995). Or, très souvent dans les évaluations traditionnelles, cette cohérence est fragile. Les enseignants observent un produit final (une réponse écrite, un calcul, un document textuel normé...) et en déduisent des apprentissages. Pourtant, les recherches en psychologie cognitive montrent qu'un même produit peut résulter de processus très différents (Ericsson & Simon, 1993). Une réponse correcte peut masquer une compréhension superficielle, et un raisonnement pertinent peut ne pas aboutir à un résultat attendu. Une tâche mal choisie, en particulier lorsqu'elle est fortement normée ou focalisée sur le résultat, met donc à mal la validité de l'interprétation. En outre, le pôle de la cognition est souvent défini de manière implicite, voire réductrice. Les tâches évaluatives se centrent généralement sur la restitution des connaissances présentées en cours, et visent rarement à observer des compétences complexes telles que l'analyse de situations authentiques, la prise de décision en contexte incertain, ou la mobilisation intégrée de savoirs — pourtant centrales dans le monde professionnel. L'irruption de l'IAG dans l'enseignement supérieur met ainsi en lumière deux déséquilibres anciens :

- Premièrement, les évaluations manquent de validité, car elles sont souvent insuffisamment alignées avec les apprentissages visés, tant du point de vue de la tâche que du point de vue de l'interprétation.
- Deuxièmement, les apprentissages visés sont souvent décalés des attentes réelles des apprenants et de leurs futurs environnements professionnels.

Loin d'être la cause première de la crise actuelle de l'évaluation, l'IAG en accentue les tensions et en révèle les angles morts. Elle oblige à interroger nos formats traditionnels d'évaluation, mais aussi les conceptions pédagogiques qui les sous-tendent. Revenir au triangle de l'évaluation permet de comprendre l'origine de ces tensions et d'envisager leur dépassement, non comme une contrainte, mais comme une opportunité de transformation.

2. Ce que l'IAG accentue... et comment dépasser les tensions

Le volet de la cognition : redéfinir ce qui est attendu à l'aune du monde professionnel

Le volet de la cognition est sans doute celui que l'IAG oblige le plus radicalement à reconsidérer. Il renvoie à la détermination de ce qu'un étudiant doit savoir (concepts, principes, faits...), de ce qu'il doit savoir faire (mobilisation de stratégies, de procédures, de raisonnements...), et de la façon dont ces éléments s'organisent et se développent avec l'apprentissage (étapes typiques, erreurs fréquentes, niveaux d'expertise). Aussi, ce volet implique de s'interroger sur ce que l'on attend réellement des étudiants à l'issue de leur formation : des connaissances théoriques et conceptuelles ? Des procédures maîtrisées ? Ou

la capacité à mobiliser des ressources variées pour résoudre des problèmes complexes, situés et évolutifs ?

De nombreux employeurs expriment depuis longtemps des attentes claires : autonomie, esprit critique, capacité d'adaptation, collaboration, résolution de problèmes complexes et non routiniers (Li, 2024). Or, ces compétences sont rarement évaluées de manière explicite.

De plus, aujourd'hui, une majorité d'employeurs considère les compétences en IAG comme essentielles. Selon une enquête Access Partnership pour Amazon Web Services¹, 93 % des employeurs prévoient d'utiliser l'IAG au sein de leur organisation dans les cinq prochaines années, ce qui génère une demande accrue de compétences en termes de maîtrise et d'intégration de l'outil. Loin d'être une solution de facilité, utiliser une IAG de manière pertinente suppose des compétences complexes : formuler une demande adéquate, évaluer la qualité d'une réponse, corriger une production, arbitrer ce qui est acceptable ou non. Cela relève moins de la restitution de connaissances que de la régulation de l'activité cognitive dans un environnement outillé.

Dans ce contexte, une question devient incontournable : qu'attend-on que l'étudiant sache faire sans assistance, et qu'attend-on qu'il sache faire avec des outils ? L'IAG oblige ainsi à expliciter le contrat cognitif de l'évaluation. Cette clarification est indissociable du contrat plus large liant enseignement supérieur et monde professionnel, qui attend désormais des diplômés capables d'utiliser l'IAG de manière efficace, critique et responsable — des compétences qui ne vont pas de soi, et qui doivent être enseignées et évaluées.

Repenser la tâche : de la conformité à l'authenticité

L'IAG brouille profondément la frontière entre ce qui est internalisé par l'étudiant et ce qui est externalisé dans l'outil. Une production peut être correcte sans que l'on puisse déterminer si elle résulte d'une véritable compréhension, d'un raisonnement partiellement délégué ou d'une simple orchestration efficace de l'outil. Face à cela, la première tentation est souvent de modifier la tâche pour la rendre « résistante » : interdictions d'outils, sujets plus complexes, contraintes supplémentaires. Mais une approche plus pertinente consiste à interroger la nature même des tâches proposées, en particulier au regard du volet de la cognition. Il s'agit pour l'enseignant de se demander : quelle tâche me permet d'observer de manière pertinente si l'objectif d'apprentissage est atteint ? Dans le triangle de l'évaluation, cela revient à renforcer le lien entre la tâche et la cognition visée, plutôt qu'à complexifier artificiellement les consignes.

Lorsque la performance outillée est exclue, l'objectif de l'évaluation est de s'assurer que l'étudiant dispose d'un socle non externalisé, mobilisable sans assistance. L'usage de l'IAG entre alors directement en contradiction avec ce qui est visé, puisqu'il empêche toute inférence valide sur la maîtrise interne de l'étudiant. Des modalités d'évaluation classiques,

¹ <https://assets.aboutamazon.com/e1/a0/17842ee148e8af9d55d10d75a213/aws-accelerating-ai-skills-us-en.pdf>

fortement contraintes (examens en présentiel, ressources limitées, temps restreint), restent ici pleinement légitimes.

En revanche, lorsque la performance outillée est autorisée, voire requise, l'usage doit être explicitement intégré dans la tâche, dans les consignes et dans les critères d'évaluation. L'objet évalué ne se limite plus au produit lui-même, mais inclut la capacité à piloter l'IAG : formuler des demandes pertinentes, évaluer la qualité des réponses proposées, les corriger, les contextualiser et prendre des décisions éclairées à partir de celles-ci. L'IAG ne masque pas la compétence ; elle devient au contraire un révélateur de la qualité du raisonnement et de la régulation cognitive de l'étudiant. L'attention portée au processus d'accomplissement de la tâche (démarches, choix, régulations, justifications...) devient une condition centrale de validité.

Vers une interprétation valide : rendre visibles les processus

L'IAG questionne également le statut de l'interprétation. Lorsque seule la production finale est évaluée, il devient difficile de distinguer ce qui relève du travail de l'étudiant de ce qui a été délégué à l'outil. A ce stade, et après avoir clarifié le volet de la cognition et le volet de l'observation, l'enseignant doit se demander : quelles traces me permettent d'observer de manière valide si l'objectif d'apprentissage est atteint ?

Un premier levier consiste à diversifier les observations pour interpréter les apprentissages. Lorsque l'évaluation repose sur une trace unique (un devoir écrit, un examen ponctuel, un livrable final), toute ambiguïté sur les conditions de production fragilise fortement l'inférence. À l'inverse, le croisement de plusieurs traces (productions écrites et orales, traces de processus réflexifs pendant l'accomplissement de la tâche, évaluations étalées dans le temps...) permet de mieux apprécier les processus cognitifs mobilisés. Cette logique de multiplication des traces, bien documentée en sciences de l'éducation, réduit les risques de sur- ou de sous-interprétation et renforce l'équité des décisions (Pellegrino et al., 2001).

L'explicitation des critères d'interprétation, indispensable pour concevoir des tâches d'évaluation pertinente, est tout autant un levier concernant l'interprétation des traces. Elles permettent de rendre l'inférence plus transparente, tant pour l'enseignant que pour l'étudiant.

Enfin, les enjeux liés à l'IAG rappellent que les décisions évaluatives à forts enjeux (certification, diplomation, sélection) ne devraient pas reposer sur une interprétation individuelle et isolée. Des dispositifs tels que les jurys, la modération des notes ou les décisions collectives permettent de confronter les interprétations, de limiter les biais individuels et de reconnaître la dimension fondamentalement sociale et institutionnelle de l'évaluation (Birenbaum et al., 2006). Dans un contexte où les conditions de production des performances sont plus hétérogènes, cette dimension collective devient un élément clé de la crédibilité des décisions. Dans le cadre du triangle de l'évaluation, cela revient à enrichir le lien entre production et cognition, en multipliant les indices observables.

3. Des évolutions aux bénéfiques multiples : l'apport explicite de l'approche par compétences

Une réponse cohérente aux limites révélées

Les évolutions esquissées précédemment convergent vers une approche déjà largement documentée en pédagogie : l'approche par compétences (APC). Être compétent ne consiste pas à posséder des connaissances isolées, mais à savoir les mobiliser de manière pertinente, combinée et située, en fonction des exigences d'une situation donnée (Le Boterf, 2010; Tardif, 2006). Dans cette perspective, l'acquisition de connaissances est une étape fondamentale de l'enseignement supérieur, mais pas une finalité : si les connaissances théoriques et conceptuelles sont des ressources centrales et indispensables, elles ne prennent sens que dans leur mobilisation effective au service de l'action.

La compétence est ainsi définie comme un « savoir-agir complexe », qui repose sur la sélection la mobilisation et l'articulation de ressources multiples : des ressources internes (connaissances, habiletés, expériences, dispositions) mais aussi des ressources externes, telles que des outils, des documents, des normes, des dispositifs techniques ou des interactions avec autrui. Ainsi, l'activité compétente n'est jamais purement individuelle ni décontextualisée : elle est toujours située et médiée par des ressources disponibles dans l'environnement. L'IAG peut ainsi être considérée, dans certaines situations, comme une de ces ressources externes, au même titre que d'autres outils numériques ou documentaires.

Si l'on se réfère au triangle de l'évaluation, l'APC favorise un alignement renforcé entre les trois pôles de l'évaluation, en promouvant des tâches authentiques, proches des situations professionnelles, des productions diversifiées intégrant des traces de processus, et des attentes explicitement définies du point de vue cognitif, centrées sur la mobilisation et le transfert des ressources.

Un intérêt formatif renforcé

L'un des apports majeurs de ces approches réside dans leur potentiel pédagogique. La distinction classique entre évaluation sommative et évaluation formative permet d'éclairer ce changement de perspective. L'évaluation formative vise prioritairement à soutenir l'apprentissage, en fournissant des informations exploitables par l'étudiant et par l'enseignant pour ajuster les stratégies d'apprentissage et d'enseignement (Black & Wiliam, 1998). De nombreux travaux montrent que des feedbacks réguliers, spécifiques et orientés vers les processus favorisent des apprentissages plus profonds, le développement de l'autorégulation et le maintien de la motivation (Hattie, 2012; Winne & Hadwin, 1998).

L'évaluation devient alors un outil au service de l'apprentissage, et non plus seulement un instrument de vérification, de certification ou de sélection. Elle offre ainsi un espace d'apprentissage à part entière, qui contribue à réduire l'anxiété liée à l'évaluation, à

renforcer l'engagement cognitif et à limiter le recours à des stratégies purement instrumentales ou de contournement (Nicol & Macfarlane-Dick, 2006).

Dans un tel cadre, l'IAG peut être intégrée comme un outil au service de l'apprentissage et de l'action, à condition que son usage fasse l'objet d'un encadrement pédagogique explicite. L'enjeu n'est alors plus d'exclure l'outil, mais de rendre visibles et discutables les processus cognitifs qu'il soutient ou qu'il transforme.

Des compétences réellement transférables

Dans le prolongement de ces apports formatifs, l'APC apparaît également déterminante pour le développement de compétences réellement transférables. en particulier à travers le renforcement de l'autorégulation et de la réflexivité des apprenants. En confrontant les apprenants à des situations complexes et authentiques, elle les engage dans des activités de planification, de prise de décision, d'ajustement en cours d'action et d'analyse réflexive de leurs choix. Ces processus d'autorégulation et de réflexivité sont largement reconnus comme centraux dans la construction d'apprentissages durables et transférables à de nouvelles situations (Winne & Hadwin, 1998). Ces évolutions permettent également de mieux préparer les étudiants aux réalités professionnelles : évaluer la capacité à analyser une situation, à prendre des décisions argumentées, à collaborer ou à s'adapter à des contraintes nouvelles, c'est répondre à des attentes largement partagées, mais longtemps restées en marge des dispositifs d'enseignement et d'évaluation.

À l'ère de l'IAG, cet enjeu devient particulièrement saillant. Lorsque certains produits peuvent être générés ou assistés par des outils, la valeur formative des situations d'apprentissage se déplace vers la capacité des apprenants à piloter leur activité, à justifier leurs choix, et à porter un regard critique sur les ressources mobilisées, y compris l'IAG elle-même. En ce sens, l'APC offre un cadre particulièrement fécond pour développer des compétences transférables, ancrées dans la compréhension de l'action et dans la capacité à apprendre de ses propres pratiques.

4. Un enjeu systémique : accompagner les enseignants et transformer les cadres institutionnels

Le rôle central des institutions dans la soutenabilité du changement

Ces transformations ne peuvent reposer uniquement sur les enseignants. D'une part, concevoir des évaluations plus authentiques, analyser des productions complexes, multiplier les traces d'apprentissage, offrir des feedbacks de qualité... tout cela a un coût en temps et en énergie. D'autre part, et plus fondamentalement, le changement de paradigme que ces évolutions impliquent nécessitent un accompagnement solide. Aussi, pour que le changement soit soutenable, un engagement institutionnel fort est nécessaire : reconnaissance du travail pédagogique, formations, espaces de mutualisation, évolution des

règles de certification... Sans cette réflexion systémique, le risque est grand de faire peser sur les enseignants une injonction paradoxale : innover sans moyens supplémentaires. Or, la recherche en évaluation comme en conduite du changement montre que les transformations durables des pratiques pédagogiques ne s'opèrent que lorsqu'elles sont soutenues par des cadres institutionnels clairs, des ressources adaptées et des espaces de concertation reconnus.

Dans le contexte de la démocratisation de l'IAG, le rôle des institutions est d'abord de rendre explicites les attentes : clarifier les principes qui doivent guider les pratiques d'évaluation, définir des lignes directrices communes concernant les usages autorisés ou attendus des outils, et garantir une cohérence minimale entre les formations. Ce cadrage a vocation à sécuriser le contrat évaluatif, tant pour les enseignants que pour les étudiants et les partenaires extérieurs (employeurs, instances d'accréditation, jurys).

Des approches programme au cœur du processus de transformation

Les approches par compétences (APC) évoquées précédemment constituent un levier central, à condition de ne pas être mises en œuvre de manière fragmentée. Leur pleine portée se déploie lorsqu'elles s'inscrivent dans une logique de programme, c'est-à-dire dans une conception décroisonnée de l'enseignement et de l'évaluation, à l'échelle d'un parcours de formation cohérent (Van Der Vleuten & Schuwirth, 2005). Il ne s'agit alors plus de penser les évaluations comme des dispositifs isolés, attachés à des unités d'enseignement indépendantes, mais comme les composantes d'un système d'évaluation distribué, articulé et progressif.

Dans une logique de programme, l'évaluation d'une compétence ne repose ni sur une seule épreuve, ni sur un seul enseignant, ni sur un moment unique du cursus. Elle s'appuie sur un ensemble de preuves construites progressivement, à travers une pluralité de situations d'apprentissage et d'évaluation, mobilisant des contextes, des formats de production et des critères complémentaires. Au-delà de ses vertus pédagogiques pour les apprenants, le décroisonnement favorise une mutualisation du travail d'évaluation : construction collective de référentiels de compétences, élaboration partagée de critères, harmonisation des attentes et modération des jugements entre enseignants. Ce travail collectif contribue à réduire la charge individuelle tout en augmentant la robustesse des évaluations.

Enfin, l'approche programme rend le changement plus soutenable dans le temps. En répartissant les responsabilités évaluatives à l'échelle d'un collectif pédagogique, elle évite que l'adaptation aux nouveaux enjeux (notamment ceux liés à l'IAG) ne repose sur des initiatives isolées ou des efforts individuels difficilement tenables. Elle constitue ainsi un cadre structurant pour accompagner durablement la transformation des pratiques, en alignant exigences pédagogiques, contraintes organisationnelles et objectifs de formation.

L'IAG comme ressource pour l'enseignant

Enfin, si l'IAG pose des défis importants du point de vue de la validité des évaluations, elle peut également être envisagée, à certaines conditions, comme une ressource au service du travail enseignant. Sans en faire ici un axe central, puisqu'il ne s'agit pas du propos de cet article, on peut noter que ces outils offrent des potentialités pour soutenir la conception de situations d'évaluation, la formulation de critères, la conception de situations authentiques ou encore accompagner certaines étapes du feedback. Ces usages, lorsqu'ils sont pensés de manière critique et encadrée, peuvent contribuer à alléger certaines charges, sans se substituer au jugement pédagogique et évaluatif, qui demeure fondamentalement humain.

En conclusion

L'irruption de l'IAG ne signe pas la fin de l'évaluation dans l'enseignement supérieur ; elle en révèle les fragilités conceptuelles. En mettant sous tension la validité des inférences que nous tirons à partir des productions étudiantes, elle contraint à réinterroger trois pôles : ce que l'on vise réellement du point de vue cognitif, les tâches que l'on mobilise pour l'observer, et les interprétations que l'on en déduit.

Ce déplacement est exigeant, car il ne se limite pas à ajuster des formats ou à renforcer des contrôles. Il suppose de clarifier le contrat cognitif et professionnel de la formation, de concevoir des situations d'apprentissage et d'évaluation plus authentiques, de rendre visibles les processus d'apprentissage, et d'inscrire ces évolutions dans des cadres institutionnels cohérents et collectifs. En ce sens, l'IAG agit moins comme une menace que comme une opportunité pour une transformation déjà documentée par la recherche en sciences de l'éducation : passer d'une logique centrée sur la restitution de connaissances à une logique centrée sur la mobilisation située des ressources internes et externes.

L'enjeu dépasse donc la question des outils. Il touche à la crédibilité des décisions certificatives, à la cohérence des parcours de formation et, plus fondamentalement, à la définition même de ce que signifie « être certifié compétent » à l'ère des environnements cognitifs outillés. Plutôt que de chercher à restaurer artificiellement un modèle fragilisé, l'enseignement supérieur est invité à assumer cette transition : concevoir des évaluations capables de distinguer la simple production de la véritable compétence, et reconnaître que, dans un monde équipé d'IAG, la compétence inclut aussi la capacité à en faire un usage réfléchi, critique et responsable.

Références bibliographiques

- Birenbaum, M., Breuer, K., Cascallar, E., Dochy, F., Dori, Y., Ridgway, J., Wiesemes, R., & Nickmans, G. (2006). A learning Integrated Assessment System. *Educational Research Review*, 1(1), 61–67. <https://doi.org/10.1016/j.edurev.2006.01.001>
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *International Journal of Phytoremediation*, 21(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Dweck, C. S. (2013). Self-theories: Their role in motivation, personality, and development. *Self-Theories*. <https://doi.org/10.4324/9781315783048/SELF-THEORIES-CAROL-DWECK>
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol Analysis: Verbal Reports as Data*. The MIT Press. <https://doi.org/10.7551/mitpress/5657.001.0001>
- Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge. https://scholar.google.com/scholar?hl=fr&as_sdt=0%2C5&q=hattie+2012&btnG=
- Le Boterf, G. (2010). *Repenser la compétence: pour dépasser les idées reçues: quinze propositions* (Editions Eyrolles). <https://books.google.com/books?hl=fr&lr=&id=EsZFPPbwGPKC&oi=fnd&pg=PA7&dq=L+e+Boterf,+2010+&ots=yaOz4eQv1y&sig=qAhI0GEqH6UGV4vR19bh7Ololc>
- Li, L. (2024). Reskilling and Upskilling the Future-ready Workforce for Industry 4.0 and Beyond. *Information Systems Frontiers*, 26(5), 1697–1712. <https://doi.org/10.1007/s10796-022-10308-y>
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741–749. <https://doi.org/10.1037/0003-066X.50.9.741>
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199–218. <https://doi.org/10.1080/03075070600572090>
- Pellegrino, J. W. ., Chudowsky, Naomi., & Glaser, Robert. (2001). *Knowing what students know : the science and design of educational assessment*. National Academy Press.
- Rege-Colet, N., & Berthiaume, D. (2013). Comment choisir des méthodes d'évaluation adaptées ? - In P. Lang (Ed.), *La pédagogie de l'enseignement supérieur : repères théoriques et applications pratiques* (Vol. 1, pp. 241–254).
- Tardif, J. (2006). *L'évaluation des compétences: documenter le parcours de développement*. Chenelière éducation.
- Van Der Vleuten, C. P. M., & Schuwirth, L. W. T. (2005). Assessing professional competence: From methods to programmes. *Medical Education*, 39(3), 309–317. <https://doi.org/10.1111/j.1365-2929.2005.02094.x>

Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., Šigut, P., & Waddington, L. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity*, 19(1), 26.
<https://doi.org/10.1007/s40979-023-00146-z>

Wiggins, G. (1990). The Case for Authentic Assessment. *Practical Assessment, Research, and Evaluation*, 2(1).

Winne, P., & Hadwin, A. (1998). Studying as self-regulated learning. *Metacognition in Educational Theory and ...*
<https://books.google.fr/books?hl=fr&lr=&id=EzWRAGAAQBAJ&oi=fnd&pg=PA277&dq=Studying+as+self-regulated+learning.&ots=lvDhYDAv7E&sig=ddxn1kJDukFoydBo3EIs5FlcHeg>