

OPTIMISER L'ÉVALUATION POUR TOUS:

Développement de tâches curriculaires qui intègrent les compétences du 21ème siècle





Auteurs

Esther Care est membre senior du Programme d'Économie Globale et Développement (« Global Economy and Development Program ») à la Brookings Institution.

Helyn Kim a été membre du programme d'Economie Globale et Développement (« Global Economy and Development Program ») à la Brookings Institution.

Aynur Gul Sahin a été une boursière Fullbright effectuant un stage post-études à la Brookings Institution.

Optimiser l'Évaluation Pour Tous (« Optimizing Assessment for All » (OAA)) est un projet de la Brookings Institution. L'objectif de l'OAA est de soutenir les pays dans l'amélioration de l'évaluation, de l'enseignement et de l'apprentissage des compétences du 21^{ème} siècle à travers le développement de la connaissance évaluative auprès des acteurs éducatifs au niveau national et régional et en se concentrant sur l'usage constructif de l'évaluation en éducation ainsi que sur le développement de nouvelles méthodes pour évaluer les compétences du 21^{ème} siècle.

Remerciements

La Brookings Institution est une organisation à but non lucratif dédiée à la recherche indépendante et aux solutions politiques. Sa mission est de conduire une recherche indépendante de haute qualité, et de fournir des recommandations innovantes et pratiques pour les décideurs politiques et pour le public sur la base de sa recherche. Les conclusions et recommandations de toute publication Brookings n'engagent exclusivement que leur(s) auteur(s), et ne reflètent pas les vues de l'Institution, de son administration ou de ses autres chercheurs. Les auteurs remercient le travail de traduction de Carla Campos Cascales.

De plus, Brookings tient à remercier l'appui fourni par Porticus.

Brookings reconnaît que la valeur qu'elle procure réside dans son engagement avec la qualité, l'indépendance et l'impact. Les activités soutenues par ses donateurs reflètent cet engagement et l'analyse et les recommandations ne sont pas déterminées ou influencées par une donation quelle qu'elle soit.

Crédit photo de couverture: Erdmiin Dalai Complex School, Dundgovi, Mongolia



TABLE DES MATIÈRES

04 INTRODUCTION

06 DÉVELOPPEMENT DES ÉVALUATIONS

06 RÉDACTION PRÉLIMINAIRE DES TÂCHES

07 PENSER À VOIX HAUTE

08 Révision

08 EXPERTISE

09 Révision

10 NOTATION

13 ESSAI PILOTE

14 Différences culturelles

14 Questions de traduction

14 Qualité des tâches et des items

15 Qualité des grilles de notation

15 DISCUSSION

17 BIBLIOGRAPHIE

18 ANNEXES

19 Annexe A: Pensée critique – Revenus du magasin

25 Annexe B: Collaboration – Mélanges hétérogènes

31 Annexe C: Collaboration (et résolution de problèmes) – Serrekunda

34 Annexe D: Résolution de problèmes – Le citronnier



INTRODUCTION

Ce quatrième rapport « Optimiser l'Évaluation pour Tous » (OAA) décrit le processus de développement des tests et des items mené par le Cambodge, la Mongolie et le Népal dans l'OAA Asie et par la République Démocratique du Congo, la Gambie et la Zambie dans l'OAA Afrique. En tant que quatrième de la série, ce rapport synthétise les deux approches légèrement différentes adoptées dans chacune des régions.

L'OAA en soi a été conçu pour explorer de quelle manière évaluer les compétences du 21^{ème} siècle (21CS) tout en situation cette exploration dans une philosophie qui utilise l'évaluation comme un outil constructif pour l'enseignement et pour chercher comment l'intégration des 21CS dans les classes peut approfondir l'apprentissage. Ce rapport décrit les principaux processus de développement et d'affinement des tâches d'évaluation, en plus de mettre l'accent sur les protocoles de notation, et est complété par une série d'annexes qui fournissent des exemples des processus itératifs grâce auxquels les tâches d'évaluation de l'OAA et leurs items ont été de plus en plus affinés. Le rapport est rédigé pour fournir des conseils qui peuvent être utilisés ou adaptés pour le développement ultérieur des items des 21CS.

L'Objectif de Développement Durable 4 (ODD4) des Nations Unies est d' « Assurer l'accès de tous à une éducation de qualité, sur un pied d'égalité, et promouvoir les possibilités d'apprentissage tout au long de la vie ». Pour réussir dans le monde d'aujourd'hui, les apprenants doivent posséder un large éventail de connaissances, de compétences, d'habitudes et de traits qui vont au-delà de ce qui a été traditionnellement enseigné en classe, comme par exemple les compétences basées sur la mémorisation par cœur. Ces vastes ensembles de compétences, telles que la pensée critique, la collaboration, la résolution de problèmes, les compétences en matière de technologie et d'information et la communication, entre autres, sont souvent appelés compétences du 21^{ème} siècle (21CS).

Malgré un accord majeur sur le fait que les écoles doivent adopter un agenda pour enseigner les 21CS, il y a encore un grand sujet de débat autour des 21CS – de quelles sont les compétences les plus importantes à quelles compétences devraient être enseignées et évaluées. En l'absence d'un consensus clair sur les moyens les plus efficaces d'enseigner et d'évaluer ces compétences, les éducateurs rencontrent des difficultés pour adapter le paradigme d'apprentissage et d'enseignement des 21CS dans la salle de classe.

L'OAA est un projet conçu pour améliorer l'enseignement et l'apprentissage des 21CS en classe par le développement d'une expertise dans l'évaluation de ces compétences. Les trois principaux objectifs du projet sont les suivants :

- Soutenir les acteurs nationaux et régionaux dans la mesure de l'apprentissage au sein des 21CS ;
- Démontrer comment aligner l'évaluation des compétences avec la manière dont elles sont inscrites dans le curriculum et la manière dont elles sont enseignées ; et
- Renforcer la capacité des systèmes éducatifs à intégrer les 21CS dans leur enseignement-apprentissage

Pour plus d'information concernant le projet OAA, consultez le premier rapport de la série ([Care & Kim, 2020](#)).

Le but de ce rapport est d'aider les éducateurs à adapter et à développer de nouvelles méthodes d'évaluation des 21CS dans leurs classes. Le rapport fournit des informations sur le processus suivi dans l'OAA et comprend des exemples de stratégies d'évaluation pour saisir les 21CS, en particulier la résolution de problèmes, la collaboration et la pensée critique. Le rapport comprend des tâches d'évaluation développées par les trois pays cibles en Asie (Cambodge, Mongolie et Népal) et les trois pays en Afrique (République démocratique du Congo, Gambie et Zambie), proposant un guide étape par étape pour le développement et l'adaptation des tâches d'évaluation des 21CS.



Alors que les nations du monde entier cherchent à intégrer un large éventail de compétences dans leurs curricula – ainsi que dans l'enseignement et l'apprentissage – la nécessité de préciser les résultats d'apprentissage et les normes de performance pour ces compétences est devenue saillante. L'évaluation est la clé non seulement de ces spécifications, mais aussi de l'enseignement des compétences. Les compétences du 21^{ème} siècle sont des compétences transférables, reconnues comme difficiles à mesurer. Pour les éducateurs qui se sont concentrés pendant de nombreuses années sur les mathématiques, les langues, les sciences et les sciences sociales, l'intégration de compétences cognitives et sociales dans l'enseignement et l'évaluation constitue un défi. L'approche décrite ici rassemble les types d'évaluation connus en classe avec le processus nécessaire pour aborder ces compétences difficiles à mesurer.

Les processus de développement des tâches entrepris par des équipes nationales en Asie et en Afrique, se sont concentrés sur l'évaluation des compétences en matière de collaboration, de pensée critique et de résolution de problèmes des élèves inscrits en fin d'école primaire et en début d'école secondaire.

Aucun des pays qui se sont engagés dans ce travail ne dispose encore de définitions ni de descriptions complètes des compétences. Par conséquent, le développement des évaluations a été précédé par la formulation de définitions et de descriptions qui ont fourni la structure à partir de laquelle les tâches d'évaluation ont pu être créées. Les pays n'ont pas non plus encore spécifié formellement les résultats de performance pour ces compétences. Cette situation reflète la réalité de la plupart des pays – où, en raison de l'intégration récente des 21CS dans l'éducation formelle, les parcours d'apprentissage qui décrivent les compétences ne sont pas encore développés. Par conséquent, la définition du niveau de difficulté des tâches d'évaluation a été réalisée par les équipes nationales et leurs enseignants participants en estimant le niveau probable de compétence des élèves. La précision de ces estimations peut être constatée dans l'analyse des résultats des programmes pilotes des tâches dans les régions de l'OAA Asie et de l'OAA Afrique. Ces estimations figurent dans les deux rapports précédents de cette série. ([Care, Vista, & Kim, 2020](#); [Kim & Care, 2020](#)).

Figure 1. Étapes de développement des tâches





DÉVELOPPEMENT DES ÉVALUATIONS

Le projet de l'OAA a entrepris le développement de l'évaluation de manière légèrement différente dans ses deux régions d'Asie et d'Afrique. Toutefois, aux fins du présent rapport, la Figure 1 illustre ces différences synthétisées en une série d'étapes combinées : (1) rédiger les tâches de manière préliminaire ; (2) mener des séances de pensée à voix haute et les réviser ; (3) organiser des groupes d'expertise et réviser ; et (4) piloter les tâches.

DRAFT TASKS

Cette étape comprend quatre activités :

- Identifier, définir et la déconstruire les compétences ciblées
- Identifier des tâches d'évaluation existantes ou en développer des nouvelles qui puissent « capturer » les 21CS
- Adapter ou modifier ces tâches pour cibler de manière explicite les 21CS
- Identifier les compétences, les composantes et les sous-composantes que la tâche et ses items visent.

Identifier, définir et déconstruire les compétences. Lors de la rédaction d'une tâche d'évaluation, la première étape consiste à identifier la compétence d'intérêt. Sa définition et sa description sont nécessaires. De nombreuses 21CS sont des compétences complexes ; elles comportent plusieurs sous-compétences contributives. Par conséquent, les descriptions doivent clarifier ces sous-compétences afin qu'elles puissent être prises en compte dans les tâches d'évaluation.

La littérature sur le développement des tests et des échelles fait référence à ces sous-compétences et à leurs composantes de plusieurs façons. Dans ce rapport, les compétences complexes sont « déconstruites » à partir du niveau global de « compétence » en composantes, sous-compétences et sous-sous-compétences. Cela permet de rédiger des tâches avec des items qui ciblent différentes parties de la compétence complexe.

Identifier, développer ou adapter des tâches d'évaluation pour cibler les 21CS. Ensuite, plutôt que de partir de zéro pour développer des items complètement nouveaux qui ciblent les 21CS, des outils d'évaluation existants peuvent être révisés pour examiner s'ils peuvent être modifiés afin d'évaluer les 21CS, en prenant en considération les différents composantes et sous-compétences des compétences cibles à mesurer. Au moment de modifier les évaluations, quelques points en termes de qualité doivent être considérés. Plus particulièrement, une évaluation de haute qualité doit comprendre :

- Une intention claire ;
- Un langage compris par la plupart d'apprenants ;
- Un contexte simple et authentique ;
- Une seule réponse correcte ou un ensemble de réponses acceptables ; et
- De bonnes propriétés psychométriques.

Identifier les compétences, composantes et sous-compétences que chaque tâche devrait viser. Finalement, une fois que les tâches d'évaluation ont été rédigées de manière préliminaire, les compétences spécifiques, les composantes et les sous-compétences sont identifiés pour chaque tâche. Les évaluations décrites dans ce rapport consistent en ce que l'on appelle une « tâche », chacune comprenant plusieurs items et s'appuyant sur l'une des disciplines suivantes : mathématiques, sciences ou sciences sociales. Ainsi, la discipline du programme enseignée fournit la substance ou les connaissances sur lesquelles les élèves peuvent exercer leurs compétences.

Pour un programme d'évaluation solide et à grande échelle, un schéma directeur des compétences avec leurs composantes et sous-compétences, ainsi que du contenu curriculaire, serait établi pour encadrer le développement des évaluations. Dans le cas du programme OAA, avec son approche exploratoire du développement des tâches, le plan était moins strict.



Il identifiait les compétences à cibler dans les différentes disciplines, selon les niveaux scolaires sélectionnés. Le nombre de tâches et d'items ciblant des composantes et des sous-compétences spécifiques a été compté après que les séries finales de tâches d'évaluation aient été convenues.

Le projet OAA a reconnu que les différents pays définiront et décriront les 21CS de différentes manières. Une vaste littérature de recherche s'est accumulée au cours des dernières décennies (par exemple, Care, Griffin & Wilson, 2018 ; Jones, Bailey, Brush, Nelson & Barnes, 2016), mais elle n'est en aucun cas définitive. La manière dont les 21CS sont perçues par les différentes nations et cultures est reconnue comme un élément essentiel pour la manière dont l'enseignement et l'apprentissage sont délivrés. Et ces perceptions encadrent la manière dont les compétences sont intégrées dans les programmes, la manière dont elles sont enseignées et la manière dont elles sont évaluées. Les structures légèrement différentes des composantes des compétences identifiées par l'OAA Asie et l'OAA Afrique témoignent de ces différences. Ces différences ont naturellement des implications sur la création et la conception des tâches.

Pour amorcer la rédaction des tâches, deux approches ont été adoptées. L'une consistait à prendre des évaluations déjà existantes et à les adapter pour intégrer les 21CS ; l'autre consistait à vérifier les curricula pour les sujets communs aux pays participants et à sélectionner des sujets parmi ceux-ci pour créer des tâches d'évaluation. Quelle que soit l'approche adoptée, le processus d'élaboration des tâches suivant a été suivi. L'annexe de ce rapport fournit des exemples concrets d'adaptation des tâches pour intégrer les 21CS.

PENSER À VOIX HAUTE

Après la rédaction liminaire d'une tâche, il faut vérifier si les compétences prévues avec leurs composantes et sous-compétences sont réellement sollicitées par la tâche. La méthode de la « Pensée à voix haute » (également appelée laboratoire cognitif) est une méthode qui permet d'étudier comment un apprenant réagit à une tâche, si la tâche semble susciter les connaissances, les aptitudes ou les compétences prévues et si elle est appropriée en termes de facteurs tels que le genre, la culture et la gamme de performances attendues.

Dans ce cas, le processus de pensée à voix haute a été développé pour vérifier si les tâches ont suscité chez les élèves des processus cognitifs et sociaux cohérents avec les compétences, les composantes et les sous-compétences visés. Au cours de ce processus, l'élève accomplit une tâche tout en rendant compte oralement de ses processus mentaux en expliquant sa pensée et son raisonnement. Le flux de conscience est enregistré et utilisé pour évaluer si les compétences ciblées sont effectivement suscitées par les tâches.

En plus des observations générales au fur et à mesure que les élèves accomplissent la tâche, les questions suivantes peuvent guider les séances de pensée à voix haute et identifier tout problème lié à la tâche et à ses items :

- Est-ce la tâche ou l'item semblent tester la compétence ciblée ?
- Est-ce que la tâche ou l'item semblent tester la composante ou la sous-compétence visés ?
- Est-ce que le niveau de difficulté spécifié dans la tâche ou l'item est approprié pour les élèves ciblés ?
- Est-ce que le temps alloué à la tâche ou à l'item est approprié ?
- Est-ce que les élèves comprennent exactement ce qu'on leur demande de produire ?
- Y a-t-il d'autres types de réponses qui n'ont pas été prises en compte dans le guide de notation ?



Pour vérifier si les tâches sont accessibles de la même manière pour tous les élèves, plusieurs élèves de tout l'éventail de capacités doivent participer au processus. Lorsque le processus est entrepris pour vérifier les compétences en matière de collaboration, des groupes à la fois homogènes et hétérogènes (en termes de capacités) devraient participer.

Révision

Les activités de pensée à voix haute apportent aux enseignants et développeurs de tâches des informations sur la manière dont les élèves abordent une tâche et sur tout problème fonctionnel qui devrait être adressé au moment de réviser la tâche. Sur la base de ces preuves empiriques, des révisions sont apportées à cette étape de développement des tâches. Les preuves mettent généralement en évidence les questions d'administration, de contenu et de rapport.

Administration: Ce processus permet de rendre explicites les difficultés des élèves lorsqu'ils s'engagent dans des tâches, mais qu'ils résolvent par la suite. Cela aide les concepteurs de tâches à réfléchir à quels échafaudages (ou conseils) supplémentaires pourraient être nécessaires dans les instructions de tâches pour aider les élèves moins aptes. Ces difficultés immédiates rencontrées par les élèves peuvent être utilisées pour s'assurer que les instructions administratives sont adéquates.

Contenu: Les tâches qui sont élaborées en vue d'évaluer les 21CS utilisent des informations disciplinaires comme base. Il est important que les apprenants soient familiarisés avec ces informations avant de pouvoir les utiliser ou les manipuler. Les résultats de la pensée à haute voix aident à clarifier si les concepts ou les informations reflétés dans les tâches ont été couverts dans le curriculum délivré.

Rapport: Les étudiants peuvent répondre et réagir aux tâches d'une manière qui n'aurait pas été prévue par les concepteurs des tâches.

Ces réponses fournissent non seulement des indices précieux sur la façon dont les étudiants interprètent les tâches et les instructions, mais aussi des indications sur les capacités des élèves qui n'auraient pas été prises en compte dans les catégories de notation initiales. Les tâches qui stimulent des réponses pouvant être notées à différents niveaux de qualité constituent une riche source d'informations pouvant éclairer les stratégies d'enseignement.

EXPERTISE

Une fois que les révisions sont faites sur la base de l'étape de pensée à voix haute, les items de test révisés passent par un autre cycle de vérification, appelé « paneling ». Le paneling ou expertise est un processus de collaboration visant à examiner les aspects techniques et de fond des tâches d'évaluation. Il sert de contrôle de la qualité dans le processus de développement en examinant la validité et la pertinence contextuelle des matériaux (y compris les questions de partialité). Le processus d'expertise se concentre sur la tâche et les items tels qu'ils sont présentés, afin de déterminer s'ils échantillonnent les compétences aux niveaux prévus, s'ils sont susceptibles de discriminer entre les élèves de différentes capacités et si, en étant combinés avec d'autres tâches, le concept ciblé sera échantillonné de manière adéquate.

Un comité d'examen indépendant se compose d'experts en la matière, d'experts en évaluation et d'experts en langues, ainsi que d'enseignants des élèves pour lesquels la tâche est conçue. Une caractéristique majeure d'un groupe d'experts bien sélectionné est que les compétences et les expériences des différents membres du groupe sont complémentaires plutôt que redondantes. Au moins un représentant du créateur de la tâche fait partie du panel. Ce représentant sert de ressource pour le panel, c'est-à-dire qu'il fournit des explications selon les besoins, et non pas qu'il défend ou défend le matériel. La tâche, et tout matériel associé doivent être présentés de manière à ce qu'ils puissent se suffire à eux-mêmes.



Au cours des séances d'expertise, le groupe d'experts examine en détail toutes les parties du matériel d'évaluation, y compris le stimulus, les questions posées, la disposition et la présentation du matériel, ainsi que les guides et les critères de notation associés. Le but de ce processus est de s'assurer que l'administration, l'achèvement et la notation des items du test puissent être entrepris de manière efficace et efficiente. Les rôles et responsabilités des participants au panel sont décrits ci-dessous.

Un animateur/trice de groupe qui:

- Collecte et distribue le matériel d'évaluation à examiner ;
- S'assure que la sécurité soit maintenue
- Dirige et suit la discussion ;
- Veille à ce que l'objectif de l'élaboration du test établisse les paramètres de la discussion ; et
- Résume ce qui doit être fait pour réviser le matériel.

Membres individuels du panel qui:

- Travaillent sur le matériel, du point de vue du développeur et du candidat, en travaillant sur les tâches ou les items, en répondant aux questions et en révisant les critères de notation ;
- Critiquent de manière approfondie les items ;
- Font preuve d'esprit critique et résolvent les problèmes ;
- Identifient les faiblesses des tâches, lorsqu'elles existent, et identifient les alternatives ; et
- Rédigent des commentaires qui sont un « instantané » de leur propre pensée sans être influencés par les autres.

Le rédacteur/trice de l'article, sur demande:

- Exprime les raisons du choix du stimulus ou du contenu et de la création du format ou de la structure des items; et
- Articule les compétences ou le contenu spécifiques que l'item vise

Le processus final du groupe consiste à résumer si une tâche ou un item:

- Est acceptable tel qu'il est;
- Devrait être modifié de manière à corriger les lacunes ou problèmes identifiés ; ou bien
- N'est pas acceptable.

Révision

Lorsque l'expertise est terminée, les réponses du panel d'experts concernant la qualité des items devrait être consolidée et utilisée pour appliquer les changements nécessaires avant la finalisation des items. La révision des matériaux d'évaluation se poursuit jusqu'à ce que l'animateur/trice soit convaincu(e) que toutes les questions principales ont été adressées. Cela ne veut pas nécessairement dire qu'un consensus soit atteint sur chacune des questions, mais toutes les perspectives et opinions doivent être examinés avec soin.

Il est important de garder à l'esprit que la finalisation des matériaux à ce stade n'est pas la dernière étape du développement du test. Cette étape survient à l'examen du processus et des résultats de l'essai pilote des évaluations.



NOTATION

La notation est une partie intrinsèque du développement des matériaux d'évaluation – et un élément de conception qui doit être pris en compte à la fois au début de la création des tâches et tout au long du processus. Les règles de notation initiales sont généralement modifiées après les activités de pensée à voix haute et à nouveau après le pilotage des évaluations, sur la base des preuves présentées dans les réponses des étudiants.

Lorsqu'un item d'une tâche fournit une réponse incontestablement correcte ou incorrecte, la notation est simple. Cependant, lorsque ce n'est pas le cas, des grilles de notation sont fournies. Une grille de notation définit ce qui est attendu en termes de réponse et sur la manière dont elle doit être interprétée. Une bonne grille de notation décrit les critères à utiliser pour évaluer une réponse, avec des descripteurs de performance qui sont complets et mutuellement exclusifs – ce qui permet au notateur d'identifier différents niveaux de qualité de réponse (Figure 2).

Pour les items à réponse construite (c'est-à-dire ceux qui nécessitent une réponse libre de l'étudiant), les grilles sont particulièrement importantes car les niveaux de qualité sont généralement plus difficiles à identifier ou à quantifier. Comme exemple de grilles claires, bien que minimalistes, pour une telle situation, voir une tâche (D8) développée par l'équipe de la République Démocratique du Congo pour échantillonner les compétences de résolution de problèmes dans les sciences sociales (Figure 3).



Figure 2. Éléments de notation

Pré-requis pour la rédaction d'une grille

- Développer une liste de qualités (compétences, composantes ou sous-compétences) qu'on attend que les élèves démontrent
- Décider combien de niveaux de performance ou compétence peuvent être identifiés
- Développer des descripteurs pour chaque niveau de compétence, en commençant par les deux extrémités de l'échelle, puis par le milieu

Critères pour les catégories de notation

- Doivent être formulées dans un langage clair et univoque
- Doivent refléter la qualité de la performance ; chaque niveau de qualité reconnaissable doit être mutuellement indépendant et clairement distinguable
- Doivent permettre une déduction au sujet de l'emplacement d'un élève sur un parcours d'apprentissage

Règles de notation

- Se concentrent sur le comportement affiché
- Décrivent le comportement aux différents niveaux de compétence possibles
- Dédussent la compétence à partir du comportement

Niveaux de qualité

- Discriminent entre les niveaux d'apprentissage ou de compétence
- Reflètent les critères de notation basés sur l'analyse d'échantillons de travail
- Sont conformes aux descripteurs des compétences visées

Critères d'écriture

- Décrivent une performance qui soit observable (faire, dire, écrire)
- N'adressent qu'un seul élément central de la connaissance, du savoir-faire ou de la compétence
- Ne sont pas exprimés en termes comparatifs (par exemple, "meilleur") ou subjectifs (par exemple, "bons") car ne sont pas facilement normalisés



Figure 3. Tâche de résolution de problèmes D8 en sciences sociales avec des grilles génériques

Dans certaines villes ou zones rurales, lorsqu'une grande quantité de pluie tombe, il y a des inondations. Les maisons et les rues sont inondées, et il y a des conséquences importantes. a) Quelles sont les causes des inondations dans les villes et zones rurales ? Identifies-en au moins deux. b) Quelles sont les conséquences d'une inondation ? Identifies-en au moins deux. c) En utilisant l'une des causes et conséquences identifiées, quelle pourrait être une possible solution ?	D8a CI-Oi-Dec 0 = Pas de réponse 1 = Réponse non pertinente 2 = Fournit une cause possible 3 = Fournit deux causes possibles
	D8b CI-Oi-Dec 0 = Pas de réponse 1 = Réponse non pertinente 2 = Fournit une conséquence 3 = Fournit deux conséquences
	D8c PS-Ge-Hyp 0 = Pas de réponse 1 = Réponse non pertinente 2 = Fournit une solution possible mais qui n'est pas en lien avec une des causes ou conséquences identifiées 3 = Fournit une solution possible qui est en lien avec une des causes ou conséquences identifiées 4 = Fournit une solution possible qui est liée à une des causes et conséquences identifiées

Clé. CI = Collecte d'Informations; Oi = Organiser l'information; Dec = Décrire; PS = Planification d'une Solution; Ge = Générer des idées; Hyp = Faire des hypothèses

Les grilles fournissent au correcteur des indications schématiques mais adéquates pour déterminer si un élève a abordé la question sous l'angle des aspects des compétences. Les grilles ne couvrent pas d'autres aspects des compétences de l'élève, tels que la qualité de l'écriture ou même la qualité des idées ; elles couvrent la question de savoir si les causes ont été identifiées, si des associations ont été faites et si des solutions ont été apportées.



Pour la notation des données d'un grand nombre d'apprenants, ces grilles sont adéquates. Au niveau de la classe, cependant, l'enseignant peut souhaiter inclure des critères de notation supplémentaires qui concernent la qualité des connaissances en sciences sociales que l'élève apporte au problème ou la qualité de la littératie de l'élève. Ainsi, dans cet exemple présenté en Figure 3, les grilles fournissent des indications adéquates pour identifier si les processus de compétences requises ont été entrepris, grâce aux preuves du produit ; elles ne fournissent cependant pas de guidage sur la manière d'estimer la qualité du produit.

Un autre exemple, en Figure 4, illustre la provision de conseils aux correcteurs en énumérant les réponses acceptables ainsi qu'en fournissant les lignes directrices génériques. Cette tâche de l'OAA Asie, et également pilotée par l'OAA Afrique, a été conçue pour saisir les compétences de résolution de problèmes dans un contexte mathématique. Le graphique ne comprend que deux items qui sont basés sur cette tâche. Notez que la catégorie de score zéro dans la figure n'a été utilisée que par l'OAA Afrique en raison de son point de vue selon lequel l'effort devrait être récompensé. Le fait qu'un tel effort ne soit pas directement associé à la cible de l'évaluation est un problème au niveau psychométrique.

Une question essentielle rendue explicite par ces exemples de réponses d'élèves est l'interaction entre les compétences de résolution de problèmes, les connaissances et compétences mathématiques et les capacités de lecture et d'écriture. Cette interaction est un défi pour le correcteur car le niveau de performance dans l'un de ces trois domaines affecte la capacité de l'élève à faire preuve de compétence dans d'autres domaines.



Figure 4. Tâche de résolution de problèmes A1 en mathématiques avec des grilles génériques et des exemples pour répondre aux critères de notation pour deux des six items

Un agriculteur doit remplir d'eau un nouvel étang à poissons.  a) De quelle information a besoin l'agriculteur pour remplir l'étang ? Par exemple, tu dois savoir d'où provient l'eau. Écris deux ou trois informations utiles.	A1 a) CI Score 0 = Pas de réponse Score 1 = Donne un ou aucun facteur pertinent Score 2 = Donne deux facteurs différents liés au remplissage de l'étang à poissons Score 3 = Donne trois facteurs différents liés au remplissage de l'étang à poissons Exemples d'information utile: <i>Comment transférer l'eau vers l'étang (depuis la source), par exemple à l'aide d'un tuyau d'arrosage, d'une pipe, en creusant un canal, etc. Quelle est la taille de l'étang (ou combien d'eau il peut contenir) Combien d'eau y a-t-il déjà à l'intérieur (ou combien de plus il faut)</i> • <i>Combien de temps l'étang mettra-t-il à se remplir</i> • <i>Combien cela va coûter de le remplir</i> • <i>S'il doit se maintenir rempli ou pas</i> • <i>La quantité d'eau qui s'écoulera à travers les parois et le fond de l'étang</i> • <i>Comment les poissons vont accéder à l'étang (cela n'est pas directement lié au remplissage d'eau, mais peut être un point important)</i>
b) Un tuyau d'arrosage peut pomper 300 litres d'eau par minute. Combien de litres d'eau pompe ce tuyau en une demie heure ? A. 10 L B. 150 L C. 9,000 L D. 18,000 L c) L'agriculteur trouve un autre tuyau d'arrosage et l'utilise pendant 5 minutes. Il a pompé environ 1,000 litres. Est-ce que ce tuyau est plus rapide que le premier (300 litres par minute) ? Donne une justification pour ta réponse. d) L'étang à poissons a besoin de 36,000 litres d'eau pour être rempli. En utilisant le tuyau d'arrosage qui pompe 300 litres par minute, combien d'heures il faut pour que l'étang se remplisse ? e) L'agriculteur veut mettre 1,000 poissons dans l'étang, mais il pense que l'étang pourrait déborder. Que faudrait-il faire ?	A1 d) CI-Oi-Ana/Dec Score 0 = Pas de réponse Score 1 = Réponse incorrecte ou insuffisante • <i>Oui, 1,000 est plus élevé que 300 (réponse incorrecte)</i> • <i>Non, il n'est pas plus rapide (réponse correcte, mais aucune justification n'est donnée)</i> Score 2 = Explique pourquoi il N'EST PAS plus rapide, en montrant qu'il a un débit plus faible • <i>1,000 litres en 5 minutes est la même chose que 200 litres en 1 minute, ce qui est inférieur au premier tuyau</i> • <i>Le premier tuyau peut pomper 1,500 litres en 5 minutes, il est donc plus rapide</i> • <i>Non. $5 \times 300 = 1,500$ qui est plus élevé que 1,000</i>

Clé. CI = Collecte d'informations; Oi = Organiser l'information; Ana/Dec = Analyser/Décrire



Cela reflète bien sûr la réalité de la vie où il est rare que la démonstration de compétences ne soit pas affectée par d'autres facteurs, et constitue une justification majeure de l'enseignement et de l'évaluation des 21CS dans le contexte des études de domaines disciplinaires.

En résumé, des questions telles que les suivantes peuvent aider à revoir les grilles de notation:

- Est-ce que les descripteurs de performance sont clairs ?
 - Les élèves seraient-ils capables de s'auto-évaluer en ayant les descripteurs ?
 - Est-ce que les descripteurs pourraient donner aux élèves assez d'information pour savoir comment s'améliorer ?
- Est-ce que la grille d'évaluation reflète les niveaux auxquels les élèves sont susceptibles d'effectuer les tâches ?
- Les critères sont-ils suffisamment définis pour garantir que la notation est précise, impartiale et cohérente ?
- Est-ce que la grille reflète-t-elle le processus, le produit ou les deux ?

ESSAI PILOTE

Des essais des tâches d'évaluation ont été menés en Asie et en Afrique. La plupart des tâches administrées au sein d'une région a été développée dans cette région, bien que l'OAA Afrique ait également administré trois des tâches de l'OAA Asie. Le nombre d'élèves fournissant des résultats a varié : chaque pays de l'OAA Asie a testé plus de 300 élèves, et les pays de l'OAA Afrique ont testé des nombres variant de 120 à 209 et plus de 300. Les réponses des apprenants de ces essais pilotes ont été analysées (les résultats quantitatifs sont présentés dans les rapports OAA Asie et OAA Afrique de cette série) ([Care, Vista, & Kim, 2020](#); [Kim & Care, 2020](#)).

L'activité pilote a été conçue pour déterminer si les tâches et leurs items ont fonctionné comme prévu. En d'autres termes, la fonction de l'essai est de vérifier la qualité des évaluations, et non la performance des apprenants. L'analyse des réponses de nombreux élèves aux tâches d'évaluation permet de savoir si les tâches sont représentatives d'un éventail de capacités, si elles peuvent être trop difficiles ou trop faciles pour les élèves et si elles ciblent des groupes ayant les mêmes compétences ou aptitudes. Lorsque les analyses montrent que les élèves réagissent de manière inattendue à certaines tâches ou à certains items, cela fournit aux concepteurs de tests des indices permettant de déterminer si les tâches ont été mal conçues, si les grilles de notation n'ont pas été claires ou ont été appliquées de manière aléatoire, ou si la saisie et la gestion des données n'ont pas été suffisamment standardisées.

Étant donné qu'il se concentre sur l'évaluation de concepts difficiles à mesurer, l'OAA s'est particulièrement intéressé à l'adéquation des grilles de notation appliquées aux réponses des élèves. Cet intérêt est dû en grande partie au manque de familiarité reconnu des enseignants et des étudiants avec les compétences ciblées et les modes d'évaluation utilisés. Ce manque de familiarité est un thème récurrent qui est ressorti des commentaires des enseignants dans les six pays, et qui a été mis en évidence à la fois par l'engagement et les hésitations occasionnelles des élèves. Pour les enseignants qui ont participé aux processus de notation, la tâche était donc différente de leurs expériences de notation précédentes. Les exemples de réponses des élèves présentés dans la figure 5 sont issus des évaluations pilotes.

Lorsque les données au niveau des tâches et des items étaient anormales, les matériaux de stimulation eux-mêmes et les grilles de notation ont été examinés. Des possibles raisons pour ces anomalies incluent les différences culturelles, les problèmes de traduction et les problèmes de qualité des tâches, des items ou des grilles de notation.



Différences culturelles

Les tâches de l'OAA Asie ont été développées et alignées par les trois pays, l'idée étant que tous les pays pilotent toutes les tâches. Par conséquent, dans le cadre du processus de développement, les pays ont audité leurs curricula pour s'assurer que les étudiants de chacun des pays auraient tous eu accès aux mêmes apprentissages de contenus disciplinaires. Malgré cet effort général de similitude, les versions finales des tâches ont été légèrement modifiées pour s'assurer que les élèves de chaque pays puissent accéder aux tâches dans des conditions similaires – ce qui nécessitait l'utilisation de mots différents pour des items particuliers, de devises différentes et de conventions de dénomination différentes.

Pour l'OAA Afrique, la conception n'a pas supposé que tous les contenus thématiques seraient accessibles de la même manière dans les trois pays ou que tous les pays administreraient les mêmes tâches. Cependant les trois pays ont administré des tâches développées à l'origine par leurs propres équipes ainsi que celles développées par d'autres. Les trois équipes nationales ont collaboré sur tous les sujets tout au long du processus, de sorte qu'une fois encore, il y a eu une certaine modération des tâches qui les a rendues plus accessibles que ce qui aurait été le cas autrement.

Au cours du processus de développement des tâches de l'OAA Afrique, il y a eu de nombreuses discussions sur des phénomènes nationaux qui étaient vrais pour tous les pays – des questions telles que la surpopulation, les inondations et les maladies, toutes faisant l'objet de tâches dans le domaine des sciences sociales. La manière dont les différents pays pourraient répondre à ces problèmes et la façon dont leur compréhension de ceux-ci pourrait varier en fonction de la gravité des problèmes pourraient bien affecter les réponses des étudiants.

Il n'y a pas de doute sur le fait que les attitudes et les valeurs pourraient influencer ce qui est perçu comme acceptable ou inacceptable en termes de réponses pour certaines de ces tâches et leurs items.

Cela met en évidence une autre question pour l'évaluation de ces concepts difficiles à mesurer, où l'habileté démontrée ne reflète pas que les connaissances et les compétences, mais aussi les valeurs et les attitudes. Ainsi, lorsque les données de réponses à l'essai pilote montrent des anomalies au niveau d'un seul pays, un facteur à considérer et si les questions culturelles ont pu influencer la performance des élèves.

Questions de traduction

Les tâches ont été traduites depuis la langue originelle de développement à chacune des langues d'instruction de l'OAA Asie : Khmer, Népalais et Mongole. Lorsqu'une tâche était rédigée pour la première fois dans l'une de ces langues, elle était d'abord traduite en anglais pour la modération, puis dans les langues d'instruction.

Pour l'OAA Afrique, les tâches ont été fournies en anglais et en français, avec l'anglais comme langue d'instruction pour la Gambie et la Zambie. Aucune rétro-traduction n'a été entreprise, et il est probable que certaines maladroresses expressives se soient manifestées dans différentes versions. Celles-ci peuvent avoir influencé la compréhension et l'interprétation des tâches et de leurs items par les élèves. Similaire aux questions culturelles, où les données des réponses pilotes montrent des anomalies d'un seul pays, un facteur à prendre en compte est de savoir si les problèmes de langue ou de traduction peuvent affecter les performances des élèves.

Qualité des tâches et des items

Lorsque les tâches et leurs items ne semblent pas fonctionner comme prévu – en d'autres termes, si les élèves les plus compétents répondent de manière inexacte ou inappropriée aux items les plus simples, ou si les élèves les moins capables répondent de manière précise aux items supposés être les plus difficiles – cela pourrait signaler des problèmes avec les tâches ou avec les items eux-mêmes. La plupart des tâches rédigées par l'équipe de l'OAA comprenaient initialement un item relativement simple axé sur un sujet.



Les items suivants reposaient davantage sur les compétences visées, tout en tenant compte de la nécessité de rendre le contenu de la matière accessible à l'étudiant. Par conséquent, si les élèves ne répondaient pas correctement aux questions les plus simples, cela indiquait soit un défaut dans le matériel de stimulation de la tâche, soit dans la formulation de la question elle-même. Les données de réponse à une tâche peuvent être analysées pour déterminer si le matériel stimulant (c'est-à-dire l'ensemble de la tâche) est problématique ou s'il y a des problèmes isolés pour des items en particulier.

Qualité des grilles de notation

Les questions identifiées ci-dessus concernant les problèmes potentiels liés aux tâches et aux items sont difficiles à distinguer des problèmes liés aux grilles de notation. Lorsque les données de réponse des étudiants sont anormales, chacune des raisons ci-dessus doit être examinée. Lorsque les anomalies sont isolées pour un item particulier plutôt que pour plusieurs items associés à une tâche, le processus consiste à analyser la qualité de l'item et de la grille de notation qui lui est associée. Les données pilotes à grande échelle mettent en évidence le problème, mais ne peuvent pas l'identifier précisément. Et c'est justement cela qui nous amène à l'analyse qualitative des réponses des élèves, comme le montre la Figure 5.

Pour l'OAA Afrique, des réponses telles que celles de la Figure 5 ont été utilisées pour revoir les grilles utilisées pour la notation pilote. Ces réponses proviennent d'apprenants congolais et zambiens. En général, une quinzaine de réponses pour un item spécifique ont été analysées, quatre seulement étant présentées dans la présente figure. Dans un premier temps, les analystes trient les réponses du moins compétent au plus compétent en se référant aux grilles génériques. Le tri fournit une progression lâche de très faible à haute compétence. Lorsque les réponses ne correspondent pas exactement aux grilles, celles-ci sont ensuite affinées. Comme le montre la Figure 5, tous les documents ont été fournis en anglais et en français, avec les équipes nationales travaillant dans les deux langues.

Figure 5. Réponses des élèves à la tâche A1, Item (c)

(d) The farmer finds another hose and uses it for 5 minutes. It pumped about 1,000 litres. Is this hose faster than the first hose (300 litres per minute)? Give a reason for your answer.

my reason is that the first one is faster than the second because it pumped 1,000 litres and second it pumped 300 litres

d) L'agriculteur trouve un autre tuyau et l'utilise pendant 5 minutes. Il a pompé environ 1 000 litres. Ce tuyau est-il plus rapide que le premier (300 litres par minute)? Donnez une raison pour votre réponse.

Le premier est plus rapide The first is fastest.

(d) The farmer finds another hose and uses it for 5 minutes. It pumped about 1,000 litres. Is this hose faster than the first hose (300 litres per minute)? Give a reason for your answer.

they were in 5 minutes it 1000 Litres so you can in 2 minutes 300

d) L'agriculteur trouve un autre tuyau et l'utilise pendant 5 minutes. Il a pompé environ 1 000 litres. Ce tuyau est-il plus rapide que le premier (300 litres par minute)? Donnez une raison pour votre réponse.

Non, parce que le premier pompe 1500 litres en 5 minutes.

No, because the first pumps 1500 liter in 5 minutes.

DISCUSSION

Pour générer des évaluations de grande qualité qui ciblent les 21CS, il est important de s'assurer que les tâches et leurs items capturent des performances qui s'appuient sur les compétences ciblées. Puisque l'évaluation des compétences est intégrée dans un contenu disciplinaire spécifique, les apprenants risquent d'avoir des difficultés, soit parce qu'ils n'ont pas les connaissances pertinentes dans la discipline, soit parce qu'ils ne peuvent pas exercer les compétences requises.

Cette difficulté n'est pas toujours facile à adresser, mais il existe des solutions : d'abord, si les mêmes compétences sont requises dans plusieurs tâches qui s'appuient sur différents domaines disciplinaires, l'analyse des résultats de toutes les évaluations peut permettre d'identifier des modèles de réussite. Deuxièmement, si les tâches sont créées en utilisant un contenu de discipline dont on sait qu'il a été atteint par tous les élèves, alors les performances variables entre les élèves peuvent être attribuées à ce qui n'est pas connu, à savoir la différence de compétences.



Malgré ces solutions, il n'existe pas d'analyse des résultats à grande échelle pour l'élaboration par les enseignants de tâches d'évaluation destinées à être utilisées en classe. Cela veut dire que les concepteurs de tests et les enseignants locaux sont les mieux placés pour contrôler les niveaux de discipline qui sont supposés être familiers et bien compris par leurs élèves. Par conséquent, il y a moins de place pour une mauvaise interprétation des raisons pour lesquelles les élèves effectuent à différents niveaux des tâches qui font appel à la fois à la démonstration de compétences et à la connaissance du programme scolaire.

Les compétences du vingt-et-unième siècle comprennent un large éventail d'aptitudes, qui couvrent aussi bien les aspects cognitifs que sociaux de notre performance et comportement. Ces compétences varient dans la mesure où elles peuvent être intégrées dans l'éducation formelle et dans les pratiques de classe. Certaines de ces compétences sont généralement démontrées par des médias qui ne sont pas compatibles avec une saisie facile, contrairement à des compétences comme l'alphabétisation, qui sont relativement faciles à saisir aux moyens de l'écriture. L'OAA a repris les pratiques d'évaluation traditionnelles et les a adaptées aux exigences d'un petit nombre de compétences. Cette approche présente des limites, en particulier pour certaines des 21CS à caractère social. Les tâches de collaboration créées dans l'OAA étaient plus difficiles à administrer que les tâches de résolution de problèmes et de pensée critique, et ont causé une certaine confusion chez les élèves, en partie en raison d'une plus grande familiarité avec les pratiques traditionnelles de compétition qui ont caractérisé l'évaluation comme étant distincte de la collaboration intégrée dans les tâches d'évaluation des 21CS. La saisie de la mise en œuvre des processus, la saisie de la contribution de tous les membres du groupe et la saisie du produit sont toutes viables en classe. La nuance de la dynamique humaine qui contribue autant aux compétences sociales est cependant moins viable, en particulier dans les classes comprenant un grand nombre d'élèves.

Dans quelle mesure pouvons-nous mesurer de manière exhaustive les 21CS par des approches telles que celles décrites ici ? Et pourquoi voulons-nous les mesurer ?

Ce que décrit ce rapport est un début. Il est possible de définir et décrire les 21CS, et il est possible de créer des tâches pour des disciplines spécifiques qui provoquent et promeuvent ces compétences. La mesure des compétences fournit aux enseignants des informations sur les habiletés des apprenants qui peuvent être ensuite utilisées pour les entraîner et les développer davantage. Ce qu'il faut garder à l'esprit c'est que nous pratiquons et affichons ces compétences tous les jours, dans des domaines non technologiques ou intégrant la technologie. L'approche décrite dans ce rapport peut ne pas sembler passionnante pour un lecteur qui s'attend à ce que les nouveaux objectifs d'apprentissage soient assortis de modalités technologiquement complexes, mais les processus quotidiens entre humains sont tout aussi nécessaires à saisir et à évaluer pour s'améliorer que ceux qui sont numériques.

Les exemples de travail fournis dans les annexes du rapport illustrent des approches qui peuvent et doivent continuer d'être développées. L'approche suppose que des cadres de compétences ont été élaborés pour soutenir la création d'évaluations. Au-delà de l'utilisation de cadres tels que ceux décrits dans les annexes, les concepteurs de tests doivent acquérir des connaissances, puis élaborer des lignes directrices, sur la manière dont chaque 21CS d'intérêt passe du plus simple au plus complexe. Cela permettra de créer de meilleures évaluations qui ciblent des gammes de performance plus larges et permettra aux enseignants de mieux estimer les performances des élèves.



BIBLIOGRAPHIE

- Care, E., Vista, A., & Kim, H. (2020). *Optimizing Assessment for All: Focus on Asia*. Washington D.C.: The Brookings Institution.
<https://www.brookings.edu/research/optimizing-assessment-for-all-focus-on-asia/>
- Care, E., & Kim, H. (2020). *Optimizing Assessment for All: Framework for understanding project goals and scope*. Washington_D.C.: The Brookings Institution.
<https://www.brookings.edu/research/optimizing-assessment-for-all-framework-for-understanding-project-goals-and-scope/>
- Care, E., Griffin, P., & Wilson, M. (Eds.) (2018). *Assessment and Teaching of 21st Century Skills: Research and Applications*. Cham: Springer.
- Jones, S., Bailey, R., Brush, K., Nelson, B., & Barnes, S. (2016). *What is the same and what is different? Making sense of the "non-cognitive domain: Helping educators translate research into practice*. Easel Lab: Harvard.
https://easel.gse.harvard.edu/files/gse-easel-lab/files/words_matter_paper.pdf.
- Kim, H., & Care, E. (2020). *Optimizing Assessment for All: Classroom-based assessments of 21st century skills in the Democratic Republic of Congo, The Gambia, and Zambia*. Washington D.C.: The Brookings Institution.
<https://www.brookings.edu/research/optimizing-assessment-for-all-classroom-based-assessments-of-21st-century-skills-in-the-democratic-republic-of-congo-the-gambia-and-zambia/>



ANNEXES

L'objectif de ces annexes est de fournir un guidage pour développer des évaluations des compétences du 21^{ème} siècle. Des items du projet de l'OAA sont utilisés comme exemples pour expliquer chacune des étapes du processus de développement des tâches et items décrits dans ce rapport. Un ou deux exemples sont donnés pour chaque compétence : la pensée critique, la résolution de problèmes et la collaboration. Les modifications des versions des tâches associées à chacune des étapes sont indiquées en rouge. L'adaptation des tâches pour intégrer les compétences du 21^{ème} siècle repose sur l'adoption d'un cadre conceptuel ou structure de la compétence ciblée. Ces cadres conceptuels sont tirés des études de l'OAA Asie ou de l'OAA Afrique. Les structures de compétences utilisées comme cadres pour les exemples travaillés se trouvent au début de chaque annexe.

Les annexes A à E fournissent des exemples travaillés de révision des tâches. Chaque exemple reprend une tâche originale conçue pour être utilisée en classe dans le contexte des études d'une discipline – comme les mathématiques, les sciences sociales ou les sciences. L'exemple est affiné pour étendre sa capacité à saisir les compétences du 21^{ème} siècle. Un ensemble complexe de compétences est ciblé pour chaque tâche, bien que dans certains cas, la tâche comprenne aussi des aspects d'un second ensemble de compétences.

Les changements réels qui sont documentés pour les tâches sont accompagnés d'annotations qui indiquent clairement de quels changements il s'agit, et pourquoi ils ont été effectués.

Pour certains des exemples, les différentes versions dans la conception et le contenu des tâches sont organisées dans le cadre des processus (1) de rédaction des tâches, (2) de l'étape de réflexion à voix haute et (3) d'expertise. Un peaufinage supplémentaire des tâches est également effectué, si besoin, après le pilotage des tâches d'évaluation. Ces tâches ne sont pas incluses dans la présente publication.

- La plupart des tâches étaient initialement conçues dans une autre langue que l'anglais ; les versions ont été faites dans une combinaison d'anglais et de langues des pays participants. Dans certains cas, cela a conduit à des maladroites d'expression en anglais – leur traitement n'a pas fait l'objet du travail rapporté ici.
- Certaines des modifications apportées aux tâches consistent à fournir davantage d'échafaudages pour les testeurs. Ces changements ont pris la forme de phrases d'instruction plus courtes, plus spécifiques, et aussi de la mise à disposition d'espaces ou de structures de réponse pour les élèves. En principe, ces changements n'apportent pas d'aperçus nouveaux au lecteur, mais comme ces changements se sont produits dans les processus, nous en présentons quelques-uns ici.
- Les changements peuvent être classés comme suit : plus grande clarté et formatage ; modifications des stimuli pour stimuler l'ensemble de compétences ciblées ; modifications dans les items individuels d'une tâche pour cibler des sous-compétences spécifiques de l'ensemble de compétences ciblées ; développement de grilles pour fournir aux enseignants et aux élèves une compréhension plus approfondie des processus cognitifs et sociaux que les tâches sont censées stimuler.

ANNEXE	COMPÉTENCE	NOM	DOMAINE DISCIPLINAIRE
A	Pensée critique	Revenus du magasin	Mathématiques
B	Collaboration	Mélanges hétérogènes	Sciences
C	Résolution de problèmes	Rendement du riz	Social sciences
D	Collaboration	Serrekunda	Social sciences
E	Résolution de problèmes	Le citronnier	Sciences



Annexe A: Pensée critique – Revenus du magasin

RÉDACTION PRÉLIMINAIRE DES TÂCHES

La pensée critique a été définie pour décrire sa structure et ses composantes contributives et sous-compétences. Le processus a été éclairé par un résumé de la littérature de recherche disponible sur le sujet. Pour les besoins du projet OAA, la pensée critique a été définie comme l'analyse objective de l'information et la formulation d'un jugement raisonné par la synthèse de l'information en utilisant le raisonnement logique. La structure de l'OAA Asie a identifié quatre composantes, liées à des indicateurs comportementaux pour les sous-compétences associées (Tableau A1).

Tableau A1. Structure de l'OAA Asie pour la pensée critique

Composantes	Sous-compétences
Argumentation	Discuter les raisons
	Identifier des alternatives
	Prendre du recul
Gestion de l'information	Rassembler des informations
	Analyser l'information
	Synthétiser l'information
Raisonnement logique	Évaluer la cause et les effets
	Faire des hypothèses
Jugement	Faire des prédictions
	Inférer
	Comparer et contraster
	Évaluer les sources
	Justifier
	Faire des recommandations

Figure A1

Stimulus: Le tableau ci-dessous montre les revenus d'un magasin pour une semaine.

Jour	Gain total
Lundi	\$370
Mardi	\$410
Mercredi	\$325
Jeudi	\$360
Vendredi	\$450
Samedi	\$320
Dimanche	-

Q1: Combien a gagné le magasin le lundi ?

- A. \$410
- B. \$320
- C. \$370
- D. \$450

Réponse: C

Q2: Quel jour est celui avec les gains les plus élevés ?

Réponse: Vendredi

Une fois que la pensée critique a été définie, une tâche existante ayant le potentiel de viser la pensée critique a été identifiée. La Figure A1 montre cette tâche pour une classe de mathématiques. Dans sa forme actuelle, la tâche n'exige pas des élèves qu'ils mettent en marche des 21CS pour répondre. Par exemple, pour répondre à la Question 1 (Q1), les élèves doivent regarder le tableau du stimulus et trouver le montant total des gains indiqué pour lundi prochain. La réponse est donnée dans le tableau, et les étudiants n'ont besoin que de quelques connaissances en matière de maîtrise de l'information pour y accéder. Il en va de même pour Q2.





Figure A2

Stimulus: Un magasin propose un certain nombre de produits à la vente, notamment des jus, des gâteaux et des biscuits. Le tableau ci-dessous indique le nombre de produits vendus pendant une semaine (un jus coûte \$3, un gâteau coûte \$2, un biscuit coûte \$1).

Jour	Jus	Gâteaux	Biscuits	Gain total
Lundi	100	10	50	\$370
Mardi	95	15	95	\$410
Mercredi	80	20	45	\$325
Jeudi	80	20	45	\$360
Vendredi	100	25	100	\$450
Samedi	80	15	50	\$320
Dimanche	90	40	52	-

Q1: Combien le magasin a gagné lundi grâce aux gâteaux ?

- A. \$10
- B. \$20
- C. \$100
- D. \$200

[Composante: Gestion de l'information]

[Sous-compétence : Rassembler des informations]

Réponse : B

Q2: Le lundi, quel est le pourcentage du total de jus vendu en une semaine ?

- A. 6.25%
- B. 16%
- C. 62.5%
- D. 100%

[Composante : Gestion de l'information]

[Sous-compétence : Analyser l'information]

Réponse: B

Q3: Le dimanche, combien a gagné le magasin au total ?

[Composante : Gestion de l'information]

[Sous-compétence: Analyser l'information]

Réponse : 402

Q4: Si le magasin veut fermer un jour par semaine, quel est le meilleur choix pour le propriétaire ?

[Composante : Jugement]

[Sous-compétence : Évaluer la source, comparer et contraster]

Réponse : Samedi

Q5: Le mercredi, la réduction est de 20 % pour les biscuits et le dimanche de 20 % pour les gâteaux. Tu souhaites acheter un gâteau et un biscuit. Quel jour veux-tu y aller ?

[Composante : Jugement]

[Sous-compétence : Comparer et contraster]

Réponse : Dimanche

La Figure A2 montre comment les tâches peuvent être modifiées pour cibler les compétences de la pensée critique.



Premièrement, le stimulus a été modifié pour inclure des catégories de ventes (Jus, Gâteaux, Biscuits) dans le tableau et fournir des détails qui rendent l'information plus complexe. Deuxièmement, les questions existantes ont été modifiées de manière à ce qu'elles exigent d'identifier les sous-compétences pour y répondre. Troisièmement, de nouvelles questions ont été ajoutées. Enfin, chaque question comprend désormais des informations sur les composantes et sous-compétences spécifiques qui sont visées. Cela peut être utile aux enseignants pour identifier les processus face auxquels les élèves éprouvent des difficultés, afin de leur apporter le soutien nécessaire.

La Figure A2 Q1 est semblable à la Figure A1 Q1, mais les légères modifications apportées avec l'insertion de la proposition « grâce aux gâteaux » à la question, ainsi que la complexité ajoutée au stimulus, exigent de l'élève qu'il engage des sous-compétences de collecte d'informations ainsi que de gestion de l'information. Q1 comprend toujours des connaissances en mathématiques, mais l'élève doit en outre identifier ce qui est pertinent, afin de générer une réponse qui n'est pas explicite dans le tableau.

Q2 Q3 ressemblent à Q1 dans le sens où elles exigent de manipuler l'information trouvée dans le stimulus.

Q4 et Q5 incluent la composante Jugement en demandant aux élèves de faire des choix basés sur des comparaisons.



Q6: Explique tes réponses à Q5.

[Composante : Jugement]

[Sous-compétence : Justifier]

Q7: Un élève a \$10. Cite deux façons dont tu pourrais dépenser ton argent. Tu veux avoir au moins un article de chaque catégorie.

[Composante : Argumentation, gestion de l'information]

[Sous-compétence : Identifier des alternatives, analyser et synthétiser l'information]

Q6 demande aux élèves d'expliquer et justifier leur choix et met ainsi en jeu des compétences métacognitives.

Q7 exige des élèves qu'ils s'engagent dans les multiples sous-compétences énumérées.

**PENSÉE
À VOIX
HAUTE**

En raison de modifications significatives apportées à l'item original (Figure A1), les enseignants ont mené des activités de pensée à voix haute avec quelques élèves supplémentaires afin d'examiner si les items modifiés suscitent les composantes de la pensée critique et les sous-compétences qui ont été identifiées, ainsi que pour vérifier une fois de plus la pertinence des tâches pour le niveau scolaire et les domaines disciplinaires ciblés.

En général, les réactions aux activités de pensée à voix haute ont confirmé que les tâches visaient les composantes et sous-compétences prévues. Certains items, en particulier Q2 Q5, étaient trop difficiles pour les élèves de tous les niveaux. Dans certains cas, les élèves n'ont pas pu répondre correctement aux questions parce qu'un sujet particulier (par exemple, l'unité de pourcentage ou l'analyse des données de tableaux) n'avait pas encore été enseigné. Les élèves étaient peu familiarisés à certains formats de questions, comme Q6, où il leur était demandé de fournir des explications. Finalement, les élèves n'ont pas compris ce qui leur était demandé pour Q7.

Sur la base des preuves empiriques des activités de pensée à voix haute, la tâche a été révisée.

Figure A3

Stimulus: Un magasin propose un certain nombre de produits à la vente, notamment des jus, des gâteaux et des biscuits. Le tableau ci-dessous indique le nombre de produits vendus pendant une semaine (un jus coûte \$3, un gâteau coûte \$2, un biscuit coûte \$1).

Jour	Jus	Gâteaux	Biscuits	Gain total
Lundi	100	10	50	\$370
Mardi	95	15	95	\$410
Mercredi	80	20	45	\$325
Jeudi	80	20	45	\$360
Vendredi	100	25	100	\$450
Samedi	80	15	50	\$320
Dimanche	90	40	52	-

Dans cette version de la Figure A3, le temps alloué et le niveau de complexité (simple, moyen et difficile) pour chaque item du test ont été ajoutés comme guide général pour les enseignants. Des informations sur les réponses possibles ont été ajoutées ou développées pour les questions à réponse construite afin de tenir compte des multiples façons de répondre (par exemple, Q3 et Q4).



Q1: Combien le magasin a gagné lundi grâce aux gâteaux ?

- A. \$10
- B. \$20
- C. \$100
- D. \$200

[Composante : Gestion de l'information]

[Sous-compétence : Rassembler des informations]

[Complexité : Simple]

[Durée : 10 min]

Réponse : B

Q2: Le lundi, quel est le pourcentage du total de jus vendu en une semaine ?

- A. 6.25%
- B. 16%
- C. 62.5%
- D. 100%

[Composante : Gestion de l'information]

[Sous-compétence : Analyser l'information, synthétiser l'information]

[Complexité : Moyenne]

[Durée : 15 min]

Réponse : B

Q3: Le dimanche, combien a gagné le magasin au total ?

[Composante : Gestion de l'information]

[Sous-compétence : Analyser l'information]

[Complexité : Moyenne]

[Durée : 10 min]

Les réponses possibles sont :

- 402
- $(90 \times 3) + (40 \times 2) + (52 \times 1) = 402$
- Quatre cent deux

Q4: Si le magasin veut fermer un jour par semaine, quel est le meilleur choix pour le propriétaire ?

[Composante : Jugement]

[Sous-compétence : Évaluer les sources, comparer et contraster]

[Complexité : moyenne]

[Durée : 05 min]

Les réponses possibles sont :

- Samedi
- \$320
- Samedi, \$320

Q5a: Le mercredi, la réduction est de 20 % pour les biscuits et le dimanche de 20 % pour les gâteaux. Tu souhaites acheter un gâteau et un biscuit. Quel jour veux-tu y aller ?

[Composante : Jugement]

[Sous-compétence : Comparer et contraster]

[Complexité : Difficile]

[Durée : 15 min]

Réponse : Dimanche

Q5b: Explique ta réponse à Q5.

[Composante : Jugement]

[Sous-compétence : Justifier]

[Complexité : difficile]

[Temps: 10 min]

Réponses possibles :

- Acheter le dimanche ; dépenser
- moins d'argent que le mercredi car les gâteaux sont plus chers que les biscuits
- Mercredi ($\$2 + \$0.8 = \$2.8$) > Dimanche ($\$1.6 + \$1 = \2.6) ainsi, en achetant le dimanche on dépense moins d'argent
- Les élèves choisissent le dimanche
- parce qu'on dépense moins d'argent
- Les élèves ne choisissent pas le mercredi parce qu'on dépense plus d'argent

Q6: Un élève a \$10. Cite deux façons dont tu pourrais dépenser ton argent. Tu veux avoir au moins un article de chaque catégorie.

[Composante : Argumentation, gestion de l'information]

[Sous-compétence : Identifier des alternatives, analyser et synthétiser l'information]

[Complexité : Difficile]

[Durée : 10 min]

Réponses possibles :

- $1j + 1ca + 5co$
- $1j + 2ca + 3co$
- $1j + 3ca + 1co$
- $2j + 1ca + 2co$

Quelques modifications mineures ont été apportées au formatage. Par exemple, dans la Figure A2, Q5 et Q6 étaient des items de test distincts. Cependant, en raison de la dépendance entre items (les élèves doivent répondre correctement à Q5 pour pouvoir répondre à Q6), ces deux items ont été fusionnés pour en former un seul (Q5a et Q5b).





EXPERTISE

La tâche révisée a ensuite été passée en revue lors d'une séance d'expertise. Au cours de cette réunion, les membres du panel (par exemple, des experts d'une discipline, des experts en évaluation et des enseignants) ont discuté de chaque question et ont donné leur avis. Sur la base de ces réactions, la tâche a été révisée une dernière fois.

Figure A4

Stimulus: Un magasin propose un certain nombre de produits à la vente, notamment des jus, des gâteaux et des biscuits. Le tableau ci-dessous indique le nombre de produits vendus pendant une semaine (un jus coûte Z 3, un gâteau coûte Z 2, un biscuit coûte Z 1).

Jour	Jus	Gâteaux	Biscuits	Gain total
Lundi	100	10	50	Z370
Mardi	95	15	95	Z410
Mercredi	80	20	45	Z325
Jeudi	80	20	45	Z360
Vendredi	100	25	100	Z450
Samedi	80	15	50	Z320
Dimanche	90	40	52	-

Q1: Combien le magasin a gagné lundi grâce aux gâteaux ?

- A. Z10
- B. Z20
- C. Z145
- D. Z375

[Composante : Gestion de l'information]

[Sous-compétence : Rassembler des informations]

[Complexité : Simple]

Réponse : B

[Durée : 10 min]

Q2: Le lundi, quel est le pourcentage du total de jus vendu en une semaine ?

- A. 6.25%
- B. 16%
- C. 62.5%
- D. 100%

[Composante : Gestion de l'information]

[Sous-compétence : Analyser l'information, synthétiser l'information]

[Complexité : Moyenne]

[Durée : 15 min]

La réponse clé est « B »

Q3: Le dimanche, combien a gagné le magasin au total ?

[Composante : Gestion de l'information]

[Sous-compétence : Analyser l'information]

[Complexité : Moyenne]

[Durée : 10 min]

Les réponses possibles sont :

- 402
- $(90 \times 3) + (40 \times 2) + (52 \times 1) = 402$
- Quatre cent deux

Dans la Figure A4, chaque pays a adapté certains aspects de la tâche. Par exemple, l'unité monétaire du stimulus a été révisée avec la monnaie locale afin de rendre la tâche plus pertinente pour la vie quotidienne locale.

Les options de réponse C et D pour Q1 ont été modifiées pour refléter les erreurs les plus communes. Voici le raisonnement qui sous-tend les options:

- A. Utilisation incorrecte du nombre de gâteaux vendus au lieu des gains
- B. Bonne réponse
- C. Additionne incorrectement le nombre de gâteaux vendus pour la semaine
- D. Utilise incorrectement le total des gains pour le lundi.

Q2 et Q3 ont été supprimés. D'après les séances d'expertise, ces items n'ont pas été considérés comme appropriés pour le niveau scolaire, ni comme reflétant le concept ciblé.



Q4: Si le magasin veut fermer un jour par semaine, quel est le meilleur choix pour le propriétaire ? **Explique pourquoi tu choisis ce jour-là.**

[Composante : Jugement]

[Sous-compétence : Évaluer les sources, comparer et contraster]

[Complexité : Moyenne]

[Temps: 05 min]

Les réponses possibles sont :

- Samedi
- \$320
- Samedi, \$320

Notation pour Q4:

Note = 1 si l'explication est incorrecte ou aucune explication n'est donnée.

Note = 2 si la réponse et l'explication sont toutes les deux correctes.

Q5a) Les mercredi, il y a une réduction de moitié prix pour les biscuits. Les dimanche, il y a une réduction de moitié prix pour les gâteaux. Tu veux acheter un gâteau et un biscuit. Quel jour veux-tu y aller ? Explique ta réponse.

[Composante : Jugement]

[Sous-compétence : Comparer et contraster]

[Complexité : Difficile]

[Durée : 15 min]

La réponse est dimanche

Q5b) Explique ta réponse à Q3(a).

[Composante : Jugement]

[Sous-compétence : Justifier]

[Complexité : Difficile]

Réponses possibles

- Acheter le dimanche ; cela demande de dépenser moins d'argent que le mercredi car les gâteaux sont plus chers que les biscuits
- Mercredi ($22 + 20.50 = 22.50$) ce sera plus cher que le dimanche ($21.00 + 21.00 = 22.00$) alors, en achetant le dimanche, tu dépenseras moins.
- Les élèves choisissent le dimanche parce qu'on dépense moins d'argent

Notation pour Q5:

Note 0 = Pas de réponse fournie, ou le jour choisi est incorrect.

Note 1 = Jour correct (dimanche), mais aucune explication/raison n'est donnée, ou l'explication est incorrecte.

Note 2 = Jour correct (dimanche) et une explication/raison valable est donnée.

Q6: Tu veux acheter au moins un jus, au moins un gâteau et au moins un biscuit. Tu veux dépenser exactement \$10. Par exemple, 1 jus + 1 gâteau + 5 biscuits = \$10 au total. Donne deux autres façons de dépenser ton argent.

[Composante : Argumentation, gestion de l'information]

[Sous-compétence : Identifier des alternatives, analyser et synthétiser l'information]

[Complexité : Difficile]

[Durée : 10 min]

Notation pour Q6:

Crédit complet = Note 2

Deux des suivantes:

1 jus + 2 gâteaux + 3 biscuits

1 jus + 3 gâteaux + 1 biscuit

2 jus + 1 gâteau + 2 biscuits

Crédit partiel = Note 1

Deux combinaisons quelconques où au moins l'un des deux s'additionne à \$10 OU comprend un de chacun des aliments mais pas les deux. Par exemple:

- 1 jus + 1 gâteau + 1 biscuit (comprend un de chaque aliment)
- 2 jus + 2 gâteaux (s'élève à \$10)
- Les autres options sont sans note = Note 0

Q4 a été légèrement modifié pour inclure une phrase demandant aux élèves d'expliquer les raisons de leur choix afin d'engager plus explicitement leurs compétences en pensée critique. La grille de notation a été ajoutée afin de distinguer les différents niveaux de performance en fonction des réponses fournies.

Q5 a été modifié parce que les enseignants ont déclaré que les élèves comprenaient la question, mais avaient eu la mauvaise réponse en raison de la difficulté des opérations mathématiques utilisant des pourcentages. Afin de résoudre ce problème, le pourcentage de réduction a été ramené à la moitié du prix pour faciliter les opérations et pour dissocier les connaissances mathématiques des compétences en matière de pensée critique. En outre, des détails supplémentaires sur la grille et la notation ont été ajoutés afin de déterminer le niveau des réponses des élèves.

Pour Q6, d'après les données de la session d'expertise, des explications supplémentaires ont été recommandées car il était difficile pour les élèves de comprendre le format de réponse attendu. Ainsi, fournir un exemple de réponse dans le cadre de la question peut aider les apprenants à mieux comprendre ce qui est attendu. En outre, la grille de notation a été incluse pour donner un crédit partiel pour une conclusion correcte même si l'élève a donné une mauvaise réponse en raison d'une erreur de calcul.



Annexe B: Collaboration – Mélanges hétérogènes

RÉDACTION PRÉLIMINAIRE DES TÂCHES

La collaboration a été définie pour identifier ses composantes et ses sous-compétences. Pour le projet OAA, la collaboration a été définie comme un processus de travail en commun et de communication afin de négocier différentes perspectives et de prendre des décisions pour atteindre un objectif commun. Il y a collaboration lorsque la réalisation de l'objectif nécessite plus que ce que chaque individu est capable de gérer seul, il est donc nécessaire de mettre en commun les ressources avec d'autres. La collaboration a été définie pour identifier ses composantes et ses sous-compétences. Pour le projet OAA, la collaboration a été définie comme un processus de travail en commun et de communication afin de négocier différentes perspectives et de prendre des décisions pour atteindre un objectif commun. Il y a collaboration lorsque la réalisation de l'objectif nécessite plus que ce que chaque individu est capable de gérer seul, il est donc nécessaire de mettre en commun les ressources avec d'autres. Associés aux principales composantes, les indicateurs comportementaux (ou sous-compétences) ont été identifiés : interagir efficacement avec les autres et avoir des conversations significatives ; savoir quand il est approprié d'écouter ou de parler (régulation sociale) ; travailler efficacement au sein d'équipes diverses (par exemple, résolution de conflits ; gestion d'équipe) ; introduire de nouvelles idées et partager les ressources ; assumer une responsabilité partagée pour le travail en équipe ; et prendre du recul (Tableau B1).

Table B1. La structure de l'OAA Asie pour la collaboration

Composantes	Sous-compétences
Participation	Interagir avec les autres membres du groupe
	Montrer de la responsabilité
	Montrer de la flexibilité
Communication	Partager de l'information et des idées
	Écouter
	Répondre aux autres
	Exprimer ses propres émotions de manière adéquate
Négotiation	Identifier des conflits
	Construire des arguments
Adoption de perspective	Reconnaître les autres
	Donner un retour aux autres
	Adapter en fonction du récepteur
Prise de décisions	Attribuer des rôles/du travail
	Faire un plan
	Identifier des alternatives possibles

Figure B1

Stimulus: Travaillez en groupe pour découvrir comment séparer le mélange d'eau et de sel pendant le cours de sciences.

Une fois que la collaboration a été définie, une tâche existante qui avait le potentiel de cibler les compétences de collaboration a été identifiée. La Figure B1 montre un exemple de tâche utilisée dans une salle de classe. À première vue, cela peut ressembler à une tâche qui évalue la collaboration car les élèves sont invités à travailler en groupe. Cependant, pour répondre à cette question, les élèves n'ont pas besoin de collaborer – c'est-à-dire de participer, de reconnaître le point de vue des autres, de fournir un retour d'information aux autres, etc. Il est possible de trouver la réponse sans l'aide des autres.



Figure B2

Stimulus: Faites une solution d'eau et de sel dans un pot et faites un mélange d'eau et de sable dans un autre pot. Séparez ensuite le sel et le sable du mélange. Faites également une démonstration devant la classe et rédigez un rapport.

Préparation et attribution de rôles:

Au sein de chaque groupe, répartissez les rôles, apportez le matériel et l'équipement nécessaires à l'expérience et remplissez le Tableau 1.

Tableau 1

Élèves	Matériel	Objectif

[Composante : Prise de décisions]

[Sous-compétence : Attribuer des rôles/du travail]

[Niveau de difficulté : Moyen]

Q2: Faire un plan

Tous les élèves font un plan de l'expérience et remplissent la première colonne du Tableau 2.

Tableau 2.

Tâche 2	Tâche 3	
Plan de l'élève 1	Retour de l'élève 2	Retour de l'élève 3
Plan de l'élève 2	Retour de l'élève 1	Retour de l'élève 3
Plan de l'élève 3	Retour de l'élève 1	Retour de l'élève 2

[Composante : Prise de décisions]

[Sous-compétence : Faire un plan]

[Niveau de difficulté : Difficile]

Q3: Retour aux autres élèves sur leur plan

Vous devez tous donner un retour d'information aux autres membres de votre groupe dans la deuxième ou troisième colonne du Tableau 2.

[Composante : Prise de Perspective]

[Sous-compétence : Donner un retour aux autres; reconnaître les autres; adapter en fonction du récepteur]

[Niveau de difficulté : Moyen]

Q4: Présentez le meilleur plan du groupe aux autres.

Chacun des groupes sélectionnera son meilleur plan et partagera les informations relatives aux expériences. Chaque groupe fera un retour sur les plans des autres groupes.

[Composantes: Communication]

[Sous-compétences: Partager de l'information et des idées; écouter; répondre aux autres]

[Niveau de difficulté : Moyen]

La Figure B2 montre comment la tâche peut être modifiée afin de cibler les compétences de collaboration. 

La nature de la tâche a été modifiée, passant d'une question directe à une tâche basée sur un projet ou une performance. Des instructions sont fournies pour la tâche globale, avec plusieurs sous-tâches (appelées questions ou items), qui structurent l'exécution de la tâche.

La version Figure B2 de la tâche est plus complexe. Elle présente un échantillon d'activités multiples et de composantes et sous-compétences de collaboration. Par exemple, Q1 concerne la préparation et l'attribution des rôles pour l'exécution des instructions principales. Q1 est structuré à l'aide d'un tableau, dans lequel les élèves identifient chacun leur rôle unique. Cela permet aux élèves et aux enseignants de savoir clairement quelles sont les composantes et les sous-compétences visées. Ensuite, des items supplémentaires ont été ajoutés pour guider les élèves tout au long du processus. Enfin, plutôt qu'une question générique de collaboration, chaque item comprend maintenant les composantes et sous-compétences spécifiques qui sont ciblées, ainsi que le niveau de difficulté.



PENSÉE À VOIX HAUTE

En raison des modifications importantes apportées à la tâche initiale (Figure B1), les enseignants ont conduit des activités de pensée à voix haute avec les élèves afin d'examiner si la tâche modifiée avait permis d'acquérir les compétences de collaboration, les composantes et les sous-compétences prévues, et si elle était adaptée au niveau scolaire et aux discipline ciblés.

La plupart des retours sur les séances de pensée à voix haute ont confirmé que la tâche visait les composantes et sous-compétences prévues. Les niveaux de difficulté ont été jugés appropriés pour tous les items. Les commentaires des enseignants ont mentionné qu'il faudrait indiquer le temps prévu pour chaque item, afin de guider les enseignants. En outre, Q1 nécessitait davantage d'explications pour que les élèves comprennent ce qu'il faut faire. Dans Q2, il a été demandé de réviser les instructions pour aider les enseignants à expliquer la tâche.

Sur la base des preuves empiriques fournies par les séances de pensée à voix haute, la tâche a été révisée (voir la Figure B3).

Figure B3

Stimulus: Faites une solution d'eau et de sel dans un pot et faites un mélange d'eau et de sable dans un autre pot. Séparez ensuite le sel et le sable du mélange. Faites également une démonstration devant la classe et rédigez un rapport.

Q1: Préparation et attribution des rôles

Travaillez dans des groupes de 3. Au sein de chaque groupe, répartissez les rôles pour chacun des membres, apportez le matériel et l'équipement nécessaires à l'expérience et remplissez le Tableau 1.

Tableau 1.

Élèves	Matériel	Objectif
Prénom de l'élève 1		
Prénom de l'élève 2		
Prénom de l'élève 3		

[Composante : Prise de décisions, Sous-compétence: Attribuer des rôles/du travail]

[Niveau de difficulté : complexe]

Level of difficulty: medium

[Durée : 5 minutes]

Grille de notation

Catégorie 0	Catégorie 1	Catégorie 2	Catégorie 3
Aucune raison n'est fournie	Les raisons sont sans rapport avec la tâche ; elles ne démontrent aucune preuve de travail vers une compréhension commune ou un objectif commun.	Les raisons de l'attribution des rôles sont quelque peu liées à la tâche et montrent que l'on tente de travailler vers une compréhension ou un objectif commun, mais qu'autrement, elles sont hors tâche.	Les raisons de l'attribution des rôles sont liées à la tâche et montrent que l'on tente de travailler à la réalisation d'une compréhension ou d'un objectif commun.

Dans la version de la Figure B3, la langue des items est révisée afin de clarifier la manière de les aborder. Par exemple, des informations sur le nombre d'élèves dans chaque groupe ont été ajoutées à Q1.

En plus, la durée pour chaque item du test a été ajoutée. Le fait de connaître la durée de chaque tâche permet aux enseignants de planifier la structure de la séance et de laisser suffisamment de temps aux élèves. Des détails sur la notation de chaque item du test ont été ajoutés. Ces grilles fournissent aux enseignants des informations spécifiques sur ce qui est évalué, ainsi qu'une échelle de notation composée de descripteurs comprenant les niveaux de progression, de qualité et/ou de compétence. L'utilisation d'une grille de notation qui inclut les comportements observables des élèves permet aux enseignants de porter des jugements objectifs.

**Q2: Faire un plan**

Chaque élève fait un plan de l'expérience et complète la première colonne du Tableau 2.

Tableau 2.

Tâche 2	Tâche 3	
Plan de l'élève 1	Retour de l'élève 2	Retour de l'élève 3
Plan de l'élève 2	Retour de l'élève 1	Retour de l'élève 3
Plan de l'élève 3	Retour de l'élève 1	Retour de l'élève 2

[Composante: Prise de décisions, Sous-compétence: Faire un plan]

[Niveau de difficulté : Difficile]

[Durée : 10 minutes]

Grille de notation:

Catégorie 0	Catégorie 1	Catégorie 2	Catégorie 3
Le(s) point(s) positif(s) sont des affirmations du plan sans raison spécifique.	Le(s) point(s) positifs font référence au plan dans son ensemble.	Au moins un point positif attribué indiquant un résultat spécifique du plan correspondant.	Deux points positifs sont donnés en lien tous les deux avec des résultats spécifiques du plan correspondant.
Les points négatifs indiquent simplement que l'idée n'est pas bonne sans explication.	Les points négatifs indiquent des critiques larges et génériques de la solution qui peuvent s'appliquer à la plupart d'idées.	Au moins un point négatif est présenté qui décrit une contrainte de la vraie vie ou une limitation qui a lien à la solution correspondante.	Deux points négatifs sont présentés qui décrivent des contraintes de la vraie vie ou des limitations qui ont lien aux solutions correspondantes.

Q3: Retour aux autres élèves sur leur plan

Chaque élève doit fournir un retour aux autres membres de ton groupe dans la deuxième ou troisième colonne du Tableau 2.

[Composante: Prise de Perspective, Sous-compétence: Donner un retour aux autres; reconnaître les autres; adapter en fonction du récepteur]

Niveau de difficulté : Moyen]

[Durée : 15 minutes]

Grille de notation:

Catégorie 0	Catégorie 1	Catégorie 2	Catégorie 3
Le(s) point(s) positif(s) sont des affirmations du plan sans raison spécifique.	Le(s) point(s) positifs font référence au plan dans son ensemble.	Au moins un point positif attribué indiquant un résultat spécifique du plan correspondant.	Deux points positifs sont donnés en lien tous les deux avec des résultats spécifiques du plan correspondant.
Les points négatifs indiquent simplement que l'idée n'est pas bonne sans explication.	Les points négatifs indiquent des critiques larges et génériques de la solution qui peuvent s'appliquer à la plupart d'idées.	Au moins un point négatif est présenté qui décrit une contrainte de la vraie vie ou une limitation qui a lien à la solution correspondante.	Deux points négatifs sont présentés qui décrivent des contraintes de la vraie vie ou des limitations qui ont lien aux solutions correspondantes.

Q4: Présente le meilleur plan du groupe aux autres

- Chacun des groupes sélectionnera leur meilleur plan et partagera l'information des expériences.
- Chacun des groupes donnera un retour aux plans des autres groupes.

Retour des autres groupes

[Composante : Communication; Sous-compétences: Partager de l'information et des idées; écouter; répondre aux autres]

[Niveau de difficulté : Moyen]

[Durée : 20 minutes]

Grille de notation:

Catégorie 0	Catégorie 1	Catégorie 2	Catégorie 3
Le(s) point(s) positif(s) sont des affirmations du plan sans raison spécifique.	Le(s) point(s) positifs font référence au plan dans son ensemble.	Au moins un point positif attribué indiquant un résultat spécifique du plan correspondant.	Deux points positifs sont donnés en lien tous les deux avec des résultats spécifiques du plan correspondant.
Les points négatifs indiquent simplement que l'idée n'est pas bonne sans explication.	Les points négatifs indiquent des critiques larges et génériques de la solution qui peuvent s'appliquer à la plupart d'idées.	Au moins un point négatif est présenté qui décrit une contrainte de la vraie vie ou une limitation qui a lien à la solution correspondante.	Deux points négatifs sont présentés qui décrivent des contraintes de la vraie vie ou des limitations qui ont lien aux solutions correspondantes.

En Q4, le format de la tâche a été modifié pour inclure un espace permettant aux élèves de noter les commentaires des autres groupes afin d'indiquer clairement ce qui est nécessaire pour mener à bien la tâche.





EXPERTISE

La tâche révisée a ensuite été passée en revue lors d'une séance d'expertise afin de vérifier et d'améliorer certains points si nécessaire. Au cours de la séance, les membres du panel (par exemple, des experts des disciplines, des experts en évaluation et des enseignants) ont discuté de chaque question et ont donné des retours. Sur la base des commentaires recueillis lors de la réunion, la tâche a fait l'objet d'un dernier cycle de révision, comme le montre la Figure B4.

Figure B4

Stimulus: Dans un pot, il y a un mélange de sel et de sable (ou de toute autre matière non soluble similaire au sable). Votre tâche consiste à séparer le sel et le sable du mélange. Faites un plan au sein de votre groupe que vous pouvez tous faire ensemble, chaque membre du groupe ayant un rôle à jouer. Décrivez le plan que votre groupe ferait pour accomplir la tâche.

Q1: Faire un plan

Les élèves travaillent par groupes de trois. Chaque membre du groupe doit faire un plan de l'expérience qui permettra d'accomplir la tâche et de remplir la première colonne du Tableau 1.

Tableau 1.

Remplissez la première colonne uniquement pour cette tâche.

Tâche 1	Tâche 2	
Plan de l'élève 1	Retour de l'élève 2	Retour de l'élève 3
Plan de l'élève 2	Retour de l'élève 1	Retour de l'élève 3
Plan de l'élève 3	Retour de l'élève 1	Retour de l'élève 2

[Composante: Prise de décisions]

[Sous-compétence : Faire un plan]

[Niveau de difficulté : Difficile]

Grille de notation:

Catégorie 0	Catégorie 1	Catégorie 2	Catégorie 3
Aucun plan fourni.	Le plan n'adresse pas les trois exigences et n'a pas de relation avec le rôle.	Le plan n'adresse pas les trois exigences mais ce qui est présenté a une relation avec le rôle.	Le plan adresse les trois exigences et a une relation avec le rôle.

Q2: Retour aux autres élèves sur leur plan

Vous devez tous donner un retour d'information aux autres membres de votre groupe dans la deuxième ou troisième colonne du Tableau 1. Fournissez un ou plusieurs commentaires positifs et négatifs sur les projets des autres.

[Composante : Prise de Perspective, Sous-compétence: Donner un retour aux autres; reconnaître les autres; adapter en fonction du récepteur]
[Difficulté : Moyenne]

Scoring rubrics:

Catégorie 0	Catégorie 1	Catégorie 2	Catégorie 3
Le(s) point(s) positif(s) sont des affirmations du plan sans raison spécifique.	Le(s) point(s) positifs font référence au plan dans son ensemble.	Au moins un point positif attribué indiquant un résultat spécifique du plan correspondant.	Deux points positifs sont donnés en lien tous les deux avec des résultats spécifiques du plan correspondant.
Les points négatifs indiquent simplement que l'idée n'est pas bonne sans explication.	Les points négatifs indiquent des critiques larges et génériques de la solution qui peuvent s'appliquer à la plupart d'idées.	Au moins un point négatif est présenté qui décrit une contrainte de la vraie vie ou une limitation qui a lien à la solution correspondante.	Deux points négatifs sont présentés qui décrivent des contraintes de la vraie vie ou des limitations qui ont lien aux solutions correspondantes.

Afin d'accroître l'authenticité et donner aux groupes la possibilité de générer des idées créatives, le stimulus a été modifié pour séparer le mélange de sable et de sel, plutôt que l'eau et le sel.

En outre, l'ordre de Q1, Q2 et Q3 a été modifié, de sorte que l'enchaînement des tâches soit plus logique. Le tableau de Q1 a été modifié pour permettre les réponses à Q2 (c'est-à-dire le retour aux autres élèves).



Q3: Préparation et attribution des rôles

Discutez des plans élaborés par chaque membre et décidez du meilleur plan ou combinez les plans en un plan final.

Pour réaliser le meilleur plan qui a été décidé, attribuez les rôles de chaque membre, décrivez les matériaux et l'équipement nécessaires à l'expérience et remplissez le tableau ci-dessous.

Enfin, partagez ce plan avec la classe pour obtenir des retours des autres groupes. Exemple de rôles (assurez-vous que chaque membre a un rôle et que les rôles ne sont pas doublés) :

- Responsable d'amener les matériaux
- Responsable de séparer le sable de la solution

Élèves	Rôle	Objectif du rôle
Prénom de l'élève 1		
Prénom de l'élève 2		
Prénom de l'élève 3		

[Composante : Prise de décisions]

[Sous-compétence : Attribuer des rôles/du travail]

[Difficulté : Complexe]

Grille de notation:

Catégorie 0	Catégorie 1	Catégorie 2	Catégorie 3
Aucune raison fournie	Les raisons sont sans rapport avec la tâche ; elles ne démontrent aucune preuve de travail vers une compréhension commune ou un objectif commun Les rôles attribués sont incomplets et n'accomplissent pas la tâche.	Les raisons de l'attribution des rôles sont quelque peu liées à la tâche et montrent que l'on tente de travailler vers une compréhension ou un objectif commun, mais qu'autrement, elles sont hors tâche. Certains des rôles sont doublés ou se chevauchent.	Les raisons de l'attribution des rôles sont liées à la tâche et montrent que l'on tente de travailler à la réalisation d'une compréhension ou d'un objectif commun. Aucun des rôles n'est doublé ou se chevauche.

Q4: Présentez le meilleur plan du groupe aux autres

Chaque groupe fournira un retour au plan de l'autre groupe.

Chaque groupe sélectionnera le meilleur plan de tous les groupes et donnera une raison de son choix.

Retour aux autres groupes	Groupe 1	Groupe 2	Groupe 3	Groupe 4
Points positifs:				
Points négatifs:				
Plan choisi comme meilleur parmi les autres et raisons du choix :				

[Composantes : Communication]

[Sous-compétences : Partager de l'information et des idées; écouter; répondre aux autres]

[Difficulté : Moyenne]

Grille de notation:

Catégorie 0	Catégorie 1	Catégorie 2	Catégorie 3
Le(s) point(s) positif(s) sont des affirmations du plan sans raison spécifique.	Le(s) point(s) positifs font référence au plan dans son ensemble	Au moins un point positif attribué indiquant un résultat spécifique du plan correspondant.	Deux points positifs sont donnés en lien tous les deux avec des résultats spécifiques du plan correspondant.
Les points négatifs indiquent simplement que l'idée n'est pas bonne sans explication.	Les points négatifs indiquent des critiques larges et génériques de la solution qui peuvent s'appliquer à la plupart d'idées.	Au moins un point négatif est présenté qui décrit une contrainte de la vraie vie ou une limitation qui a lien à la solution correspondante.	Deux points négatifs sont présentés qui décrivent des contraintes de la vraie vie ou des limitations qui ont lien aux solutions correspondantes.

Retour du groupe

Choix du meilleur plan

Catégorie 0	Catégorie 1	Catégorie 2
Ni meilleur choix ni raison fournis.	Le plan choisi est incomplet et ne pourra pas accomplir la tâche ou ne le fera que partiellement. Par exemple, il ne fera que séparer le sable de la solution salée. Les raisons ne s'appuient pas sur des concepts scientifiques.	Le plan choisi est complet et accomplira aussi bien la séparation du sable de la solution salée que du sel de l'eau. Les raisons s'appuient correctement sur des concepts scientifiques.

En Q3, des détails ont été ajoutés dans les instructions afin de mieux guider les élèves dans le remplissage du tableau. La grille de notation a également été révisée.

La question Q4 a été élargie pour que les élèves comprennent clairement ce qu'ils sont censés faire. Dans la version précédente, il était demandé aux élèves de donner leur avis aux autres groupes afin de choisir le meilleur plan parmi les différents groupes. La tâche est toujours la même ; cependant, le tableau demande spécifiquement un retour positif et négatif pour chaque groupe. En plus, les groupes doivent maintenant collaborer pour élaborer le meilleur plan et fournir une raison ensemble. Cela exige des élèves qu'ils engagent de manière plus explicite leurs compétences de collaboration, y compris en partageant des idées, en écoutant, en répondant aux autres et en prenant des décisions pour parvenir à un consensus.

La grille de notation a été révisée afin de mieux correspondre aux attentes de la tâche. Deux grilles de notation distinctes sont incluses, l'une qui évalue le retour d'information aux groupes, et l'autre qui évalue la sélection du meilleur plan et les raisons pertinentes fournies par le groupe.



Annexe C: Collaboration (et résolution de problèmes) – Serrekunda

RÉDACTION PRÉLIMINAIRE DES TÂCHES

La collaboration a été définie pour identifier sa structure et ses correspondantes composantes et sous-compétences. Le processus a été informé par un sommaire de la littérature de recherche disponible sur le sujet. Les indicateurs de comportement observables pour la sous-compétence ont été identifiés, en lien avec les quatre composantes du concept de collaboration. La structure de la collaboration utilisée pour la révision des tâches par l'OAA Afrique est affiché dans le Tableau C1. Consultez l'Annexe D pour la structure de la résolution de problèmes.

Tableau C1. Structure de l'OAA Afrique pour la collaboration

Composante	Sous-compétence
Participation (P)	Prise de responsabilité (Pr)
	Partager (Pa)
	Se relayer (Sr)
	S'engager (En)
Communication (C)	Recevoir (Re)
	Exprimer (Ex)
Négotiation (N)	Faire des compromis (Co)
	Prendre du recul (Pre)
Prise de décisions (D)	Analyser (An)
	Évaluer (Ev)
	Planifier (Pl)

Figure C1

Stimulus: Dans ton pays, plusieurs familles vivent ensemble dans des maisons surpeuplées. Identifie les causes et les conséquences de ce phénomène, et les solutions que tu préconises pour résoudre ce problème de population.

La tâche était initialement conçue pour la résolution de problèmes.



Figure C2

Stimulus: Serrekunda est une ville surpeuplée. Les familles vivent dans de très petites maisons avec un nombre d'occupants de plus en plus nombreux, ce qui soulève des préoccupations sanitaires. Dans votre groupe, travaillez de manière autonome afin de déterminer les causes, les conséquences et ensuite proposez des solutions à ce problème. Échangez au sein de votre groupe.

- Imagine que tu es le/la maire de Serrekunda, quelles mesures mettrais-tu en place pour adresser ces préoccupations sanitaires ?
- Quels défis es-tu susceptible de rencontrer en adressant ce problème ?
- Comment tu comptes dépasser ces défis ?

[Collaboration : Participation, communication, négociation, et prise de décisions]

[Résolution de problèmes: Collecte d'informations, Planifier une solution]

La Figure C2 montre que bien que la tâche conserve ses caractéristiques de résolution de problèmes, elle est reformulée de manière à exiger un travail en groupe.



Les sous-compétences spécifiques de collaboration que la tâche cible sont identifiées. Cette itération de la tâche a été conçue pour rendre le problème plus réel en nommant la ville, et en précisant que la surpopulation était un problème communautaire plutôt que simplement familial. Pour cibler les réponses des élèves, le rôle du maire a été introduit.



PENSÉE À VOIX HAUTE

En raison des modifications importantes apportées à l'item original (Figure C1), les enseignants ont mené des activités de pensée à voix haute avec les élèves pour examiner si la tâche modifiée permettait d'acquérir les composantes des compétences et les sous-compétences prévues, ainsi que la pertinence de la tâche pour le niveau scolaire et les matières ciblés. Le retour d'information de ce processus a conduit à une nouvelle série de révisions.

Figure C3

Serrekunda est une ville surpeuplée. De très grandes familles comprenant des grands-parents, leurs enfants adultes avec leurs maris et épouses et leurs propres enfants, vivent dans de petites maisons d'une ou deux chambres.

Ensemble de Questions 1

Dans des groupes de trois élèves, discutez :

- a) Pourquoi des familles aussi nombreuses vivent-elles ensemble dans de petits espaces ?
- b) Quels problèmes peuvent survenir en raison du grand nombre de personnes vivant dans d'aussi petites maisons ?

Ensemble de Questions 2

Les questions suivantes devraient être distribuées dans le groupe, une question pour chaque élève :

- a) En tant que grand-parent, que penses-tu du fait que le/la maire propose que les familles élargies soient logées dans des maisons différentes ? Donne une réponse positive et une réponse négative possibles.
- b) En tant que mère de quatre enfants, que penses-tu du fait que le/la maire propose que les familles élargies soient logées dans des maisons différentes ? Donne une réponse positive et une réponse négative possibles.
- c) En tant qu'enfant, que penses-tu du fait que le/la maire propose que les familles élargies soient logées dans des maisons différentes ? Donne une réponse positive et une réponse négative possibles.

Ensemble de Questions 3

En prenant le rôle du/de la maire, comment le groupe planifie-t-il de relever les défis identifiés par chaque membre de la famille ?

Ensemble de Questions 1

- a) *Les réponses plausibles comprennent un manque de bâtiments, la pauvreté, la culture*
- b) *Les réponses plausibles comprennent le partage des responsabilités par la famille étendue*

Plus de détails ont été fournis dans le matériel de stimulation au début de la tâche.



En raison des multiples questions désormais intégrées dans la tâche, celle-ci a été restructurée en trois ensembles de questions pour faciliter l'administration en classe.

Afin de faciliter la collaboration, des instructions plus détaillées pour travailler ensemble ont été incluses.

Les modifications ont eu pour effet d'augmenter les composantes et les sous-compétences visées. Certaines réponses plausibles ont été fournies à titre d'illustration.



EXPERTISE

La tâche révisée a ensuite été passée en revue lors d'une séance d'expertise afin de vérifier et d'améliorer certains points si nécessaire. Au cours de la séance, les membres du panel (par exemple, des experts des disciplines, des experts en évaluation et des enseignants) ont discuté de chaque question et ont donné des retours. Sur la base des commentaires recueillis lors de la réunion, la tâche a fait l'objet d'un dernier cycle de révision (Figure C4).

Figure C4

Serrekunda est une ville surpeuplée. De très grandes familles comprenant des grands-parents, leurs enfants adultes avec leurs maris et épouses et leurs propres enfants, vivent dans de petites maisons d'une ou deux chambres. Un des problèmes auxquels ces familles font face est qu'il n'y a pas assez de charbon pour que chacun puisse l'utiliser selon ses besoins.

- Différents membres de la famille ont des besoins et des usages quotidiens différents du charbon.
- La mère a besoin de charbon pour faire à manger.
- La grand-mère a besoin de charbon pour se réchauffer pendant la journée afin de rester au chaud et surveiller sa petite-fille qui est encore un bébé.
- Le jeune élève a besoin de charbon pour avoir de la lumière pour l'étude.

Chacune de ces activités nécessite des quantités journalières de charbon différentes.

- 7 unités de charbon pour préparer à manger
- 8 unités de charbon pour avoir de la lumière
- 8 unités de charbon pour réchauffer

La famille n'a que 15 unités de charbon au total par jour.

Dans un groupe de trois élèves identifiez qui prendra le rôle de quel membre de la famille (la mère, la grand-mère, le jeune élève).

- Individuellement, établissez et justifiez quelle est la meilleure issue pour toi dans ton rôle attribué.
- En tant que groupe, discutez et décidez de la meilleure issue pour tous les membres de la famille. Expliquez votre réponse.

a) Ge>ConCom

0 = Pas de réponse ou réponse non pertinente

1 = Fournit un résultat pertinent pour soi-même

b) Ge>ConCom

0 = Pas de réponse ou réponse non pertinente

1 = Fournit une issue pertinente pour le groupe

P>Pa

0 = Moins de trois personnes partagent le meilleur choix pour eux-mêmes

1 = Les trois membres partagent le meilleur choix pour eux-mêmes

C>Ex

0 = Pas de contributions

expressives/contributions non pertinentes

1 = Au moins une contribution expressive pertinente

Prise de décisions (D)

0 = Ne se sont pas mis d'accord sur une meilleure issue pour tout le monde

1 = Se sont mis d'accord sur une meilleure issue pour tous

Ju>Ex

0 = Pas de réponse ou réponse non pertinente

1 = Donne une raison de pourquoi cette issue a été choisie

2 = Donne une raison et explique ce que cette décision implique pour chacun des membres de la famille

Les révisions apportées à la tâche reflètent la structure du « modèle de tâche de collaboration » (Kim & Care, 2020). Cela a été un changement significatif qui a permis de passer de se focaliser sur une tâche spécifique, à une restructuration d'une tâche selon une structure modèle. En retour, cela signifie un plus grand potentiel pour le développement de tâches multiples qui suivent la même structure et qui offrent donc la possibilité de collecter de multiples indicateurs du même domaine/sous-compétence à travers un ensemble de tâches.

Le premier grand changement apporté a été de rendre le scénario du problème plus précis. Cela a permis de créer un contexte plus solide dans lequel les élèves pouvaient générer des idées, envisager et comparer les options, justifier et expliquer, participer et partager, communiquer et prendre des décisions dans le cadre d'une collaboration.

Le deuxième changement majeur a consisté à prévoir une simplification du marquage et de la notation. Comme on peut le voir, les réponses b) prévoient une notation sur plusieurs composantes/sous-compétences.





Annexe D: Résolution de problèmes – Le citronnier

RÉDACTION PRÉLIMINAIRE DES TÂCHES

La résolution de problèmes a été définie pour identifier sa structure et ses correspondantes composantes et sous-compétences. Le processus a été informé par un sommaire de la littérature de recherche disponible sur le sujet. Les indicateurs de comportement observables pour la sous-compétence ont été identifiés, en lien avec les quatre composantes du concept de collaboration. La structure de résolution de problèmes utilisée pour la révision des tâches par l'OAA Afrique est affichée dans le Tableau D1.

Tableau D1. Structure de l'OAA Afrique pour la résolution de problèmes

Composantes des compétences	Sous-composantes	Processus
Collecte d'Informations (CI)	Poser des questions reliées au problème (Pq) Organiser l'information	Classer (Cla) Analyser (vérifier, distinguer, comparer) (Ana) Décrire (Dec)
Planification d'une Solution (PS)	Générer des idées, des options, des hypothèses (Ge)	Faire des hypothèses (Hyp) Considérer des options et les comparer (ConCom)
	Développer un plan (Dp)	Distinguer (Dis) Identifier des liens (Lie) Prédire (Pre)
Gestion de l'Information (GI)	Suivre un plan (Sp)	
	Comparer les résultats au plan (Cf)	Comparer les preuves avec les prévisions(Com) Vérifier le flux logique (Vfl)
	Justifier le processus (Ju)	Expliquer (Exp)
	Synthétiser (Sy)	Résumer (Res)

Figure D1

Stimulus: Dans le cadre d'un projet de sciences, Chabota et ses deux amis ont planté des citronniers. Chabota a planté le sien sous un grand arbre. Ses amis ont planté les leurs à l'endroit où des bûches de bois avaient été brûlées. Il a découvert que sa plante ne poussait pas aussi bien que celles de ses amis. Comment expliquer les différences de croissance entre le citronnier de Chabota et ceux de ses amis ?

La tâche montrée dans la Figure D1 a été dérivée d'un sujet de sciences, et adaptée à la résolution de problèmes.

Figure D2

Dans le cadre d'un projet de sciences, Chabota et ses deux amis ont planté des citronniers. Chabota a planté le sien sous un grand arbre. Ses amis ont planté les leurs à l'endroit où des bûches de bois avaient été brûlées. Il a découvert que sa plante ne poussait pas aussi bien que celles de ses amis.

a) Comment expliquer les différences de croissance entre le citronnier de Chabota et ceux de ses amis ? **Identifiez au moins deux facteurs et expliquez comment ces facteurs expliquent les différences.**

Dans la Figure D2, la question initiale a été donnée de manière plus structurée afin d'éclairer ce qui était attendu de l'élève.



b) Les amis de Chabota l'ont trouvé en train de se préparer à couper l'arbre qui était proche de son citronnier. Si vous étiez l'un des amis de Chabota, quelles recommandations lui donneriez-vous et pourquoi ?

- a)** *CI>Pq>Ana/Dec*
b) *PS>Ge/Dp*
GI>Exp

Une question supplémentaire a été ajoutée à la tâche, afin de stimuler les élèves à la génération d'idées sur les possibles facteurs ou causes.

Les composantes de la résolution de problèmes, les sous-compétences et sous-sous compétences pertinentes pour chaque processus ont été identifiées.

PENSÉE À VOIX HAUTE

En raison des modifications importantes apportées à l'item original (Figure D1), les enseignants ont mené des activités de pensée à voix haute avec les élèves pour examiner si la tâche modifiée permettait d'acquérir les composantes des compétences et les sous-compétences prévues, ainsi que la pertinence de la tâche pour le niveau scolaire et les matières ciblées. Comme résultat de l'activité, la tâche a été significativement modifiée est élargie comme cela est montré dans la Figure D3.

Figure D3

Dans un projet de sciences, Chabota a découvert que son citronnier ne poussait pas aussi bien que ceux de ses deux amis. Les arbres de ses amis se trouvaient dans un endroit où des bûches de bois avaient été brûlées alors que le sien était sous un arbre.

- a)** Comment expliquer les différences de croissance ? Identifiez au moins deux facteurs et expliquez comment ces facteurs expliquent les différences.
- b)** Les amis de Chabota l'ont trouvé en train de se préparer à couper l'arbre qui était proche de son citronnier. Si vous étiez l'un des amis de Chabota, quelles recommandations lui donneriez-vous et pourquoi ?
- c)** Suggère deux manières par lesquelles Chabota pourrait s'assurer que son arbre grandisse aussi bien que ceux de ses amis.

- a)** *CI>Pq>Ana/Dec*
0 = Pas de réponse/sans lien
1 = Présentation d'un facteur
2 = Présente des facteurs associés aux deux environnements
- b)** *PS>Ge/Dp*
GI>Exp
0 = Pas de recommandation donnée/sans lien
1 = Recommandation donnée sans justification
2 = Recommandation donnée avec une justification
- c)** *PS>Dp/Pre*
GI>Ju/Ex
0 = Pas de suggestion/sans lien
1 = Une suggestion
2 = Deux suggestions

La tâche a été légèrement reformulée pour plus de clarté, et les lieux de plantation n'ont pas été directement attribués aux trois amis. Cette mesure visait à aider les étudiants interrogés à se concentrer sur la situation actuelle, plutôt que sur les personnes qui ont pu commettre des erreurs dans la localisation initiale des meilleures zones de plantation.

La Question c) a été ajoutée, qui exige que l'étudiant interrogé tienne compte des conseils donnés et de la manière dont cela se manifesterait dans l'action – en cherchant à élaborer puis à suivre un plan.

Les composantes pour la résolution de problèmes, les sous-compétences et sous-sous-compétences pertinentes à chaque processus ont été identifiées pour la question c).

Des grilles de notation ont été ajoutées, qui portent sur les processus et les produits associés à chacune des questions. Il est à noter que ces grilles ne permettent pas d'évaluer la qualité des réponses, mais seulement de savoir si les processus et produits requis ont été mis en œuvre. Un autre niveau de grilles pourrait être ajouté pour discriminer les niveaux de qualité.



EXPERTISE

La tâche révisée a ensuite été passée en revue lors d'une séance d'expertise afin de vérifier et d'améliorer certains points si nécessaire. Au cours de la séance, les membres du panel (par exemple, des experts des disciplines, des experts en évaluation et des enseignants) ont discuté de chaque question et ont donné des retours. Sur la base des commentaires recueillis lors de la réunion, la tâche a fait l'objet d'un dernier cycle de révision comme le montre la Figure D4.

Figure D4

Dans un projet de sciences, Chabota a découvert que son citronnier ne poussait pas aussi bien que ceux de ses deux amis. Les arbres de ses amis se trouvaient dans un endroit où des bûches de bois avaient été brûlées alors que le sien était sous un arbre.

- a) Comment expliquer les différences de croissance ? Identifiez au moins deux facteurs et expliquez comment ces facteurs expliquent les différences.
- b) Les amis de Chabota l'ont trouvé en train de se préparer à couper l'arbre qui était proche de son citronnier. Si vous étiez l'un des amis de Chabota, quelles recommandations lui donneriez-vous et pourquoi ?
- c) Suggère deux manières par lesquelles Chabota pourrait s'assurer que son arbre grandisse aussi bien que ceux de ses amis.

a-i) CI>Pq>Ana

- 0 = *Aucun facteur n'est fourni*
- 1 = *Des facteurs sont fournis mais sans lien*
- 2 = *Un facteur pertinent est donné (par exemple, seulement la question de l'ombre ou des nutriments)*
- 3 = *Deux facteurs pertinents sont donnés*

a-ii) CI>Oi/Dec

- 0 = *Aucune explication n'est fournie*
- 1 = *L'explication n'est pas liée aux facteurs*
- 2 = *L'explication est fournie uniquement pour un des facteurs (par exemple, l'ombre d'un grand arbre empêche la lumière du soleil d'atteindre l'arbre de Chabota ; [alors que] les arbres de ses amis étaient à l'air libre et recevaient beaucoup de lumière du soleil ; le citronnier de Chabota est en concurrence avec les racines du grand arbre pour les nutriments ; [alors que] les citronniers de ses amis ont moins ou pas de concurrence ; les plantes des amis de Chabota contiennent de la potasse).*
- 3 = *Une explication est donnée pour les deux facteurs*

b-i) PS>Ge/Dp

- 0 = *Aucune recommandation n'est donnée*
- 1 = *Recommandation donnée mais sans lien*
- 2 = *Recommandation pertinente donnée*

b-ii) GI>Exp

- 0 = *Aucune suggestion n'est fournie*
- 1 = *La justification n'est pas pertinente*
- 2 = *La justification est reliée à la recommandation spécifiquement donnée.*

c) PS>Dp/Pre**GI>Ju/Ex**

- 0 = *Aucune suggestion /sans lien*
- 1 = *Une suggestion*
- 2 = *Deux suggestions*

Le stimulus et les questions sont inchangés dans la version D4.



Des modifications ont été apportées aux grilles de notation, et à la notation en elle-même. Ces modifications garantissaient que des crédits soient accordés aux élèves répondant à plusieurs parties de chaque élément, comme aux points a) et b).

Des exemples d'explications valables ont été fournies comme guides pour la notation des tâches et des items comme ceux-ci.

La quantité d'informations incluses pour la notation et l'évaluation, bien qu'utile pour l'élaboration des tests et des questions, peut dépasser les besoins de l'enseignant en termes d'identification des compétences de l'élève. Dans de tels cas, la notation et la correction peuvent être réduites aux niveaux de la compétence seule, des composantes ou des sous-compétences.