



International Baccalaureate®  
Baccalauréat International  
Bachillerato Internacional

# Principes et pratiques de l'évaluation de l'IB – Guide de l'évaluation destiné aux examinateurs en chef et aux examinateurs principaux

Le présent guide a pour but de fournir une vue d'ensemble des processus d'évaluation d'importance pour les examinateurs en chef et les examinateurs principaux. Ces derniers participent en général à toutes les étapes du processus d'examen. Si vous avez déjà probablement connaissance de la plupart des informations qui suivent, nous espérons néanmoins qu'en détaillant les responsabilités de chaque rôle, nous vous permettrons de comprendre l'importance de chaque étape de l'évaluation.

## Quelles sont vos différentes responsabilités ?

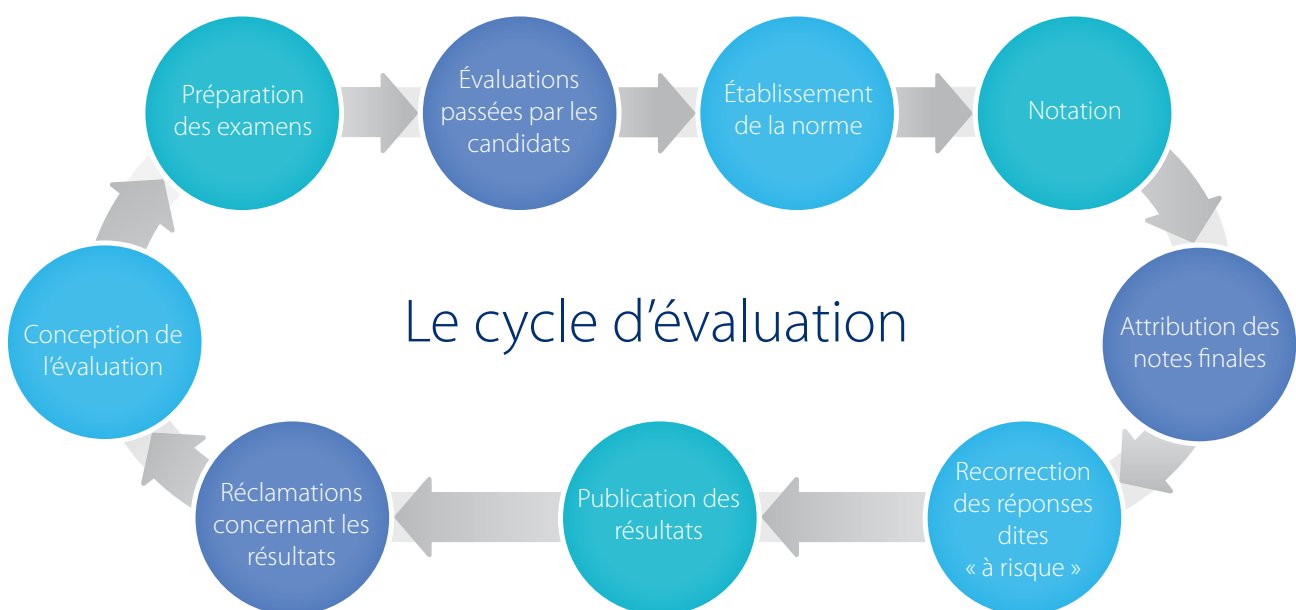
Les évaluations sommatives doivent offrir aux candidats des résultats significatifs, justes et fiables :

- **significatifs**, en ce que les évaluations doivent mesurer ce qui est censé l'être, tel qu'établi dans le guide pédagogique ;
- **justes**, en ce que les évaluations ne doivent pas favoriser des groupes de candidats en particulier ;
- **fiables**, en ce que les résultats des candidats ne doivent pas varier en fonction de la personne qui a noté leurs évaluations ou du moment où la correction a été effectuée.

Si une évaluation n'est pas significative et juste et si les résultats qui en découlent ne sont pas fiables, cette évaluation ne saurait être utilisée pour une évaluation sommative importante.

Le rôle de l'examinateur en chef est de superviser l'évaluation d'une matière, dont il est responsable, et de toujours garder en tête l'objectif de l'évaluation. Dans ce rôle, il doit également veiller à ce que les normes soient maintenues d'une année sur l'autre.

L'examinateur principal est responsable de la norme de notation d'une composante au sein d'une matière. Tous les examinateurs corrigeant une composante sont tenus de respecter la norme de l'examinateur principal de manière à garantir que tous les candidats obtiennent les bons résultats (en d'autres termes, les résultats d'un candidat ne doivent pas dépendre de l'examinateur qui a corrigé son travail). Le diagramme suivant montre les étapes du cycle d'évaluation, un circuit ininterrompu dont l'objectif est de contribuer à l'amélioration continue. En d'autres termes, des efforts continus sont consacrés à rendre les évaluations plus significatives, fiables et justes.



## Quels sont les éléments importants de la préparation des évaluations et des barèmes de notation ?



Les examinateurs en chef et les examinateurs principaux sont responsables de la préparation des examens pour chaque session. Dans ce cadre, ils bénéficient du soutien d'autres examinateurs expérimentés. Cette préparation peut commencer de 18 mois à 2 ans avant la session d'examens. Il convient de réfléchir à tous les examens et évaluations constituant l'ensemble requis pour la matière et de s'assurer qu'ils sont équilibrés et couvrent la gamme complète des objectifs spécifiques. L'examineur superviseur peut rédiger lui-même l'intégralité de l'examen ou compiler des questions soumises par d'autres examinateurs.

Au niveau de la création des évaluations, la question de leur caractère prévisible constitue un sujet important. Chaque examen doit imposer aux candidats de s'atteler à des questions et/ou des tâches qui sont d'une manière ou d'une autre différentes de ce qu'ils ont pu faire auparavant (afin d'éviter l'impact négatif d'un enseignement orienté en fonction de l'examen). Parallèlement à cela, l'examen doit être comparable aux examens précédents (pour permettre aux candidats de se préparer de manière efficace). L'IB cherche à mesurer les compétences cognitives supérieures. Dès lors, il est nécessaire de présenter des questions où les candidats doivent résoudre un problème ou réfléchir de manière créative pour appliquer leurs connaissances à un contexte nouveau, plutôt que de simplement employer des aptitudes bien affûtées ou réciter des connaissances.

Les questions et les barèmes de notation doivent être produits ensemble de manière qu'ils soient alignés dans ce qu'ils cherchent à évaluer. Le même effort doit être consacré à la préparation du barème de notation qu'à celle des questions pendant le processus d'élaboration de l'examen. Même pour les examens comportant des tâches ou des questions ouvertes notées selon les mêmes critères d'évaluation à chaque session d'examens, il est nécessaire de préparer des remarques sur la notation fournissant des instructions aux examinateurs sur la manière d'appliquer les critères dans le contexte de chaque question. Les barèmes de notation sont tout aussi importants que les examens pour la qualité du processus d'évaluation. Chaque barème de notation représente bien plus qu'une simple série de réponses types. En effet, il fournit des instructions sur la manière de noter les approches différentes que les candidats sont le plus communément susceptibles d'adopter dans leur réponse à une question ou sur la manière d'aborder les erreurs ou idées fausses les plus courantes chez les candidats.

Le barème de notation ou les remarques sur la notation doivent être clairs et réduire la possibilité que les examinateurs adoptent une approche de notation différente. Il s'agit du principal outil pour communiquer aux examinateurs ce qui est attendu d'eux.

## Comment communiquez-vous votre norme à vos examinateurs ?

La normalisation ne se limite pas à une seule réunion mais s'inscrit sur une période de temps dont les objectifs sont :

- pour l'examineur principal, de fixer la norme d'évaluation ;
- de tester et parfaire le barème de notation ;
- de produire des réponses comportant les notes définitives (réponses d'entraînement, de validation et de contrôle) ;
- de partager une compréhension avec tous les examinateurs ;
- de confirmer que les examinateurs comprennent la norme.

Il est important de comprendre les différents objectifs des trois types de réponses comportant les notes définitives. L'objectif des réponses d'entraînement est de permettre aux examinateurs de se familiariser avec la norme de notation. Les réponses d'entraînement ont pour but de prouver que les examinateurs sont capables de se conformer, dans leur notation, à la norme correcte. Les réponses de contrôle servent à démontrer que la notation des examinateurs reste conforme à la norme correcte.

Les réponses sélectionnées doivent répondre à une norme claire puisque leur raison d'être n'est pas de « coincer » les examinateurs mais de vérifier qu'ils appliquent de manière appropriée le barème de notation ou les critères et respectent la norme de notation de l'examineur principal. Il est important que les réponses sélectionnées contiennent un éventail de réponses et de notes de manière à vérifier que les examinateurs ont une compréhension large et approfondie du barème de notation ou des critères. Une telle compréhension est nécessaire pour reproduire la norme de notation de l'examineur principal.

L'étape d'établissement de la norme est réussie si les examinateurs ont une bonne compréhension de la norme et si l'IB dispose d'une série de réponses comportant les notes définitives (réponses d'entraînement, de validation et de contrôle) pour aider les examinateurs à apprendre la norme et à démontrer qu'ils la maîtrisent.

## Comment savez-vous que vos examinateurs comprennent votre norme ?

La fiabilité de la notation (c'est-à-dire la capacité d'un processus d'évaluation à fournir presque la même note à un travail, quel que soit l'examineur qui l'a corrigé et quel que soit le moment où la notation a été effectuée) joue un rôle essentiel pour garantir des résultats justes à tous les candidats. Dans chaque composante, l'examineur principal fixe la norme de notation, qu'il doit communiquer clairement pour que chaque examinateur puisse s'y conformer dans son travail de notation. Tout écart de cette norme entraînera une baisse de la fiabilité des résultats des candidats, ce qui signifie que l'évaluation n'aura pas atteint son objectif.

La méthode employée pour assurer la qualité de la notation des composantes corrigées en ligne est celle de la **notation de réponses de contrôle**. Dans ce processus, les examinateurs corrigent des réponses

ayant déjà été notées par l'examineur principal et l'équipe des examinateurs superviseurs. Puis, les deux séries de notes sont comparées. Si les différences entre les notes se situent dans une marge de tolérance prédéterminée, la notation est jugée acceptable. Une marge de tolérance est utilisée parce que l'on reconnaît que des différences légitimes peuvent exister entre des notes attribuées au même travail par des examinateurs différents. Il est important de limiter autant que possible ces différences pour que l'influence sur les résultats des candidats soit minime.

Les marges de tolérance varient en fonction des différentes tâches de manière à prendre en compte le fait qu'il est plus facile de convenir de notes objectives que de notes subjectives. Dans de nombreux cas, l'on peut raisonnablement s'attendre à ce que tous les examinateurs donnent exactement la même note (par exemple, si la réponse est clairement correcte ou non). Dans les réponses plus développées, telles que dans une dissertation, les examinateurs doivent s'accorder sur la qualité générale de la réponse mais peuvent néanmoins attribuer une note divergeant d'un point ou deux.

Pour assurer que les travaux des candidats sont notés avec précision et pour réduire le recours aux recorections à la fin de la période de notation, les examinateurs doivent corriger un certain nombre de réponses de validation et prouver qu'ils sont capables de se conformer à la norme requise avant de pouvoir noter des réponses « en conditions réelles ». Une fois que les examinateurs notent des réponses en conditions réelles, ils font l'objet d'un suivi permettant de vérifier qu'ils continuent de respecter la norme requise. À cette fin, des réponses de contrôle sont insérées de manière aléatoire parmi les réponses affectées aux examinateurs. Ceux-ci ne doivent pas savoir lesquelles sont des réponses de contrôle de manière qu'ils les corrigent de la même façon que toutes les autres réponses.

Dans chaque cas, les notes des examinateurs sont comparées à celles de l'examineur principal pour vérifier que les candidats reçoivent des résultats fiables. L'IB ne formule aucun avis sur ce que la note correcte devrait être.

## Fournir des commentaires aux examinateurs

Si, pendant la notation en conditions réelles, les notes de l'examineur diffèrent de manière conséquente des notes définitives/convenues, il est important que l'examineur reçoive un retour d'information immédiat expliquant comment et pourquoi les notes ont été attribuées. L'examineur reçoit un exemplaire de la réponse de contrôle où figurent les notes définitives et les annotations détaillées de l'équipe des examinateurs superviseurs. Si, malgré ces commentaires, la notation de l'examineur se situe toujours en dehors de la marge de tolérance, il fera l'objet d'une suspension temporaire de la notation jusqu'à ce qu'un examinateur superviseur ou un chef d'équipe se mette en rapport avec lui afin de discuter de la notation.

L'objectif de cette discussion est pour les chefs d'équipe de fournir des commentaires de qualité qui permettront aux examinateurs de comprendre comment rectifier leur notation. Les réponses de contrôle doivent donc être envisagées comme des activités de perfectionnement professionnel pour les examinateurs. La plupart des examinateurs recevront des commentaires sur des réponses de contrôle à un moment donné du processus de notation. Ces commentaires doivent donc être utilisés pour contribuer à garantir que les examinateurs respectent la norme de notation établie par l'examineur principal.

Lorsque le chef d'équipe considérera que l'examineur est prêt à reprendre la notation, il lui permettra d'accéder à de nouvelles réponses. Toutefois, si la notation de l'examineur sur les réponses de contrôle ne s'améliore pas, il sera à nouveau suspendu de la notation. Cette fois-ci, l'information sera

communiquée au personnel de l'IB et une décision sera prise pour déterminer si l'examinateur doit faire l'objet d'une suspension permanente de la notation afin d'empêcher que les candidats ne reçoivent des notes incorrectes.

La proportion d'examineurs dont les notes aux réponses de contrôle se situent dans la marge de tolérance donne une indication du niveau de fiabilité de la notation d'une composante donnée. Si une proportion relativement petite d'examineurs parvient à noter les réponses de contrôle dans la marge de tolérance, l'examinateur en chef et l'examinateur principal doivent déterminer la cause du faible niveau de fiabilité de la notation et la résoudre pour les sessions ultérieures. Par exemple, si un barème de notation n'explique pas de manière suffisamment claire comment les notes doivent être attribuées, les futurs barèmes de notation doivent être révisés de manière à fournir des instructions plus claires. Si la tâche d'évaluation elle-même ne permet pas une notation fiable, cela doit être pris en compte dans la révision du programme de manière que la tâche puisse être modifiée ou remplacée entièrement.

L'établissement des notes définitives des réponses de contrôle doit être fait avec le plus grand soin, non seulement parce qu'elles sont utilisées dans le processus d'assurance de la qualité, mais aussi parce que les examinateurs sont censés utiliser les notes et les annotations pour parfaire leur propre notation. Les conséquences d'une erreur dans une réponse comportant les notes définitives peuvent être nombreuses : des examinateurs sont injustement suspendus de la notation ; des examinateurs ne sont pas suspendus de la notation (entraînant des résultats injustes pour les candidats) ; des instructions incorrectes sont fournies aux examinateurs ; les candidats dont le travail comporte une note définitive reçoivent une note incorrecte.

## Groupes d'éléments de question

Certains examens seront notés par groupes d'éléments de question. Ces derniers peuvent correspondre à une partie ou à l'intégralité de la réponse d'un candidat. Si l'examen est décomposé en différentes parties, des examinateurs différents corrigeront les différentes sections. En décomposant l'examen en différentes parties, chaque partie étant un groupe d'éléments de question, les examinateurs peuvent choisir quel groupe d'éléments de question ils souhaitent corriger en fonction de leurs connaissances et de leurs préférences. La notation par groupes d'éléments de question est vérifiée au moyen de la notation de réponses de contrôle.

## Révision de notation de l'évaluation interne

Pour les travaux évalués en interne, la révision de notation constitue le principal outil pour garantir la fiabilité de la notation des enseignants. À la différence de la notation, la révision de notation vise à vérifier que les enseignants appliquent correctement le barème de notation tout en partant du principe que leur jugement est cohérent.

## Comment traiter des preuves contradictoires lors de l'attribution des notes finales ?

Du point de vue des candidats, la réunion de délibérations représente le point culminant du processus d'évaluation. L'attribution des notes finales permet de donner un sens aux notes attribuées aux candidats. Les notes peuvent varier d'une session à une autre puisque les questions ne sont pas les

mêmes. En revanche, la norme d'attribution des notes finales établie dans les sessions précédentes doit être maintenue. Par exemple, une note finale de 4 attribuée au cours de la session d'examens actuelle représente le même niveau qu'une note finale de 4 attribuée lors des deux sessions précédentes. Pour renforcer cette cohérence, l'équipe d'attribution des notes finales examine des réponses issues des sessions précédentes.

La réunion de délibérations a pour tâche première de réfléchir à la mise en oeuvre de chaque composante. Tous les participants auront été activement impliqués dans la notation d'au moins une composante et la plupart auront contribué à la rédaction des épreuves. Outre leur propre expérience, les examinateurs superviseurs doivent examiner les commentaires relatifs aux épreuves soumis formellement par les enseignants, les rapports des examinateurs sur leur perception des épreuves et les données statistiques issues de la notation des épreuves. Il s'agit là d'informations contextuelles importantes aidant à convenir des seuils d'attribution des notes finales appropriés.

Puis, l'équipe doit examiner chaque composante pour laquelle de nouveaux seuils d'attribution des notes finales doivent être fixés à chaque session. La modification des seuils doit en principe être mineure puisque tout doit être mis en oeuvre pour que chaque nouvelle version d'un examen soit élaborée au même niveau de difficulté générale que la version précédente. Il est de la responsabilité de l'examineur en chef de prendre en compte toutes les preuves disponibles (commentaires sur les résultats à l'examen, informations statistiques et jugements de l'équipe des examinateurs superviseurs) pour fixer le seuil recommandé qui permettra de maintenir au mieux la norme.

Les seuils des composantes évaluées en interne et des composantes évaluées en externe ne faisant pas l'objet d'examens ne sont pas modifiés à chaque session puisque les exigences de la tâche restent les mêmes, à moins que le cours n'ait changé. Dans ces cas, les seuils d'attribution des notes finales sont généralement fixés une fois au début du cours, bien qu'il soit possible de les réviser à chaque session d'examens.

Lorsque les résultats finals sont jugés justes et corrects, l'équipe des examinateurs superviseurs et les autres examinateurs expérimentés peuvent résoudre les questions en suspens ayant trait à la fiabilité de la notation. Le domaine principal de la recorection se concentrera sur les candidats dits « à risque ». En général, il s'agit de candidats dont le travail a été corrigé par des examinateurs dont la notation a été signalée comme pouvant potentiellement poser problème ou de candidats dont la note finale est de deux points ou plus en deçà de ce qui était prévu et qui n'ont besoin que de deux points de pourcentage pour obtenir une meilleure note finale dans la matière. Dans de tels cas, il est nécessaire de vérifier la justesse de la notation, étant donné qu'il existe toujours une marge d'erreur. Par conséquent, la révision des travaux de candidats dits « à risque » doit être effectuée par les correcteurs les plus fiables.

Toutes les discussions et décisions prises lors de la réunion doivent être consignées dans un rapport et utilisées pour contribuer à l'élaboration des évaluations de la prochaine session.

## Commentaires destinés aux établissements

Le processus d'évaluation vise à fournir des notes finales valides aux candidats et non à fournir des conseils aux établissements en matière d'enseignement. Néanmoins, l'IB s'engage à soutenir les établissements dans l'amélioration des normes. S'il n'est pas approprié de dépenser des ressources pour fournir des commentaires individuels, les rapports pédagogiques doivent comprendre des informations complètes sur les résultats généraux à la session d'examens. Cela a le double avantage de renforcer le soutien apporté par l'évaluation sommative à l'enseignement en classe et de permettre aux établissements, enseignants

et candidats qui utilisent ce système d'évaluation de mieux comprendre son fonctionnement.

Après chaque session d'examens, les examens et leurs barèmes de notation sont disponibles à la vente et les établissements peuvent ainsi se les procurer. Dans votre rôle d'examineur principal ou d'examineur en chef, il est de votre responsabilité de rédiger le rapport pédagogique en utilisant les informations qui vous sont fournies par l'équipe des examinateurs. Le rapport pédagogique couvre tous les aspects généraux du travail des candidats dans chaque composante, montre dans quelle partie les candidats ont bien travaillé ou se sont montrés moins compétents et fournit des recommandations visant à mieux préparer les candidats.

## Comité d'attribution des notes finales

Le comité d'attribution des notes finales se réunit après que toutes les réunions de délibérations ont eu lieu et juste avant la publication des résultats. Ce comité attribue de manière officielle les diplômes et certificats aux candidats ayant satisfait aux exigences. La réunion est présidée par le président du Bureau des examinateurs et comprend un petit nombre d'autres examinateurs en chef et de membres du personnel de l'IB. En outre, un observateur venant d'un établissement scolaire est invité à participer.

Le comité d'attribution des notes finales dispose d'un sous-comité sur l'intégrité intellectuelle. Ce dernier autorise la prise de mesures appropriées dans les cas de fraude présumée. Le sous-comité se compose d'examineurs superviseurs, d'enseignants superviseurs des écoles du monde de l'IB et de membres du personnel de l'IB. Chaque décision est prise par une équipe d'au moins trois personnes comprenant au moins un représentant de chaque groupe.

Le comité examine également les recommandations de politiques du personnel de l'IB sur les aménagements de la procédure d'évaluation pour les candidats ayant des besoins éducationnels spéciaux. Le comité traite aussi des cas de mauvaise administration touchant au non-respect des dates limites et/ou des procédures par les établissements. L'autorisation accordée à un établissement d'enseigner le Programme du diplôme peut lui être retirée lorsqu'il s'agit de cas graves qui impliquent une atteinte sérieuse à la sécurité et à l'intégrité des examens, ou lorsqu'il y a eu mauvaise administration durant plusieurs sessions.

## Réclamations concernant les résultats – Appliquez-vous la même norme de notation que lors de la session d'examens ?

Les services de réclamation concernant les résultats visent à permettre aux établissements et aux candidats de mettre en valeur leurs inquiétudes relatives à la notation des évaluations et à donner de la transparence sur la manière dont les notes sont attribuées. Il ne s'agit pas d'un mécanisme conçu pour octroyer de meilleurs résultats aux élèves, et les examinateurs en chef et les examinateurs principaux doivent diriger leurs équipes d'examineurs de manière à garantir que les mêmes normes sont appliquées dans le cadre des réclamations concernant les résultats que dans la période de notation initiale. Tous les candidats doivent être traités équitablement, qu'ils aient ou non soumis une réclamation concernant les résultats. Une question pertinente est de savoir si le travail à recorriger est d'une qualité suffisante pour mériter une note finale plus élevée.



Si des problèmes de notation sont identifiés dans le processus de réclamation concernant les résultats (c'est-à-dire si les notes sont modifiées), vous devez, dans votre rôle d'examinateur en chef ou d'examinateur principal, réfléchir aux façons de réduire ces divergences pour les sessions ultérieures et de renforcer la fiabilité de manière à assurer des résultats corrects pour tous les candidats.

## Comment améliorez-vous les évaluations en vue de la prochaine session ?

Il est important que nous tirions les enseignements de notre expérience au cours du cycle d'évaluation de manière à améliorer les futurs examens et les modèles que nous utilisons pour évaluer les candidats. Les examinateurs en chef et les examinateurs principaux doivent fournir des indications claires pour identifier ces améliorations et doivent soutenir l'IB pour les mettre en oeuvre.

Les modèles d'évaluation sont révisés dans le cadre du processus de révision du programme de chaque matière. En principe, les programmes d'études des matières sont révisés tous les sept ans par un groupe de révision composé d'enseignants, d'examineurs, de membres du personnel de l'IB et de conseillers spécialisés ne faisant pas partie de l'organisation. Durant le processus de révision des programmes, qui s'appuie sur des consultations, des propositions sont formulées puis envoyées aux établissements scolaires autorisés à enseigner le Programme du diplôme afin de recueillir leurs commentaires. Les recommandations et propositions des groupes chargés de la révision des programmes sont également envoyées au comité de révision du Programme du diplôme pour être examinées. Ce comité est responsable de la qualité globale des cours qui composent le Programme du diplôme et approuve les programmes d'études et modèles d'évaluation qui sont proposés. Il s'occupe particulièrement :

- du niveau scolaire et de la comparabilité des différents programmes d'études ;
- de réduire au maximum le chevauchement du contenu ou des objectifs spécifiques des matières, et d'encourager au contraire la complémentarité des programmes d'études ;
- de contrôler la charge de travail qu'impose l'évaluation aux élèves, aux enseignants et à l'IB, et ce afin de garantir que tous puissent y faire face ;
- d'éliminer toute redondance dans l'évaluation.

En ce qui concerne l'évaluation, les groupes chargés de la révision des programmes et le comité de révision du Programme du diplôme se réfèrent à la politique du Programme du diplôme en matière d'évaluation, qui présente les paramètres encadrant l'élaboration des modèles d'évaluation. Tout problème relatif aux évaluations actuelles doit être soulevé et il convient d'examiner les améliorations possibles des niveaux actuels de validité, de fiabilité et d'équité.