



International Baccalaureate®  
Baccalauréat International  
Bachillerato Internacional

# Principes et pratiques de l'évaluation – Une évaluation de qualité à l'ère du numérique



International Baccalaureate®  
Baccalauréat International  
Bachillerato Internacional

---

# **Principes et pratiques de l'évaluation – Une évaluation de qualité à l'ère du numérique**

## Principes et pratiques de l'évaluation – Une évaluation de qualité à l'ère du numérique

Version française de la publication parue originellement en anglais en mai 2026 sous le titre *Assessment principles and practices: Quality assessment in a digital age*

Publiée en mai 2026

Publiée pour le compte de l'Organisation du Baccalauréat International, fondation éducative à but non lucratif  
sise Rue du Pré-de-la-Bichette 1, 1202 Genève, Suisse.

Site Web : <https://ibo.org/fr/>

© Organisation du Baccalauréat International 2026

L'Organisation du Baccalauréat International (couramment appelée l'IB) propose quatre programmes d'éducation stimulants et de grande qualité à une communauté mondiale d'établissements scolaires, dans le but de bâtir un monde meilleur et plus paisible. Cette publication fait partie du matériel produit pour appuyer la mise en œuvre de ces programmes.

L'IB peut être amené à utiliser des sources variées dans ses travaux, mais vérifie toujours l'exactitude et l'authenticité des informations employées, en particulier dans le cas de sources participatives telles que Wikipédia. L'IB respecte les principes de la propriété intellectuelle et s'efforce toujours d'identifier les titulaires des droits relatifs à tout matériel protégé par le droit d'auteur et d'obtenir, avant publication, l'autorisation de réutiliser ce matériel. L'IB tient à remercier les titulaires de droits d'auteur qui ont autorisé la réutilisation du matériel apparaissant dans cette publication et s'engage à rectifier dans les meilleurs délais toute erreur ou omission.

La volonté d'inclusion de l'IB passe par les communications écrites. Ainsi, dans le respect de l'esprit international cher à l'IB, le français utilisé dans le présent document se veut standard et compréhensible par tout le monde, et non propre à une région particulière. De même, pour assurer une communication plus inclusive, l'IB recommande l'utilisation d'un langage non sexiste, qui représente une première étape vers l'adoption de l'écriture inclusive et non genrée.

Tous droits réservés. Aucune partie de cette publication ne peut être reproduite, mise en mémoire dans un système de recherche documentaire, ni transmise sous quelque forme ou par quelque procédé que ce soit, sans autorisation écrite préalable de l'IB ou sans que cela ne soit expressément autorisé par le [règlement de l'IB en matière d'utilisation de sa propriété intellectuelle](#).

Vous pouvez vous procurer les articles et les publications de l'IB sur le [magasin en ligne de l'IB](#) (adresse électronique : [sales@ibo.org](mailto:sales@ibo.org)). Toute utilisation commerciale des publications de l'IB (qu'elles soient commerciales ou comprises dans les droits et frais) par des tierces parties exerçant dans le milieu de l'IB, mais sans relation formelle avec lui (ce qui comprend notamment les organisations consacrées au tutorat, les prestataires de perfectionnement professionnel, les maisons d'édition spécialisées dans le domaine de l'éducation et les spécialistes de la planification de programmes d'études ou de la gestion de plateformes numériques contenant des ressources pédagogiques) est interdite et nécessite par conséquent l'obtention d'une licence écrite accordée par l'IB. Veuillez envoyer toute demande de licence à l'adresse [copyright@ibo.org](mailto:copyright@ibo.org). Des informations complémentaires sont disponibles sur le [site Web public de l'IB](#).

## Déclaration de mission de l'IB

Le Baccalauréat International (IB) a pour but de développer chez les jeunes la curiosité intellectuelle, les connaissances et la sensibilité nécessaires pour contribuer à bâtir un monde meilleur et plus paisible, dans un esprit d'entente mutuelle et de respect interculturel.

À cette fin, l'IB collabore avec des établissements scolaires, des gouvernements et des organisations internationales pour mettre au point des programmes d'éducation internationale stimulants et des méthodes d'évaluation rigoureuses.

Ces programmes encouragent les élèves de tout pays à apprendre activement tout au long de leur vie, à faire preuve de compassion et à comprendre que les autres, dans leurs différences, puissent aussi être dans le vrai.



## Profil de la communauté d'apprentissage de l'IB

Tous les programmes de l'IB ont pour but de former des personnes sensibles à la réalité internationale, conscientes des liens qui unissent les êtres humains, soucieuses de notre responsabilité partagée envers la planète et désireuses de contribuer à l'édification d'un monde meilleur et plus paisible.

En tant que membres de la communauté d'apprentissage de l'IB, nous nous efforçons de démontrer les aptitudes suivantes :

### ESPRIT DE RECHERCHE

Nous cultivons notre curiosité tout en développant des capacités d'investigation et de recherche. Nous savons apprendre indépendamment et en groupe. Nous apprenons avec enthousiasme et nous conservons notre plaisir d'apprendre tout au long de notre vie.

### CONNAISSANCE

Nous développons et utilisons une compréhension conceptuelle, en explorant la connaissance dans un ensemble de disciplines. Nous nous penchons sur des questions et des idées qui ont de l'importance à l'échelle locale et mondiale.

### RAISONNEMENT

Nous utilisons nos capacités de réflexion critique et créative, afin d'analyser des problèmes complexes et d'entreprendre des actions responsables. Nous prenons des décisions réfléchies et éthiques de notre propre initiative.

### COMMUNICATION

Nous nous exprimons avec assurance et créativité dans plus d'une langue et de différentes façons. Nous écoutons également les points de vue d'autres personnes et groupes, ce qui permet une collaboration efficace.

### INTÉGRITÉ

Nous adhérons à des principes d'intégrité et d'honnêteté, et possédons un sens profond de l'équité, de la justice et du respect de la dignité et des droits de la personne, partout dans le monde. Nous sommes responsables de nos actes et de leurs conséquences.

### OUVERTURE D'ESPRIT

Nous portons un regard critique sur nos propres cultures et expériences personnelles ainsi que sur les valeurs et traditions d'autrui. Nous recherchons et évaluons un éventail de points de vue et nous cherchons à en tirer des enrichissements.

### ALTRUISME

Nous faisons preuve d'empathie, de compassion et de respect. Nous accordons une grande importance à l'entraide et nous œuvrons concrètement à l'amélioration de l'existence d'autrui et du monde qui nous entoure.

### AUDACE

Nous abordons les incertitudes avec discernement et détermination. Nous travaillons de façon autonome et collaborative pour explorer de nouvelles idées et des stratégies innovantes. Nous faisons preuve d'ingéniosité et nous savons nous adapter aux défis et aux changements.

### ÉQUILIBRE

Nous accordons une importance équivalente aux différents aspects de nos vies – intellectuel, physique et affectif – dans l'atteinte de notre bien-être personnel et de celui des autres. Nous reconnaissons notre interdépendance avec les autres et le monde dans lequel nous vivons.

### REFLEXION

Nous abordons de manière réfléchie le monde qui nous entoure ainsi que nos propres idées et expériences. Nous nous efforçons de comprendre nos forces et nos faiblesses afin d'améliorer notre apprentissage et notre développement personnel.

Le profil de la communauté d'apprentissage de l'IB présente dix aptitudes mises en valeur par les écoles du monde de l'IB. Nous avons la conviction que ces aptitudes, et d'autres qui leur sont liées, peuvent nous aider à devenir des membres responsables au sein de nos communautés locales, nationales et mondiales.

# Table des matières

|   |            |
|---|------------|
| <b>Introduction et présentation</b>   | <b>1</b>   |
| Introduction  | 1          |
| Ressources complémentaires  | 3          |
| Comment bien utiliser cette ressource   | 5          |
| Principes et pratiques de l'évaluation  | 11         |
| Une bonne évaluation  | 12         |
| Le rôle de la technologie dans l'évaluation de l'IB                               | 22         |
| <b>Section A : les principes de l'évaluation</b>                                  | <b>26</b>  |
| L'évaluation dans l'éducation   | 26         |
| La validité   | 31         |
| Les éléments de la chaîne de validité   | 34         |
| Les normes  | 52         |
| La description de la réussite des élèves permise par l'évaluation sommative       | 56         |
| La notation des évaluations de l'IB   | 59         |
| Le renforcement d'un état d'esprit éthique  | 62         |
| <b>Section B : les pratiques de l'évaluation de l'IB</b>                          | <b>65</b>  |
| Définition des pratiques  | 65         |
| La transmission des résultats des élèves  | 66         |
| Les rôles et les responsabilités entourant le processus d'évaluation              | 74         |
| L'intégrité de l'évaluation   | 79         |
| Le caractère prioritaire de l'équité universelle                                  | 83         |
| Le cycle de vie de l'évaluation   | 88         |
| La production de l'évaluation   | 90         |
| La notation   | 97         |
| La révision de notation   | 107        |
| La révision de notation et les attentes du personnel enseignant                   | 109        |
| L'attribution des notes finales et la totalisation                                | 114        |
| Les vérifications de la qualité   | 123        |
| Le comité d'attribution des notes finales   | 124        |
| La préparation de la publication des résultats                                    | 126        |
| Les réclamations concernant les résultats et les demandes d'appel de l'évaluation | 129        |
| Les commentaires destinés aux établissements                                      | 135        |
| <b>Section C : les processus spécifiques aux programmes de l'IB</b>               | <b>137</b> |
| Définition des processus spécifiques aux programmes                               | 137        |
| Le socle commun à tous les programmes   | 138        |

|   |     |
|---|-----|
| Le profil de la communauté d'apprentissage de l'IB      | 140 |
| Des besoins et des solutions spécifiques aux programmes | 141 |
| Le Programme du diplôme                                 | 142 |
| Le Programme à orientation professionnelle              | 147 |
| Le Programme d'éducation intermédiaire                  | 151 |
| Le Programme primaire                                   | 155 |

---

|                |            |
|----------------|------------|
| <b>Annexes</b> | <b>158</b> |
|----------------|------------|

|   |     |
|---|-----|
| Gros plan sur la révision de notation de l'évaluation interne | 158 |
| Bibliographie   | 164 |
| Glossaire   | 167 |
| Ressources imprimables  | 179 |

# Introduction

Les programmes du Baccalauréat International (IB) sont proposés dans plus de 150 pays et régions (ou territoires), couvrant ainsi un éventail très varié de traditions et contextes éducatifs. Si certaines personnes connaissent très bien la philosophie de l'IB en matière d'évaluation et les méthodes associées, d'autres peuvent considérer qu'il s'agit d'un système complexe et inhabituel.

L'IB relève un véritable défi avec ses principes d'évaluation, en s'engageant à évaluer ce qui compte vraiment et non ce qui est facile à mesurer. Pour Alec Peterson, le premier directeur général de l'IB, ce défi prend la forme d'un équilibre exigeant qui s'inscrit dans l'écosystème éducatif.

Nous avons besoin d'un processus d'évaluation qui soit aussi valide que possible, dans le sens où il évalue réellement toute la capacité et la personnalité de l'élève par rapport à la prochaine étape de sa vie, mais qui soit en même temps suffisamment fiable pour que les élèves, les parents, [le corps enseignant] et les établissements d'accueil [aient l'assurance] que justice est faite. De plus, un tel processus ne doit pas, en raison de son effet de remous, fausser la qualité de l'enseignement, être trop lent ou absorber une trop grande quantité de nos ressources scolaires, qui sont peu abondantes.






(Peterson, 1971, p. 27 – 55)

La présente publication a pour objectif de démystifier les principes que l'IB met en œuvre afin de garantir que, dans tous ses programmes, l'évaluation est pertinente, juste et bénéfique aux élèves, et qu'elle est appliquée d'une manière authentique, appropriée et reflétant le processus d'apprentissage et d'enseignement.

Comme le montre la figure 1, l'évaluation telle qu'elle est conçue par l'IB se divise en deux catégories.

Figure 1

## L'évaluation sommative et l'évaluation formative

|  |  |
|--|--|
| <b>Le but</b>                            |  <b>Évaluation sommative</b> : prouver la maîtrise ou mesurer la capacité des élèves<br><b>Évaluation formative</b> : repérer les domaines à améliorer ou à privilégier (au niveau de la classe ou individuellement)  |
| <b>Le résultat</b>                       |  <b>Évaluation sommative</b> : un produit<br><b>Évaluation formative</b> : un processus   |
| <b>Le cadre temporel de l'évaluation</b> |  <b>Évaluation sommative</b> : l'aboutissement de l'apprentissage<br><b>Évaluation formative</b> : une pratique concomitante à l'apprentissage  |
| <b>Les moyens de preuve</b>              |  <b>Évaluation sommative</b> : dossiers de l'élève, examens et travaux réalisés dans le cadre des cours (activités généralement individuelles)<br><b>Évaluation formative</b> : observations, billets de sortie, discussion à deux et partage, et autoévaluation (activités généralement collectives) |
| <b>L'outil d'évaluation</b>              |  <b>Évaluation sommative</b> : critères et/ou bandes de notation<br><b>Évaluation formative</b> : grilles d'évaluation favorisant le développement et la progression des apprentissages   |

Il est fondamental que toute la communauté de l'IB comprenne que ces évaluations occupent des fonctions distinctes.

Les résultats que les élèves obtiennent aux évaluations ont une incidence sur leurs vies, car ils influent sur leur orientation vers des études supérieures ou une carrière professionnelle donnée. L'évaluation constitue un outil qui peut mener à des résultats aussi bien positifs que négatifs. Le personnel enseignant, les parents et les personnes exerçant une tutelle ainsi que les élèves doivent donc impérativement comprendre les points forts de l'évaluation, ses points faibles et les décisions qu'elle renferme. La présente publication aborde des questions et préoccupations courantes, telles que celles qui suivent.

- « Pourquoi l'IB intervient-il ainsi ? » – Comprendre les principes qui sous-tendent les pratiques d'évaluation de l'IB
- « Ce n'est pas juste. J'avais besoin d'une meilleure note finale. » – Expliquer l'équité et les effets des décisions en matière d'évaluation
- « Je n'avais jamais raisonné de cette manière avant. » – Analyser le processus d'évaluation
- « Je veux simplement me servir des évaluations pour mieux décider de mon enseignement. » – Souligner l'importance des évaluations de grande qualité à des fins tant formatives que sommatives

La technologie évolue en permanence. L'IB doit donc tenir compte de ces innovations pour garantir que ses évaluations continuent à être authentiques et le reflet fidèle de l'implication des élèves, de l'enseignement et de l'apprentissage. En incluant la formule « Des évaluations de qualité à l'ère du numérique », le titre de la présente ressource reconnaît que les évaluations de l'IB doivent se dérouler dans des conditions bien connues dans lesquelles les élèves se sentent à l'aise pour garder leur pertinence. Qui plus est, à l'heure de la transition, il est considéré comme impératif que l'IB reste fidèle aux principes fondamentaux sur lesquels il repose. Cela signifie que la technologie sert de soutien à l'apprentissage et à l'évaluation pour créer des évaluations encore plus valides. La formule « Des évaluations de qualité à l'ère du numérique » met en avant l'enthousiasme et la réactivité de l'IB en ce qui concerne à la fois le processus et le cadre dans lequel il évolue, tout en confirmant son engagement envers ses principes fondamentaux en matière d'évaluation.

La présente publication explique dans le détail les principes d'évaluation de l'IB mis en œuvre dans tous ses programmes : le Programme primaire (PP), le Programme d'éducation intermédiaire (PEI), le Programme du diplôme et le Programme à orientation professionnelle (POP). Ces principes couvrent tous les types d'évaluations de l'IB et expliquent pourquoi les évaluations sont conçues de la sorte. Certaines sections se concentrent spécifiquement sur l'évaluation sommative, notamment les évaluations envoyées à l'IB pour une notation externe, mais les informations fournies ont pour but d'améliorer les pratiques pédagogiques dans tous les programmes. Dans les contextes dépourvus d'évaluations sommatives externes, comprendre les éléments qui constituent une évaluation efficace et équitable peut améliorer les expériences en classe et mieux préparer les élèves aux futurs tests. La publication a aussi pour but d'expliquer les principes d'évaluation à l'ensemble de la communauté de l'IB, notamment au personnel enseignant, à la direction des établissements scolaires, aux examinateurs et examinatrices ainsi qu'aux responsables des admissions dans les universités et les établissements d'enseignement supérieur. Elle s'adresse aux personnes qui connaissent l'IB et ses évaluations, mais qui n'ont peut-être pas encore approfondi leur compréhension des théories sous-jacentes de l'évaluation. Nous espérons que cette ressource vous permettra de mieux comprendre comment les principes d'évaluation de l'IB soutiennent les objectifs pédagogiques plus larges de l'IB, qui sont de permettre aux élèves de vivre une expérience de haute qualité.

## Terminologie

Les évaluations de l'IB s'accompagnent d'un grand nombre de termes qui peuvent être méconnus du personnel enseignant et des élèves. Dans l'ensemble de ses publications, l'IB s'efforce d'éviter toute confusion, tout en affichant la plus grande précision possible.

Dans ce but, les abréviations sont expliquées la première fois qu'elles apparaissent dans une section de la présente publication. De la même manière, les termes importants sont expliqués la première fois qu'ils apparaissent, soit directement soit au moyen d'un lien vers une autre ressource.

Cette publication comporte également un glossaire qui rassemble tous les termes liés à l'évaluation. En dehors des termes généralement utilisés dans le cadre de l'évaluation, les termes spécifiques à chaque matière sont expliqués dans le guide pédagogique de l'IB correspondant.

## Ressources complémentaires

La présente publication vient rejoindre les autres ressources publiées par l'IB pour expliquer son approche de l'éducation. Il peut être compliqué de comprendre comment ces différentes ressources fondamentales sont liées les unes aux autres. Les tableaux 1 et 2 viennent donc récapituler les principales ressources et expliquer leur but et leur fonction.

Tableau 1

### Les principales publications de l'IB à utiliser pour les différents programmes

| Principales publications de l'IB communes à tous les programmes                                |   |
|--|---|
| <i>Principes et pratiques de l'évaluation – Une évaluation de qualité à l'ère du numérique</i> | Cette publication explique l'approche globale adoptée par l'IB en matière d'évaluation et les mesures prises pour la mettre en pratique. Elle se concentre sur l'évaluation sommative (les évaluations officielles) dans le système éducatif de l'IB.   |
| <i>En quoi consiste le système éducatif de l'IB ?</i>  | Cette publication a pour but de communiquer clairement ce qui est au cœur du système éducatif de l'IB. Elle détaille les idéaux qui sous-tendent tous les programmes de l'IB. Grâce à sa description de la philosophie éducationnelle de l'IB, elle sert également de guide aux établissements scolaires tout au long de leur parcours, qui commence dès leur adhésion aux principes de l'IB et couvre les nombreuses années qu'ils connaîtront en tant qu'écoles du monde de l'IB.   |
| <i>Normes de mise en œuvre des programmes et applications concrètes</i>                        | Cette publication fournit un ensemble de critères à l'aide desquels les établissements scolaires et l'IB peuvent évaluer le succès de la mise en œuvre des quatre programmes de l'IB. Elle contient les normes de mise en œuvre des programmes (les conditions générales à remplir par les établissements pour mettre en œuvre tout programme de l'IB), leurs applications concrètes (les descriptions plus détaillées des normes de mise en œuvre communes à tous les programmes) ainsi que les exigences (qui sont spécifiques à chaque programme). |

Tableau 2

### Les principales publications de l'IB spécifiques à un programme

| Principales publications de l'IB spécifiques à un programme  |  |
|--|--|
| <i>Le Programme primaire : des principes à la pratique</i><br><i>Le Programme d'éducation intermédiaire : des principes à la pratique</i><br><i>Le Programme du diplôme : des principes à la pratique</i><br><i>Le Programme à orientation professionnelle : des principes à la pratique</i> | Ces publications traitent de l'apprentissage et de l'enseignement dans le cadre d'un programme de l'IB en particulier. Elles expliquent également les exigences de ce programme. |

| Principales publications de l'IB spécifiques à un programme   |   |
|---|---|
| Guides pédagogiques (PEI et Programme du diplôme uniquement)  | Ils contiennent des informations détaillées sur une matière, un cours ou une discipline, comme la géographie ou les arts visuels. Les objectifs globaux, les objectifs spécifiques, le programme et les critères d'évaluation interne définis pour la matière en question y sont précisés. En règle générale, ces guides fournissent également des conseils supplémentaires pour l'apprentissage et l'enseignement spécifiques à la matière concernée.  |
| Continuums d'apprentissage propres aux disciplines (PP uniquement)  | Ces continuums sont des exemples dont les établissements scolaires peuvent s'inspirer pour concevoir leurs propres documents portant sur le contenu et l'enchaînement des programmes. Ils peuvent être adaptés librement pour répondre aux spécificités et aux exigences de la communauté locale. Ils donnent également des éléments de compréhension conceptuelle et des exemples de concepts pour chaque discipline et ses différents aspects.<br><br>Les descriptions générales des disciplines qui les accompagnent rendent transparents les éléments essentiels du PP en les inscrivant dans le contexte des différentes disciplines, et précisent le rôle des disciplines dans un programme transdisciplinaire. |
| <i>Procédures d'évaluation du Programme d'éducation intermédiaire</i><br><i>Procédures d'évaluation du Programme du diplôme</i><br><i>Procédures d'évaluation du Programme à orientation professionnelle</i>                    | Ces publications exposent l'ensemble des règles, règlements et processus spécifiques qui doivent être suivis pour procéder à l'évaluation d'un programme de l'IB en particulier. Elles incluent le règlement général qui regroupe les règles à suivre par les écoles du monde de l'IB.  |
| <i>Politique de préparation aux examens</i> (PEI, Programme du diplôme et POP uniquement)   | Cette publication présente la politique de l'IB en matière de préparation des examens, notamment pour ce qui est du stockage sécurisé du matériel confidentiel de l'IB dans les locaux de l'établissement et la modification du calendrier des examens.   |
| <i>Déroulement des examens</i> (Programme du diplôme uniquement ; le POP utilise la version pour le Programme du diplôme) et <i>Déroulement des examens sur ordinateur du Programme d'éducation intermédiaire de l'IB</i> (PEI) | Cette publication informe les personnes chargées de la coordination des programmes ou de la surveillance des examens des règlements relatifs à l'administration et à la conduite des examens du programme. Un exemplaire de cette publication doit se trouver dans chaque salle d'examen.   |

Outre ces publications de haut niveau, l'IB propose de nombreuses autres ressources qui fournissent des conseils précis au personnel enseignant, aux coordonnateurs et coordonnatrices ainsi qu'à ses autres parties prenantes. Ces documents sont disponibles sur le [site Web de l'IB](#) et sur le [Centre de ressources pédagogiques](#).

## Comment bien utiliser cette ressource

L'éventail des thèmes couverts dans la publication *Principes et pratiques de l'évaluation – Une évaluation de qualité à l'ère du numérique* est vaste. Celle-ci a été conçue pour proposer de multiples pistes qui serviront les intérêts de chaque lecteur et lectrice, et répondront à ses besoins. Les **questions thématiques** présentées dans cette section permettront d'en apprendre davantage sur les aspects qui comptent le plus pour chaque personne.

### Questions thématiques

La présente publication a pour objectif de donner un aperçu complet de la manière dont l'IB aborde l'évaluation. Beaucoup chercheront cependant des informations précises sur une partie du processus d'évaluation. C'est pourquoi les questions thématiques sont suivies de listes de titres de sections qui abordent des questions particulières que les lecteurs et lectrices peuvent se poser. Ces listes sont données à titre de références uniquement et n'imposent nullement l'ordre dans lequel il convient de lire les sections.

Le glossaire qui figure à la fin de la publication rassemble les termes présentant un intérêt particulier. Pour connaître le détail de tout terme du glossaire, une recherche pourra être effectuée dans la version PDF de cette publication.

#### « Quelle est l'approche de l'évaluation adoptée par l'IB ? »

**Public visé : le personnel chargé des admissions universitaires, le personnel enseignant et les parties prenantes qui souhaitent en savoir plus sur les bases théoriques des évaluations de l'IB, la définition donnée par l'IB des évaluations de qualité, ce qui guide ses décisions en matière d'établissement des processus et la manière de trouver un juste équilibre entre les exigences contradictoires de l'évaluation.**

Voici quelques exemples de questions courantes.

- En quoi consistent exactement les évaluations de l'IB ? En quoi sont-elles si particulières ?
- En quoi l'intelligence artificielle (IA) modifie-t-elle l'évaluation ?

Les sections qui suivent sont particulièrement pertinentes pour trouver une réponse à ces questions.

#### **Introduction et présentation**

- Une bonne évaluation
- Une bonne évaluation soutient les objectifs pédagogiques
- Une bonne prévisibilité
- Une bonne évaluation s'appuie sur tout un éventail de tâches d'évaluation
- Le rôle de l'évaluation se déroulant en classe et de l'évaluation interne
- Le travail en collaboration et les notes individuelles
- Une bonne évaluation tient compte des compétences plus larges de l'élève et des compétences de pensée de haut niveau
- Les compétences de pensée de haut niveau
- Le profil de la communauté d'apprentissage et le développement des compétences
- La sensibilité internationale et la compréhension interculturelle
- À quoi ressemble une bonne évaluation numérique ?

#### **Section A : les principes de l'évaluation**

- L'évaluation dans l'éducation
- Définition de l'évaluation
- Les approches de l'évaluation
- L'effet de remous et l'apprentissage
- La validité
- Définition de la validité
- La création d'un argument de validité
- Le maintien de la validité
- Les éléments de la chaîne de validité
- Trouver un équilibre entre les différents aspects de la validité
- La fiabilité
- Les résultats cohérents et les « bons » résultats
- La pertinence du construit et l'authenticité
- La facilité d'application
- L'équité et le biais
- La comparabilité
- L'approche de la validité adoptée par l'IB
- Les avantages de l'évaluation numérique pour la validité
- Les normes
- Les trois aspects des normes
- L'approche normative et l'approche critériée
- L'approche normative
- L'approche critériée
- L'approche comparée
- Le maintien des normes
- La description de la réussite des élèves permise par l'évaluation sommative
- Les effets des notes finales
- L'importance du jugement professionnel
- La notation des évaluations de l'IB
- Définition de la notation
- Les méthodes de notation
- La notation de l'évaluation formative

#### **Section B : les pratiques de l'évaluation de l'IB**

- La transmission des résultats des élèves
- La signification des notes finales de l'IB
- La distinction entre les notes et les notes finales
- Une session d'examens réussie
- Les notes finales et les accomplissements
- Définition de l'intégrité intellectuelle
- Le caractère prioritaire de l'équité universelle

#### **Ressources imprimables**

- « Les principes de l'IB en matière d'évaluation » (PDF)

## « Je suis spécialiste du PP, et nous ne procédons à aucune évaluation pour ce programme. En quoi cette publication peut-elle m'être utile ? »

**Public visé : les spécialistes du PP qui souhaitent comprendre les possibilités d'évaluation formative et sommative propres à l'établissement.**

Bien qu'aucune évaluation ne soit à envoyer à l'IB dans le cadre du PP, ce programme offre néanmoins des possibilités d'évaluation formative et sommative aux établissements et cette publication met en lumière les bonnes pratiques en la matière. Il est également important d'aborder dès la petite enfance, et de manière adaptée, les aspects réglementaires autour de l'intégrité intellectuelle et le renforcement d'un état d'esprit éthique.

Les sections qui suivent sont particulièrement pertinentes pour trouver une réponse à ces questions.

### **Section A : les principes de l'évaluation**

- Les approches de l'évaluation
- La notation de l'évaluation formative
- Le renforcement d'un état d'esprit éthique
- L'utilisation éthique de l'intelligence artificielle dans l'évaluation

### **Section B : les pratiques de l'évaluation de l'IB**

- Définition de l'intégrité intellectuelle
- Le caractère prioritaire de l'équité universelle

### **Section C : les processus spécifiques aux programmes de l'IB**

- Le socle commun à tous les programmes
- Le Programme primaire

## « Je suis spécialiste du PEI et nous procédons déjà à des évaluations numériques. En quoi cette publication peut-elle m'être utile ? »

Bien qu'une partie de cette publication traite de la transition vers l'évaluation numérique du Programme du diplôme et du Programme à orientation professionnelle, les sections en question confirment également la valeur réelle des évaluations électroniques du PEI. Cette publication s'applique donc au PEI. Veuillez vous référer à la question thématique suivante : « Je suis spécialiste du PEI, du Programme du diplôme ou du POP, et je souhaite comprendre comment sont attribuées les notes et les notes finales pour les évaluations envoyées à l'IB. Quel retour d'information est prévu pour les établissements ? »

## « Qu'est-ce que l'évaluation numérique va changer dans l'évaluation du Programme du diplôme ou du POP ? »

**Public visé : les enseignantes et enseignants connaissant bien le processus d'évaluation de l'IB, mais souhaitant savoir comment l'IB gèrera la transition vers les évaluations numériques.**

Voici quelques exemples de questions courantes.

- Qu'est-ce que l'évaluation numérique ?
- Les travaux seront-ils notés par un appareil informatique ?
- En quoi cela constitue-t-il une amélioration ?
- Les normes seront-elles les mêmes ?

Les sections qui suivent sont particulièrement pertinentes pour trouver une réponse à ces questions.

### **Introduction et présentation**

- Une bonne évaluation
- Le rôle de la technologie dans l'évaluation de l'IB
- À quoi ressemble une bonne évaluation numérique ?

- Les risques associés à l'évaluation numérique

### **Section B : les pratiques de l'évaluation de l'IB**

- Le caractère prioritaire de l'équité universelle
- Le cycle de vie de l'évaluation
- Les effets de l'évaluation numérique sur le cycle d'évaluation

## **« Je suis spécialiste du PEI, du Programme du diplôme ou du POP, et je souhaite comprendre comment sont attribuées les notes et les notes finales pour les évaluations envoyées à l'IB. Quel retour d'information est prévu pour les établissements ? »**

**Public visé : les enseignantes et enseignants souhaitant découvrir ce qui se passe après avoir envoyé les travaux des élèves à l'IB pour qu'ils soient notés, et déchiffrer le fonctionnement du système d'attribution des notes finales. Ces personnes peuvent aussi chercher à comprendre pourquoi l'IB inclut les évaluations du personnel enseignant dans ses processus, mais n'accepte pas toujours les jugements émis.**

Voici quelques exemples de questions courantes.

- Pourquoi les seuils d'attribution des notes finales sont-ils différents cette année ?
- Pourquoi est-il impossible de changer d'examineur ou d'examinatrice ?
- Qui note les copies ?
- Comment la notation des examinateurs et examinatrices est-elle vérifiée ?
- Comment puis-je déposer un recours ?
- J'aurais attribué une autre note à l'élève. En quoi cela est-il juste ?
- Le barème de notation indique que la réponse de l'élève vaut un point. En quoi cela peut-il être juste ?
- La meilleure solution pour les élèves ne serait-elle pas d'avoir une « salle de classe sans note finale » ?
- Pourquoi mes notes sont-elles modifiées ?
- Pourquoi le facteur de révision de notation change-t-il chaque année ?

Les sections qui suivent sont particulièrement pertinentes pour trouver une réponse à ces questions.

### **Introduction et présentation**

- Le rôle de l'évaluation se déroulant en classe et de l'évaluation interne

### **Section A : les principes de l'évaluation**

- L'évaluation dans l'éducation
- Définition de l'évaluation
- Les approches de l'évaluation
- L'effet de remous et l'apprentissage
- Les éléments de la chaîne de validité
- Les effets des notes finales
- La notation des évaluations de l'IB

### **Section B : les pratiques de l'évaluation de l'IB**

- La transmission des résultats des élèves
- La signification des notes finales de l'IB
- La distinction entre les notes et les notes finales
- Une session d'examens réussie
- Les notes finales et les accomplissements
- Les notes finales prévues
- Les rôles et les responsabilités entourant le processus d'évaluation

- Les examinateurs principaux ou en chef et examinatrices principales ou en chef
- Les autres examinateurs et examinatrices
- Les responsabilités du personnel de l'IB
- L'organisation hiérarchique des examinateurs et examinatrices
- Le caractère prioritaire de l'équité universelle
- Les notes et les notes finales
- Les approches de la notation
- Les barèmes de notation analytiques
- Les critères globaux : les bandes de notation
- La normalisation
- Le modèle d'assurance de la qualité
- L'entraînement
- La validation
- La notation en conditions réelles et notation des réponses de contrôle
- Les indices d'une normalisation réussie
- Les marges de tolérance
- Les copies difficiles ou inhabituelles
- Les liens avec un établissement scolaire
- Les groupes d'éléments de question
- La totalisation
- La révision de notation
- La révision de notation et les attentes du personnel enseignant
- La sélection des travaux d'élèves
- Les situations irrégulières
- L'impossibilité de trouver un facteur de révision de notation
- L'échantillonnage dynamique
- L'attribution des notes finales et la totalisation
- Les seuils d'attribution des notes finales fondés sur une appréciation et calculés par interpolation
- Les preuves utilisées dans le processus d'attribution des notes finales
- L'examen de la cohorte de l'année en cours
- Le retour d'information sur l'évaluation
- La recherche de preuves dans les copies
- L'analyse des statistiques sur les résultats
- Trouver un juste équilibre entre toutes les preuves
- Les seuils d'attribution des notes finales fixes
- L'étude et l'approbation des recommandations relatives aux seuils d'attribution des notes finales
- L'octroi du diplôme ou du certificat du programme
- Les observateurs et observatrices
- Les principes du processus d'attribution des notes finales
- Les vérifications de la qualité
- Les commentaires destinés aux établissements
- Les commentaires sur l'évaluation interne

### **Section C : les processus spécifiques aux programmes de l'IB**

- Le socle commun à tous les programmes

## « Comment l'IB conçoit-il et développe-t-il les examens ? »

**Public visé : les enseignantes et enseignants souhaitant comprendre le mécanisme d'élaboration des examens ou ayant des inquiétudes quant à leur qualité.**

Voici quelques exemples de questions courantes.

- Qui rédige les examens ?
- Comment ce travail est-il vérifié ?
- Les examens préparés dans différentes langues sont-ils identiques ?
- Testent-ils réellement ce qu'il faut ?

Les sections qui suivent sont particulièrement pertinentes pour trouver une réponse à ces questions.

### **Introduction et présentation**

- Une bonne évaluation
- Une bonne évaluation soutient les objectifs pédagogiques
- Une bonne prévisibilité
- Une bonne évaluation s'appuie sur tout un éventail de tâches d'évaluation
- Le rôle de l'évaluation se déroulant en classe et de l'évaluation interne

### **Section A : les principes de l'évaluation**

- L'évaluation dans l'éducation
- Définition de l'évaluation
- Les approches de l'évaluation
- L'effet de remous et l'apprentissage
- La validité
- Définition de la validité

### **Section B : les pratiques de l'évaluation de l'IB**

- Les rôles se rapportant à la rédaction des examens
- Le caractère prioritaire de l'équité universelle
- La production de l'évaluation
- Présentation de la production des examens
- Le processus conduisant à la validation du contenu
- La rédaction et la validation du contenu
- L'élaboration et la relecture
- La validation de la facilité d'utilisation
- Le contrôle de la qualité
- Les traductions
- Les épreuves modifiées
- Les notes et les notes finales

# Principes et pratiques de l'évaluation

Les **principes de l'évaluation** sont les propositions de base jugées importantes pour la conception, la mise en œuvre, la correction et la notation des évaluations menant aux diplômes et certificats. Ils découlent des éléments essentiels du système éducatif de l'IB. Le principe prépondérant est que les évaluations doivent être au service de l'éducation, et non la fausser.

Les **pratiques de l'évaluation** sont les transpositions efficaces et concrètes des principes de l'IB. Elles tiennent compte des exigences contradictoires et des limites pratiques posées par une utilisation à travers le monde, tout en respectant l'intégrité, une idée phare de la philosophie de l'IB.

Les principes de l'IB en matière d'évaluation sont résumés ci-après.

Qu'elle soit sommative ou formative, l'évaluation doit respecter un certain nombre de conditions.

- Être valide pour atteindre les buts visés. En d'autres termes, l'évaluation doit trouver un équilibre entre les exigences contradictoires posées par la « pertinence du construit » (c'est-à-dire le fait d'évaluer les bonnes choses), la fiabilité, l'équité (c'est-à-dire l'absence de biais), la comparabilité avec d'autres évaluations et la facilité d'application pour les élèves, les établissements scolaires et l'IB.
- Avoir un effet de remous positif, c'est-à-dire compter sur une conception favorisant un apprentissage et un enseignement de qualité.
- Convenir au plus grand nombre possible d'élèves, en leur permettant de montrer le niveau personnellement atteint.
- Disposer d'une conception propice aux aménagements pour garantir l'accès pour la totalité des élèves.
- Faire partie intégrante des programmes de l'IB, et non être considérée comme un élément isolé.
- Favoriser la simultanéité des apprentissages et soutenir l'expérience globale des élèves.
- Soutenir la mission de l'IB et développer les aptitudes suivantes chez les élèves : l'esprit de recherche, la connaissance, le raisonnement, la communication et l'ouverture d'esprit.

Les pratiques de l'évaluation sont habituellement décrites de manière très globale et elles doivent être mises en œuvre de façon adéquate, en fonction du contexte propre à chaque matière, discipline et programme. Elles expliquent de façon générale ce qui doit être fait, sans donner de détails précis sur la marche à suivre au quotidien ni sur les applications différentes en fonction des matières et des programmes concernés.

La ressource imprimable intitulée « [Les principes de l'IB en matière d'évaluation](#) » (PDF) est mise à la disposition du personnel enseignant.

## Une bonne évaluation

- Il n'existe pas de définition unique d'une « bonne évaluation » : tout dépend de l'importance accordée aux différentes priorités et du but de l'évaluation.
- La démarche de l'IB repose sur le principe qu'il faut évaluer ce qui est important et non considérer comme important ce qui peut être évalué.
- Les évaluations de l'IB cherchent à avoir un effet de remous positif sur l'apprentissage et sur l'enseignement.
- En règle générale, les bonnes évaluations doivent inclure un éventail de tâches mesurant les accomplissements et aidant à prouver le profil équilibré de l'apprentissage des élèves, et permettre d'organiser des activités plus poussées en classe ainsi que des examens.
- Pour inclure l'ensemble des élèves en couvrant toute la variété des styles et préférences d'apprentissage, les bonnes évaluations doivent être conçues et développées en suivant les principes de conception universelle de l'évaluation. Elles doivent tenir compte de divers moyens d'engagement, de représentation, d'action et d'expression.

Selon l'IB, une **bonne évaluation donne la priorité à des objectifs d'apprentissage riches de sens**. L'évaluation doit correspondre aux objectifs pédagogiques et aider les élèves à développer un large éventail de compétences et d'aptitudes.

La démarche de l'IB repose sur le principe qu'il faut **évaluer ce qui est important** et non considérer comme important ce qui peut être évalué. Ces éléments doivent être pondérés avec d'autres aspects, comme la fiabilité et la charge de travail des élèves.

Il est difficile pour chaque approche de satisfaire à toutes les priorités possibles. Plus précisément, la conception d'une bonne évaluation est différente pour l'évaluation sommative et l'évaluation formative. En partant de ce principe, l'IB considère qu'une bonne évaluation doit réunir les conditions suivantes :

- soutenir les objectifs pédagogiques ;
- s'appuyer sur tout un éventail de tâches d'évaluation ;
- tenir compte des compétences plus larges de l'élève et des compétences de pensée de haut niveau.

Chacune de ces conditions est passée en revue à tour de rôle dans les sections qui suivent.

## Une bonne évaluation soutient les objectifs pédagogiques

- Les évaluations doivent avoir un effet de remous positif, et encourager de bons apprentissage et enseignement.
- Les évaluations doivent permettre une bonne prévisibilité.

L'évaluation ne doit pas être perçue comme une activité indépendante de l'apprentissage et de l'enseignement. Les résultats de l'évaluation de l'IB reposent sur une évaluation sommative et ils n'ont pas vocation à fournir un retour d'information direct sur l'apprentissage et l'enseignement. Il est toutefois bien entendu que ce qui est inclus dans l'évaluation aura un impact sur ce qui est enseigné. Il s'agit de l'« effet de remous ».

Pour l'IB, la conception des évaluations doit encourager les résultats pédagogiques les plus souhaitables pour l'ensemble des élèves. Les effets sur l'apprentissage des élèves demeurent un aspect fondamental de

la conception des évaluations de l'IB. Cette prise en compte, associée à la pertinence du construit, aide l'IB à déterminer comment pondérer les différents éléments de la validité.

Une évaluation ayant des enjeux importants exerce une influence sur l'apprentissage et sur l'enseignement dont il est possible de tirer parti en concevant des évaluations qui encouragent une bonne pédagogie et une implication constructive des élèves dans leur propre apprentissage (voir, par exemple, Murphy, 1999).

Tout au long des programmes de l'IB, l'évaluation est étroitement liée à l'expérience d'apprentissage, dont elle est également le reflet. En suivant le modèle de l'alignement constructif (Biggs, 1996), tout est construit autour des objectifs d'apprentissage visés. Ces objectifs façonnent et orientent non seulement la manière d'enseigner aux élèves, mais encore la manière de les évaluer, en veillant à les préparer convenablement à la réussite. Pour en savoir plus, veuillez vous référer à la section intitulée « *Les approches de l'évaluation à l'IB* ».

Les caractéristiques personnelles que l'on cherche à développer chez les élèves, et qui sont exprimées dans la déclaration de mission de l'IB et explorées dans le profil de la communauté d'apprentissage, s'accordent très bien avec la **théorie constructiviste de l'apprentissage des élèves**. Dans le cadre qu'elle fixe, les élèves s'engagent activement dans le processus d'apprentissage, assument la responsabilité de leur propre apprentissage, et approfondissent leurs connaissances, leur compréhension et leurs compétences par l'intermédiaire de la recherche.

Ainsi, dans un certain nombre de matières, les modalités d'évaluation attendent explicitement des élèves un engagement actif et ouvert pour découvrir des perspectives culturelles différentes des leurs. L'altruisme et la compassion étant des qualités plus affectives, il est difficile d'en tenir compte dans l'évaluation officielle ; elles doivent néanmoins être représentées dans le système d'évaluation global. Ceci est en partie rendu possible grâce aux exigences spécifiques au programme incluses dans le programme d'études qui ne sont pas évaluées, comme l'exposition dans le PP, le projet d'engagement auprès de la communauté dans le PEI, le programme créativité, activité, service (CAS) dans le Programme du diplôme et l'engagement auprès de la communauté dans le POP.

Pour ce qui est de la conception des cours du Programme du diplôme ou du POP plus particulièrement, l'IB accorde une grande importance à la validité prédictive (la mesure dans laquelle les résultats peuvent prédire la réussite ultérieure), en sachant que la manière dont l'évaluation est menée influence grandement la façon dont les cours de l'IB sont enseignés dans les établissements scolaires. Le modèle d'évaluation utilisé dans chaque matière est conçu de manière à avoir une large base, de façon à pouvoir rassembler divers types de preuves. Le but est en effet de favoriser la pertinence du construit en fournissant un éventail de preuves aussi large que possible pour soutenir les résultats et l'apprentissage des élèves.

Alors que l'IB est parfaitement conscient de l'effet de rebrous exercé par les évaluations sur l'apprentissage et sur l'enseignement, il encourage les établissements scolaires à adopter une pédagogie développant l'intégralité des buts et des principes philosophiques de ses programmes chez les élèves.

L'IB mène régulièrement des études pour évaluer dans quelle mesure les programmes conçus constituent une bonne préparation à la poursuite de la scolarité. Pour de plus amples informations sur les études réalisées par l'IB, veuillez vous reporter à la section consacrée à la [recherche](#) sur son site Web.

## Une bonne prévisibilité

Le terme « prévisibilité » désigne la capacité à évaluer ce qui se passera ou à quel moment une chose se passera. Dans le domaine de l'évaluation, cela fait référence à l'aptitude des établissements scolaires à déterminer quels types de questions seront posées dans un examen donné et à quel moment. Une bonne prévisibilité participe à l'équité en faisant correspondre les questions au programme d'études utilisé par le personnel enseignant pour préparer les élèves à l'évaluation. Les élèves ont ainsi la garantie de savoir à quels grands types de questions s'attendre, bien que le contenu précis soit variable.

Pour proposer des options d'enseignement, les établissements scolaires doivent fournir un investissement et l'IB veut s'assurer qu'ils en sont récompensés pendant tout le cycle de vie d'un programme d'études (avant qu'il soit révisé). Tout le programme doit être examiné tel que l'exige l'évaluation. Une grande prudence est montrée à la fin d'un cours pour supprimer les risques de mauvaise prévisibilité (lorsqu'un établissement se rend compte de ce qui n'a pas encore été demandé et qui apparaîtra donc probablement dans un examen).

Le principe sous-jacent est que rien dans un examen donné ne doit constituer une surprise, ni pour l'élève, ni pour l'établissement scolaire. Quelle que soit la composante concernée, les questions posées doivent donc être explicitement inspirées du guide pédagogique correspondant. Dans la mesure du possible et s'il y a lieu, l'IB vise à éliminer les problèmes dus à la prévisibilité. Pour ce faire, il conçoit les évaluations de façon à ne pas avantager les élèves disposant de réponses préparées à l'avance.

La conception de l'évaluation s'avère primordiale pour veiller à ce que les façons d'évaluer tout type de thème, d'option ou de texte soient suffisantes, afin de limiter la mauvaise prévisibilité.

Tableau 3

*Comparaison entre une bonne et une mauvaise prévisibilité*

| Une bonne prévisibilité   | Une mauvaise prévisibilité   |
|---|--|
| Si l'on faisait l'inventaire de toutes les questions que l'on pourrait poser sur un thème, sur une option ou sur un texte en fonction de tous les mots-consignes, les questions posées par l'IB dans la matière en question y figureraient. | Le personnel enseignant prépare excessivement les élèves à certaines questions lorsqu'il remarque qu'elles varient très peu d'une session à l'autre. Cette pratique pose problème pour les questions à réponse plus développée, qui imposent de montrer toute une gamme de compétences (connaissances, analyse et évaluation).             |
| Les questions sont réutilisées de manière imprévisible.   | L'éventail des questions possibles est limité en raison de mauvaises décisions prises lors de la conception de l'évaluation (par exemple, un texte prescrit choisi pour un cours ne comporte qu'un seul thème principal sur lequel il est possible de poser des questions). Cela conduit à une mauvaise prévisibilité due à la conception. |

## Une bonne évaluation s'appuie sur tout un éventail de tâches d'évaluation

Une question à choix multiple, une question à réponse courte, une question à réponse développée, une dissertation, un projet, un travail tiré d'un dossier et un travail de recherche constituent tous des exemples de tâches d'évaluation.

La composante d'évaluation ou l'outil d'évaluation comporte une ou plusieurs tâches groupées en raison de la continuité du thème ou du contenu, ou encore par commodité. Les outils d'évaluation ou composantes d'évaluation peuvent prendre la forme d'un examen sommatif, d'un dossier de l'élève, d'un projet collectif à long terme ou d'un travail de recherche. Les concepts de tâche d'évaluation et de composante se chevauchent. L'élève n'effectue parfois qu'une seule tâche parmi un certain nombre de tâches qui lui sont proposées dans une composante.

Plusieurs raisons expliquent que l'IB utilise une grande variété de types de tâches et de composantes d'évaluation. Tout d'abord, d'un point de vue historique et pratique, Peterson (2003) a dit au sujet de l'élaboration initiale de l'évaluation dans le Programme du diplôme que l'on avait à la fois une obligation et une occasion de tenir compte des différentes techniques d'évaluation utilisées dans les pays et les régions (ou territoires) dans lesquels étaient situés les établissements scolaires que les élèves de l'IB essayaient principalement d'intégrer, et ce principe peut être étendu au PEI et au POP. Il existe également des aspects de la validité (liés à l'adéquation avec l'objectif visé) qui requièrent une approche différente de l'évaluation. Enfin, l'utilisation de diverses techniques d'évaluation aide à réduire les possibilités d'iniquité dans le domaine de l'évaluation (voir Linn, 1992 et Brown, 2002). L'éventail des composantes du modèle d'évaluation pour une matière donnée et l'ensemble des tâches inhérentes garantissent une représentation adéquate des accomplissements de l'élève par rapport à tous les objectifs définis pour cette matière.

## Le rôle de l'évaluation se déroulant en classe et de l'évaluation interne

L'évaluation se déroulant en classe offre différentes possibilités d'évaluer les élèves dans des domaines qui ne sont pas adaptés aux examens. Son principal avantage est qu'elle permet d'engager les élèves dans une tâche plus longue, grâce à laquelle il leur est possible d'étudier un problème et de montrer le développement de leur raisonnement sans les contraintes de temps caractéristiques des examens. Un large éventail de tâches d'évaluation peuvent ainsi uniquement être utilisées pour des évaluations se déroulant en classe.

Parmi les tâches qui se prêtent parfaitement à une évaluation se déroulant en classe, il est possible de citer les projets, le travail de terrain, les travaux pratiques en laboratoire, les représentations artistiques et la recherche mathématique.

L'évaluation interne a d'autres **avantages** dans le contexte d'un diplôme reconnu internationalement. Elle offre une plus grande **flexibilité** au niveau du choix des thèmes, tout en continuant de requérir un ensemble commun de compétences. Cela permet aux établissements scolaires de placer l'étude dans **un contexte local, culturel ou géographique**, ou d'établir des liens plus étroits entre la classe et le monde qui s'étend au-delà. Les écoles internationales, qui comptent des élèves au bagage culturel souvent différent de celui des gens du pays où l'établissement est implanté, peuvent utiliser les travaux soumis à l'évaluation interne pour mettre en place une participation plus importante dans la société ou l'environnement local.

En outre, l'évaluation interne donne souvent aux élèves la possibilité de choisir leur propre sujet ou thème. Ce choix leur permet de suivre leurs propres centres d'intérêt et leur donne une plus grande maîtrise de leur apprentissage. De par sa flexibilité, l'évaluation interne apporte une contribution précieuse à l'éducation des élèves ; elle renforce non seulement la validité du processus d'évaluation, mais également celle de l'ensemble de l'apprentissage.

L'évaluation se déroulant en classe pose elle aussi des **difficultés**, en tête desquelles figure la nécessité de garantir l'**intégrité intellectuelle**. Lorsque les travaux sont effectués dans des conditions non standardisées (c'est-à-dire hors surveillance) – en dehors de la salle de classe notamment –, il est bien plus compliqué de s'assurer que les élèves n'entreprennent pas des activités enfreignant la politique de l'IB en matière d'intégrité intellectuelle, comme demander à quelqu'un de faire le travail à leur place. Et ce défi présente une complexité exponentielle avec l'accès libre à Internet. C'est pourquoi l'IB considère que le personnel de l'établissement et notamment l'équipe pédagogique sont les mieux placés pour aider à authentifier les travaux des élèves et à repérer les travaux réalisés par d'autres personnes. Vous trouverez des informations et des ressources complémentaires sur la page du [site Web de l'IB](#) qui est consacrée à l'intégrité intellectuelle.

Une autre difficulté de l'évaluation se déroulant en classe est la **charge** importante qu'elle peut faire peser à la fois sur le personnel enseignant et sur les élèves. En effet, pour organiser des évaluations en classe, il faut rogner sur les heures d'enseignement, et les tâches données en interne impliquent généralement que les élèves y consacrent beaucoup de temps. Bien qu'il soit approprié que les enseignantes et enseignants passent beaucoup de temps à familiariser les élèves avec les compétences et les processus requis pour l'évaluation interne, la tentation peut être grande de s'entraîner plus que nécessaire à effectuer la tâche donnée, ce qui limite donc encore davantage le temps consacré à l'enseignement. Les guides pédagogiques du Programme du diplôme expliquent dans le détail l'aide que le personnel enseignant peut apporter pour les travaux d'évaluation interne.

L'évaluation se déroulant en classe peut être notée soit **en externe**, soit **en interne**. En cas d'évaluation externe, le travail produit par l'élève est envoyé à l'IB et noté par un examinateur ou une examinatrice. La qualité des décisions des examinateurs et examinatrices est contrôlée par l'IB (voir la section « [La notation](#) »).

Concernant l'évaluation interne au contraire, l'établissement scolaire (ou plus exactement l'enseignant ou enseignante) note le travail de l'élève, puis l'IB vérifie que la norme mondiale a été appliquée correctement, en utilisant un processus appelé « révision de notation » (voir la section consacrée à la [révision de notation](#)).

Les avis sur l'évaluation interne sont partagés à travers le monde. Certains systèmes éducatifs attachent une grande importance au fait que les enseignantes et enseignants sont les plus à même de donner une

opinion globale sur l'accomplissement de l'élève sans se limiter au seul « aperçu » permis par une évaluation unique. D'autres se basent sur les résultats des élèves pour déterminer la réussite des enseignantes et enseignants, ce qui les incite donc fortement à attribuer des notes élevées à toutes les évaluations internes.

Le jugement de l'enseignant ou enseignante peut aussi être altéré par sa connaissance des travaux antérieurs de l'élève, car ils peuvent introduire certains biais. Les enseignantes et enseignants ne connaissent parfois pas très bien les limites de leur rôle pour guider et aider les élèves lors de la réalisation des travaux d'évaluation interne, et n'ont souvent qu'une vue restreinte des niveaux d'accomplissement que l'on trouve à l'échelle mondiale dans leur discipline. L'influence des niveaux d'accomplissement généraux existant dans leur établissement peut également être très grande lorsque les enseignantes et enseignants évaluent les travaux de leurs propres élèves. Même si rien ne pousse à noter de manière bienveillante le travail envoyé pour l'évaluation, les enseignantes et enseignants auront du mal, du fait de la relation professionnelle établie avec les élèves, à prendre des décisions objectives. Les études menées sur ce biais inconscient indiquent qu'il peut avoir des effets perceptibles (voir, par exemple, Zanga et De Gioannis, 2023).

Pour l'IB, l'utilisation de tâches d'évaluation riches de sens, passant obligatoirement par un travail réalisé en classe, et évaluées par une personne ayant vu le développement s'opérer chez les élèves, comporte des avantages plus grands que les risques présentés par l'évaluation interne. C'est pourquoi le modèle d'évaluation de chaque matière comporte généralement une tâche d'évaluation interne.

### Le travail en collaboration et les notes individuelles

Des difficultés se présentent au moment d'évaluer les tâches réalisées en collaboration, dans lesquelles les élèves doivent effectuer un travail collectif pour réaliser une activité. Il s'agit d'un aspect important que les élèves rencontreront dans le milieu du travail, mais qui est souvent négligé dans les évaluations scolaires.

Dans certains cas, il est possible de déterminer facilement le rôle de chacun dans une activité de groupe. Par exemple, une représentation de danse permet de distinguer les compétences de chaque personne et, par conséquent, de procéder à une notation individuelle. Dans ce type de situation, le travail de groupe n'est pas sujet à controverse.

Dans d'autres cas, il est impossible de savoir qui est responsable de quels aspects du travail final. Le cas échéant, il est possible d'attribuer une note unique à l'ensemble du groupe, mais cette pratique ne tient pas compte des différences de contribution de chaque élève au résultat global. Ainsi, même si une personne réalise l'essentiel du travail et apporte la plus grande part d'expertise, elle ne bénéficie d'aucune reconnaissance particulière. Étant donné que les résultats sont utilisés à des fins de sélection, l'IB considère que ce traitement n'est pas juste. Il s'efforce donc d'éviter les évaluations de groupe lorsqu'elles ne permettent pas de juger l'accomplissement de chaque élève.

## Une bonne évaluation tient compte des compétences plus larges de l'élève et des compétences de pensée de haut niveau

Le système éducatif de l'IB vise bien au-delà du simple apprentissage de « faits » par les élèves. Cet objectif de longue date se retrouve dans la réflexion actuelle d'un certain nombre de gouvernements sur le besoin de doter les élèves des **compétences** nécessaires au XXI<sup>e</sup> siècle et des **aptitudes recherchées dans le monde professionnel** ou dans d'autres **initiatives** de ce genre.

La philosophie de l'IB et son approche en matière de compétences des élèves se concentrent sur le **profil de la communauté d'apprentissage** et sur son lien avec la **sensibilité internationale**. Ceci est expliqué de façon détaillée dans la « Section C : les processus spécifiques aux programmes de l'IB », qui porte sur la continuité entre les programmes. Il faut avant tout garder à l'esprit que l'approche de l'IB en matière d'évaluation garantit une évaluation portant non seulement sur la connaissance du contenu, mais encore sur un éventail d'aptitudes calquées sur des pratiques concrètes. En s'appuyant sur une conception

mûrement réfléchi, l'évaluation de l'IB vise à développer chez les élèves un goût pour l'**apprentissage permanent** et à leur fournir la préparation adéquate pour relever les défis à venir.

Par conséquent, une bonne évaluation tient compte de **toute la gamme de résultats** qu'un cours souhaite obtenir et permet aux élèves de montrer leurs compétences dans chaque domaine. Cependant, il est parfois souhaitable, voire uniquement faisable, de mesurer une partie de ces résultats, et une évaluation de qualité tient compte de ces limites. Pour l'IB, le meilleur système consiste à classer ces résultats en fonction des **compétences de pensée de haut niveau**, des **compétences** plus larges des élèves et de leur lien avec la **sensibilité internationale**.

## Les compétences de pensée de haut niveau

Les évaluations de l'IB sont conçues pour mesurer les compétences de pensée de haut niveau que sont la **capacité d'évaluation** et la **capacité d'analyse**, et non la simple aptitude à se rappeler des connaissances.

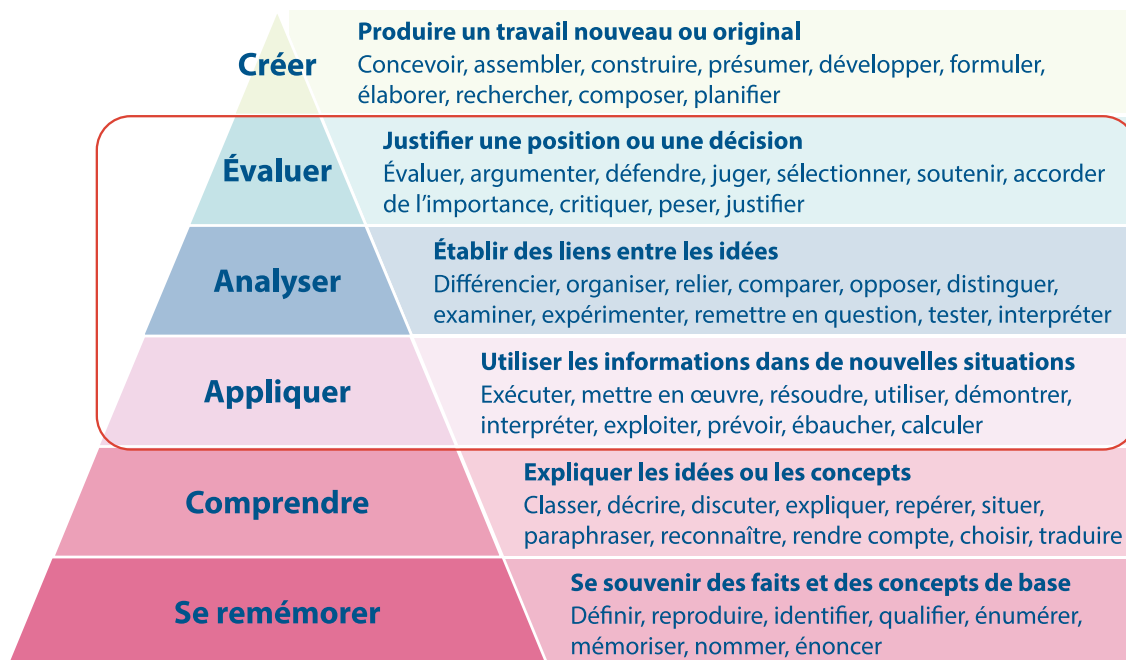
Cette remarque a été faite par Alec Peterson (2003), qui a donné forme à la philosophie éducationnelle de l'IB. Il a déclaré que ce qui importe n'est pas l'acquisition et la régurgitation de faits ou d'interprétations prémâchées de faits, mais le développement de facultés de l'esprit ou de modes de pensée qui peuvent être appliqués à de nouvelles situations et à de nouvelles présentations des faits au fur et à mesure qu'ils surviennent. Lors de la conférence des directeurs et directrices d'école du monde de l'IB, organisée à Singapour en 2011, Sugata Mitra est allé un peu plus loin, en soutenant que, puisque Internet permettait aujourd'hui d'avoir une profusion de connaissances à disposition immédiatement, les connaissances n'avaient plus vraiment de valeur intrinsèque. Il a fait valoir que c'était donc la capacité à analyser, à interpréter et à constituer des connaissances qui était exigée des citoyennes et citoyens du XXI<sup>e</sup> siècle (Mitra, 2011).

Les évaluations de l'IB accordent une attention toute particulière aux compétences de pensée de haut niveau, en définissant par exemple les attentes pour ses tâches d'évaluation à partir de taxonomies reconnues, comme celle de Bloom (Bloom et coll., 1956 ; Anderson et Krathwohl, 2001). Les objectifs pédagogiques qui y sont rassemblés constituent un cadre utile pour refléter la diversité des compétences et des processus cognitifs qui sont requis pour réussir les tâches d'évaluation et obtenir le plus de points possible. Les compétences supérieures décrites par Bloom exigent sans aucun doute un type d'évaluation différent. Les compétences en matière d'analyse, de synthèse et d'évaluation ne peuvent être correctement évaluées que s'il est demandé aux élèves d'analyser, de synthétiser et d'évaluer dans le détail. La seule façon réaliste de mesurer l'accomplissement des élèves dans ces domaines est de se pencher sur leurs réalisations et, du fait que les résultats d'une telle activité ne peuvent être rigoureusement établis, de telles évaluations doivent être relativement non structurées et ouvertes, et accepter de nombreuses réponses correctes différentes.

Figure 2

La taxonomie de Bloom

### Les capacités à vérifier chez les élèves



Pour être pertinentes, les évaluations de l'IB doivent impérativement pouvoir reconnaître et récompenser l'accomplissement des élèves dans ces compétences, en dépit des difficultés que cela pose au niveau de la fiabilité et d'autres aspects de la validité. Les tests qui ne récompensent que le rappel de connaissances, de concepts et de techniques courantes ne sont pas adaptés aux objectifs du système éducatif de l'IB, car ils n'abordent pas comme il se doit ces compétences de pensée de haut niveau indispensables.

## Le profil de la communauté d'apprentissage et le développement des compétences

Aujourd'hui, l'éducation s'apparente bien plus à un mode de pensée fondé sur la créativité, le raisonnement critique, la résolution de problèmes et la prise de décision. Elle tient également à des méthodes de travail, notamment la communication et la collaboration, et à des outils de travail, y compris la capacité de reconnaître et d'exploiter le potentiel des nouvelles technologies, mais aussi d'éviter leurs écueils. Enfin, et surtout, l'éducation concerne la capacité de vivre sa citoyenneté de façon active et engagée dans un monde protéiforme. À leur tour, les individus exercent une influence sur leurs souhaits et modes d'apprentissage, d'où une évolution du rôle des spécialistes de l'éducation.

(Schleicher, 2016, traduction libre)

D'aucuns prétendent de plus en plus que le XXI<sup>e</sup> siècle requiert des compétences radicalement différentes de celles attendues des générations précédentes, même si d'autres soutiennent que l'approche de la recherche qui sous-tend ces compétences du XXI<sup>e</sup> siècle est en réalité valorisée depuis Socrate. La majorité s'accorde toutefois à dire qu'il est important que les élèves possèdent un large éventail d'aptitudes pour recevoir la préparation adaptée à la vie qui les attend. Voir, par exemple, les arguments avancés par Llewellyn (2014).

Bien qu'il existe de nombreux cadres des compétences du XXI<sup>e</sup> siècle, notamment celui défini par l'Organisation de coopération et de développement économiques (OCDE) ou par RAND Education et le National Research Council (NRC), l'IB énumère et décrit ces compétences dans le profil de la communauté d'apprentissage.

Tous les aspects du profil de la communauté d'apprentissage ne peuvent pas être évalués au moyen d'une évaluation sommative, mais plusieurs sont englobés dans la notion de compétences de pensée de haut niveau. Une bonne évaluation reconnaît l'importance de ces caractéristiques et ouvre la voie au développement de ces compétences, au moyen d'activités formatives et sommatives. Il pourrait s'agir, par exemple, d'encourager des approches éthiques des enquêtes et des expériences en psychologie et en science, ou l'exploration approfondie d'un dilemme éthique dans le projet de réflexion, en facilitant une révision par les pairs appropriée et en présentant des contextes inattendus aux élèves.

Pour de plus amples informations sur l'approche globale adoptée par l'IB en matière de compétences des élèves, veuillez vous reporter aux documents disponibles sur son [site Web](#) ou à la publication *Des principes à la pratique* spécifique à chaque programme.

## La sensibilité internationale et la compréhension interculturelle

La mission de l'IB est d'aider les élèves à développer « la sensibilité nécessaire pour contribuer à bâtir un monde meilleur et plus paisible, dans un esprit d'entente mutuelle et de respect interculturel » et à « comprendre que les autres, dans leurs différences, puissent aussi être dans le vrai » (extraits de la déclaration de mission de l'IB, 2024). Les programmes de l'IB sont également étudiés par des élèves vivant dans de nombreux pays et régions (ou territoires) représentant toute une palette de nationalités. L'apprentissage et l'enseignement prônés par l'IB se caractérisent donc à la fois par un **contexte international** et par un accent mis sur le développement de la **sensibilité internationale**, deux éléments qui doivent se retrouver dans l'évaluation. Pour rendre cette pluralité de perspectives, il est important de faire appel à des spécialistes de l'éducation, notamment des personnes chargées du développement des programmes d'études et de l'évaluation, provenant de contextes culturels très divers. La mission de l'IB tient à ne pas occulter les différences, mais au contraire à les aborder de façon à ce que les élèves les explorent de façon constructive et positive.

Dans certaines disciplines, il est possible de favoriser la diversité culturelle en tenant compte des différentes perspectives culturelles et en mettant l'accent sur la variété culturelle du contenu dans le programme d'études. Plusieurs matières du Programme du diplôme illustrent bien cette approche, notamment la biologie, la chimie, la psychologie et les arts visuels. Dans les trois premiers cours, les options permettent aux établissements scolaires de choisir un contenu qui respectera, dans une certaine mesure, leurs traditions culturelles concernant l'enseignement de la matière.

Dans d'autres disciplines, comme les arts ou la langue et la littérature, la sensibilité internationale est encouragée au moyen du contenu pédagogique recommandé et prescrit. Cette notion peut également être mise en avant dans un éventail plus large de tâches d'évaluation interne.

La sensibilité internationale ne se réduit pas au fait de connaître et de comprendre d'autres cultures. Les **attitudes** affichées et les **actions** entreprises en sont d'autres aspects fondamentaux. Il est difficile d'évaluer les attitudes à l'aide des évaluations scolaires traditionnelles.

L'IB parvient à le faire par l'intermédiaire des éléments non évalués du tronc commun de ses programmes, comme le projet communautaire dans le PEI et la composante CAS dans le Programme du diplôme. Étant donné qu'il est impératif de terminer avec succès les composantes du tronc commun du programme pour recevoir le certificat du PEI, le diplôme de l'IB ou le certificat du POP, ces composantes influent grandement sur le résultat global du processus d'évaluation de l'IB.

Figure 3

*L'action fondée sur des principes*



Le fait de permettre aux élèves de choisir les questions auxquelles répondre peut être considéré comme la meilleure façon de satisfaire les différentes exigences internationales en matière d'évaluation. Cette approche pose cependant certains problèmes du point de vue de la **comparabilité** entre les options. Cela est souvent le cas lorsque plusieurs questions sont proposées ou que les tâches d'évaluation sont très ouvertes, ce qui rend même plus difficile de définir une « difficulté égale » lorsque les élèves proviennent de milieux scolaires très différents. En règle générale, il est plus facile de garantir la comparabilité avec des tâches communes qui permettent aux élèves de s'appuyer sur leurs expériences personnelles pour rédiger leurs réponses. Il revient alors à l'examineur ou examinatrice de réussir à maintenir une norme commune.

Il est possible d'obtenir des renseignements sur la comparabilité en analysant les résultats des élèves. Cette analyse est abordée plus en détail dans la section « [L'attribution des notes finales et la totalisation](#) ».

L'évaluation effectuée dans un contexte international présente des défis supplémentaires (en matière d'équité), qui sont plus importants que ceux normalement rencontrés dans un système national. Des questions qui sont parfaitement appropriées dans un pays peuvent devenir complètement inappropriées dans un autre. Ainsi, les questions portant sur les sports, les voyages, les divertissements, les événements historiques voire sur la météo doivent être préparées et traitées avec beaucoup de soin. Il semblerait que la seule façon de contourner ce problème soit de rédiger des questions d'examen dénuées de tout contexte socioculturel. Toutefois, non seulement cela rendrait les questions d'examen assez limitées, mais cela irait aussi à l'encontre de la philosophie de l'évaluation de l'IB et des bonnes pratiques de l'évaluation en général, qui visent à garantir la validité au moyen de tâches basées sur le contexte. Pour que l'apprentissage soit bon, il est vital que le travail et l'évaluation tiennent compte du contexte.

Il existe deux façons de résoudre ce dilemme. La première approche consiste à fournir des informations contextuelles appropriées aux élèves, grâce aux précisions apportées dans le contenu du programme de la matière, en incluant les études de cas qui sont à la base des questions, ou encore dans l'examen lui-même.

La seconde approche consiste à utiliser des tâches et des questions d'évaluation plus ouvertes, qui permettent aux élèves de choisir leur propre contexte pour élaborer leurs réponses. Adopter cette

approche signifie veiller à axer la notation sur une compréhension approfondie plutôt que sur une simple connaissance du contenu du programme de la matière, car il n'y a aucune base commune en ce qui concerne le contenu. Ceci s'accorde parfaitement avec la philosophie de l'IB en matière d'évaluation.

Figure 4

*Des contextes culturels très divers*



Quand bien même ces deux méthodes seraient utilisées, les élèves pourraient avoir à réaliser des tâches d'évaluation présentant des contextes socioculturels inconnus à leurs yeux. Cela s'accorde également avec la philosophie de l'IB en matière d'évaluation, qui encourage les élèves à faire preuve d'une **plus grande ouverture d'esprit**, d'avoir une **meilleure conscience du monde** et de se sentir plus **à l'aise** et plus **aptes** à fonctionner dans des contextes culturels qui ne leur sont pas familiers.

Un nombre important d'élèves de l'IB s'inscrivent aux examens dans une langue autre que leur première ou meilleure langue. Presque tous ces cas concernent l'anglais, car les élèves qui travaillent en français ou en espagnol (les deux autres langues les plus souvent utilisées pour les évaluations de l'IB) parlent généralement couramment ces langues. Lors des révisions rédactionnelles et éditoriales des évaluations, il est donc veillé de près à ce que la formulation des questions ne vienne pas désavantager les élèves recourant à une langue supplémentaire. Cette priorité est maintenue durant tout le processus d'édition des examens.

## Le rôle de la technologie dans l'évaluation de l'IB

- Les élèves font une utilisation très large de la technologie, ce qui s'observe notamment dans leur vie sociale, mais aussi dans leur apprentissage et la rédaction de leurs dissertations. Cet usage de la technologie se poursuivra en s'intensifiant lors de leurs études supérieures et dans leur vie professionnelle. Le recours à la technologie pour les examens apparaît donc comme une pratique conforme à leur expérience du monde.
- L'IB s'appuie sur la technologie pour proposer une évaluation pertinente et de meilleure qualité. L'utilisation de la technologie n'influence pas les approches de l'évaluation choisies.
- Il est important de faire une distinction entre les effets de la technologie permettant d'apporter un soutien aux examinateurs spécialisés et examinatrices spécialisées (notation électronique) et l'utilisation de la technologie permettant de concevoir des évaluations pertinentes pour les élèves.
- Il conviendra d'apprendre en classe à utiliser la technologie qui servira pour l'évaluation.

### L'évaluation numérique

L'évaluation numérique désigne l'évaluation qui est passée sur un appareil informatique. En fonction de ce qui est évalué, les questions et les stimuli peuvent être de nature très diverse, et aller des évaluations très interactives s'appuyant sur de nombreux médias aux examens plus traditionnels de type dissertation.

D'un point de vue pratique, l'évaluation numérique exige que chaque élève travaille sur un appareil individuel lors de l'examen. Avec le modèle actuellement utilisé pour le PEI, une connexion Internet est recommandée, sans pour autant être exigée, puisque les examens sont chargés sur les appareils informatiques au préalable. Chaque appareil est « verrouillé » pendant toute la durée d'utilisation du logiciel d'examen et des sauvegardes sont effectuées à la moindre interruption (le surveillant ou la surveillante pouvant procéder à un redémarrage en cas d'interruption justifiée de l'examen). Il est fortement recommandé de passer les examens numériques du PEI en ligne, car cela permet l'enregistrement des réponses des élèves à intervalles réguliers et allège les contraintes administratives pesant sur les établissements scolaires. Bien que la grande majorité des établissements et des élèves utilise la fonctionnalité en ligne, il ne s'agit pas d'une condition obligatoire pour passer les examens du PEI et l'IB propose un mode hors ligne. Si les examens numériques du Programme du diplôme et du POP exigent une connexion Internet, les élèves accèdent aux examens en mode verrouillage, ce qui restreint leur utilisation au simple examen.

L'évaluation numérique présente des avantages considérables : elle permet de tester ce que l'IB souhaite véritablement tester, plutôt que de se limiter à ce qui est possible avec un examen sur papier. Parmi ces avantages, décrits dans le détail dans la section intitulée « [Les avantages de l'évaluation numérique pour la validité](#) », il convient de relever que ce système permet d'utiliser des supports vidéo, offre aux élèves la possibilité d'utiliser les outils de traitement de texte courants pour rédiger leurs dissertations, et favorise l'accessibilité, en donnant par exemple la possibilité aux élèves de modifier les couleurs et la taille de la police.

L'évaluation numérique continuera d'évoluer, et l'IB sera en capacité de développer ses évaluations en accord avec les avancées technologiques. À l'heure actuelle, seule l'utilisation des claviers et des souris ou des pavés tactiles est prise en compte, mais dans quelques années, il est fort probable que les écrans tactiles seront monnaie courante et ouvriront d'autres horizons. En poussant la réflexion encore plus loin, quelle pourrait être la place de la réalité virtuelle dans les évaluations ? Un tel environnement pourrait donner encore plus d'authenticité aux évaluations en langues étrangères, et les films de science-fiction sont

un cadre favorable au développement d'idées sur la façon de gérer des tâches techniques (questions scientifiques) dans des laboratoires de réalité virtuelle.

L'IB ayant pour souci permanent la transformation vers l'évaluation numérique pour tous ses programmes, une nouvelle terminologie verra forcément le jour et se répercutera sur la communauté de l'évaluation au sens large. Un glossaire des termes utilisés par l'IB dans le cadre de l'évaluation, notamment numérique, figure à la fin de la présente publication.

## Un enseignement s'appuyant sur la technologie

La technologie modifie la façon d'enseigner : certaines classes utilisent des tableaux interactifs et des cours en ligne ouverts à tous et toutes mettent l'enseignement à la portée d'un plus grand nombre. La plupart des élèves ne rédigent plus leurs dissertations à la main, mais à l'aide d'un appareil informatique.

L'abandon de l'écriture manuscrite au profit des appareils crée en réalité un décalage entre les pratiques de l'évaluation et les pratiques d'enseignement. Pour beaucoup d'élèves, l'expérience consistant à écrire au stylo deux ou trois heures de suite constitue une expérience presque exclusivement réservée au passage des examens, ce qui altère considérablement l'authenticité des évaluations.

En raison de l'utilisation accrue des évaluations numériques, il est important de familiariser les élèves au style des évaluations, de façon à ce que l'utilisation de l'interface ne soit pas un obstacle risquant de masquer leur compréhension de la matière. Si cette difficulté peut être surmontée en fournissant aux élèves des outils de familiarisation et en leur faisant passer des examens d'entraînement, la meilleure solution reste que les élèves utilisent ces outils dans le cadre de l'enseignement.

## À quoi ressemble une bonne évaluation numérique ?

L'évaluation numérique peut améliorer la qualité des évaluations, en tirant profit de la technologie pour poser des questions qui ne pourraient pas être posées dans un examen traditionnel sur papier. L'IB considère qu'une bonne évaluation numérique :

- utilise la technologie pour présenter des tâches plus valides, plus authentiques et plus pertinentes ;
- supprime les obstacles administratifs et simplifie la logistique ;
- évite tout nouvel obstacle à la réalisation de l'évaluation.

## La notation électronique

Le terme « notation électronique » est utilisé pour décrire la manière dont l'IB note le travail envoyé pour l'évaluation, en se servant d'appareils informatiques pour afficher les copies et enregistrer les notes. Les examinateurs et examinatrices de l'IB ne reçoivent plus les copies d'examen ni les travaux d'évaluation interne au format papier, sauf circonstance exceptionnelle liée à des aménagements à des fins d'accès et d'inclusion. Les travaux au format papier sont numérisés pour que les examinateurs et examinatrices puissent accéder au format numérique en se connectant à Internet et en utilisant un logiciel de notation spécialisé.

La notation électronique présente des avantages de taille au niveau de la vitesse et de la qualité de la notation.

- L'IB conserve le travail envoyé pour l'évaluation. De ce fait, si ce travail doit être revu, pour quelque raison que ce soit, par un second examinateur ou une seconde examinatrice, il est alors immédiatement disponible sans devoir être expédié dans un autre pays.
- L'IB peut rendre le travail anonyme afin de réduire les risques de biais dans la notation.
- Il peut aussi mettre en œuvre un processus de contrôle de la qualité plus rigoureux et vérifier les normes de notation au fil de la correction.
- Les travaux envoyés pour l'évaluation provenant d'un même établissement peuvent être répartis entre l'ensemble des examinateurs et examinatrices, au lieu d'être tous notés par une seule personne. L'IB peut ainsi pousser encore plus loin le contrôle de la qualité.

L'IB a commencé à utiliser la notation électronique en 2010 et ce système n'impose pas aux élèves de rendre un travail réalisé sur un support numérique. À l'heure actuelle, les réponses aux examens sont, pour la plupart, rédigées à la main, puis les copies sont numérisées pour être utilisables au format numérique.

## Les risques associés à l'évaluation numérique

Développer une évaluation numérique demande tout d'abord de prendre conscience de plusieurs difficultés importantes, puis de les surmonter pour assurer une équité entre l'ensemble des élèves et des établissements. Quelques-uns des principaux risques à gérer par l'IB sont énumérés dans les sections qui suivent. Une analyse plus minutieuse sera réalisée dans plusieurs sections de la présente publication ainsi que dans d'autres ressources de l'IB traitant de l'évaluation numérique.

### La charge pesant sur les établissements

L'IB a conscience que l'utilisation de la technologie s'est inscrite dans le quotidien des élèves, et que ce principe doit aussi s'appliquer à leur expérience de l'évaluation. Il n'en demeure pas moins que les établissements ont besoin de temps pour modifier leurs pratiques et que les examens numériques nécessitent des aménagements particuliers par rapport aux examens sur papier. L'IB gèrera l'introduction des examens numériques du Programme du diplôme et du POP de façon à ce que tous les établissements puissent suivre les avancées technologiques, mais aussi pour qu'ils n'aient pas à s'adapter de manière trop brutale, quelle que soit leur situation géographique.

### La garantie d'une transition progressive vers l'évaluation numérique

Un travail sera réalisé avec la communauté de l'IB pour que le calendrier d'introduction des évaluations numériques permette à tous les établissements intéressés par une technologie offrant une évaluation de meilleure qualité de pouvoir en bénéficier. L'IB estime que le potentiel de la technologie doit être exploité intelligemment pour améliorer la qualité de l'apprentissage et de l'enseignement et, par conséquent, de l'évaluation.

### La défaillance technologique

Aucun degré de défaillance technologique ne peut être accepté quand il s'agit du passage des examens. En effet, chaque élève est en droit de vivre cet événement sans heurts. L'IB propose aux établissements de suivre un processus complet de familiarisation, de test et de conformité la veille des examens. L'expérience de l'IB montre en effet que la plupart des problèmes viennent de l'infrastructure des établissements, ce que le processus de familiarisation, de test et de conformité peut permettre de repérer et de résoudre bien avant les examens.

### La sécurité

Le niveau de sécurité des examens numériques doit être au moins identique à celui des examens sur papier. Les évaluations de l'IB sont conçues de façon à verrouiller toutes les autres fonctionnalités de l'appareil pendant leur déroulement. Lorsque le matériel d'évaluation est envoyé aux établissements avant les examens, comme pour les évaluations électroniques du PEI, ces fichiers sont cryptés et protégés par un mot de passe.

Même si la sécurité demande à être constamment renforcée, en raison des développements de plus en plus sophistiqués qui voient le jour, il ne faut pas perdre de vue que les examens sur papier utilisés actuellement sont eux aussi exposés à certains risques.

### La technologie pour la technologie

L'IB s'est engagé à utiliser la technologie pour améliorer la qualité des éléments pouvant être évalués, et non pour le simple fait d'utiliser cette technologie. L'IB continuera de faire ressortir la valeur ajoutée par l'utilisation de la technologie à la validité de son évaluation, et reflétera ce progrès dans le processus de développement de l'évaluation.

## Le biais

L'IB reconnaît que des élèves réussiront mieux que d'autres en raison de leur capacité à passer les évaluations en se servant d'appareils informatiques. Il ne perd cependant pas de vue que les examens sur papier peuvent également comporter un biais : les élèves ayant des difficultés à écrire pendant longtemps au stylo ou écrivant lentement subissent déjà un désavantage.

Par principe, l'IB n'a pas pour objectif de mesurer la capacité des élèves à utiliser un stylo ou la saisie informatique, et il concevra les tâches d'évaluation de façon à ne pas procurer un avantage aux personnes capables de taper ou d'écrire rapidement.

Afin de garantir la bonne application de ce principe, l'IB accorde une importance particulière aux études publiées sur les « effets de la technologie ».

## L'évolution des normes, vers le haut ou vers le bas

Les notes finales de l'IB ont un sens et ce sens sera protégé par le jugement spécialisé des examinateurs superviseurs et examinatrices superviseuses de l'IB, et fondé sur les données sur les résultats. Par exemple, il peut être plus facile pour l'élève d'élaborer une dissertation bien structurée en ayant la possibilité de couper et de coller des paragraphes au fil de la rédaction. L'IB portera une attention particulière à cet élément lors de l'établissement des seuils d'attribution des notes finales. De cette façon, l'élève autant à l'aise avec les examens sur papier qu'avec les examens numériques obtiendra le même résultat, qui rendra compte de sa compréhension et de sa capacité d'analyse dans la matière.

L'évaluation numérique permet de tester de nouveaux aspects ainsi que des aspects que l'IB avait toujours valorisés, sans pouvoir les évaluer jusque-là. Cela pourra donc nécessiter un ajustement des éléments à montrer par les élèves dans les évaluations. Pour en savoir plus, veuillez vous référer à la section intitulée « [Les normes](#) ». L'IB opérera toutefois ce changement dans le respect de ses objectifs de longue date et de façon à refléter les qualités décrites dans les descripteurs des notes finales en vigueur.

## Lectures complémentaires

Pour en savoir plus sur l'évaluation numérique dans le cadre de l'IB, veuillez consulter les ressources qui suivent.

- La page « [Transition vers les examens numériques](#) » du Centre de ressources pédagogiques
- [Manuel pour l'évaluation électronique du PEI](#)
- [L'évaluation électronique du Programme d'éducation intermédiaire \(PEI\) de l'IB](#) (vidéo)
- [Prêts pour l'évaluation électronique du PEI de l'IB ! Introduction](#)

## L'évaluation dans l'éducation

- Dans le domaine de l'éducation, l'évaluation est souvent qualifiée soit de formative (l'évaluation **au service de** l'apprentissage), soit de sommative (l'évaluation **de** l'apprentissage). La tendance actuelle va dans le sens d'une évaluation **en tant qu'**apprentissage.
- Quel que soit le contexte concerné, le type d'évaluation doit être déterminé en fonction de l'utilisation des résultats ou du but de l'évaluation.
- L'évaluation peut influencer les pratiques d'enseignement et doit être conçue de façon à rendre positif tout effet de remous (c'est-à-dire l'incidence de l'évaluation sur l'apprentissage et sur l'enseignement).

### Définition de l'évaluation

Tel qu'expliqué dans la publication intitulée *En quoi consiste le système éducatif de l'IB ?* (2019), le Baccalauréat International (IB) suit une approche constructiviste de l'apprentissage. Celle-ci confère un rôle actif à l'élève et reconnaît l'importance du contexte dans l'efficacité de l'apprentissage (Murphy, 1999). Si l'évaluation vise à renforcer l'efficacité de l'apprentissage et de l'enseignement, alors elle doit être conçue selon la théorie de l'apprentissage constructiviste. Pour consulter des travaux de recherche sur ce concept, voir Shepard (1992), Wood (1998), Black (1999) ainsi que Lambert et Lines (2000).

Le terme « évaluation » peut désigner toutes les façons de recueillir et d'évaluer des données sur l'accomplissement de l'élève. Parmi les évaluations courantes, il est possible de citer les tests, les examens, les travaux pratiques approfondis, les projets, les dossiers et les travaux oraux. Dans certaines situations, les évaluations se déroulent sur une longue période et dans d'autres, elles prennent quelques heures. En fonction de leur type et de leur but, les évaluations peuvent être jugées par un enseignant ou une enseignante, l'élève, ses pairs ou encore un examinateur ou une examinatrice.

Figure 5

Exemples de différents types d'évaluations



Les tâches d'évaluation qui vérifient le niveau de compétence de l'élève s'appuient soit sur un système de **compensation**, soit sur un système de **maîtrise** (voir la figure 6).

Le système de **compensation** est utilisé dans la plupart des examens externes. Il permet à un résultat élevé dans un domaine d'atténuer un résultat plus faible dans un autre.

À l'inverse, le système de **maîtrise** exige un niveau minimal (maîtrise) dans chaque partie d'une évaluation.

Un examen est composé de 3 questions, chacune notée sur 10.

La note de passage est 15 et pourrait être obtenue soit si l'élève obtient la note 5 à chacune des 3 questions, soit si l'élève obtient les notes 10, 3 et 2. Dans le second cas, un résultat parfait à la première question vient compenser les résultats plus faibles obtenus aux deux autres questions.

En reprenant le même exemple (un examen composé de 3 questions et chaque question étant notée sur 10), si une note de passage de 5 était exigée pour chaque question, l'évaluation serait réussie dans le premier cas, mais échouée dans le second. Des notes de 10, 10 et 4 signifieraient également un échec.

Figure 6

Comparaison entre le système de compensation et le système de maîtrise

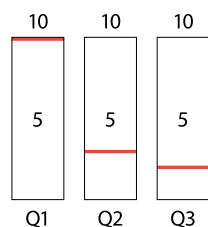
### Système de compensation

Nombre de questions : 3

Nombre de points (chacune) : 10

Note de passage minimale (globale) : 15/30

Application d'un système de compensation :  
évaluation **réussie** lorsque le total de points de toutes les questions est égal ou supérieur à la note de passage minimale requise.



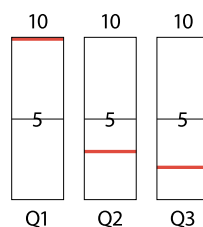
### Système de maîtrise

Nombre de questions : 3

Nombre de points (chacune) : 10

Note de passage minimale (par question) : 5/10

Application d'un système de maîtrise :  
évaluation **échouée** avec une note insuffisante pour les questions 2 et 3.



Les évaluations suivant le système de maîtrise sont généralement utilisées dans des contextes plus professionnels ou dans le milieu du travail, qui se prêtent mal à l'excellence dans un domaine et à la médiocrité dans un autre. Par exemple, pour confectionner un vêtement, aucun niveau élevé d'expertise en matière de stylisme ne peut compenser l'absence de compétences de base en couture.

L'IB s'appuie sur toute une gamme d'outils d'évaluation. Cela couvre notamment des examens destinés à être passés en fin de programme ainsi qu'une multitude de tâches d'évaluation (mémoires de recherche, travaux écrits, entretiens oraux, recherches scientifiques et mathématiques, rapports sur le travail de terrain et représentations artistiques) portant sur des matières très diverses et effectuées par les élèves à différents moments et dans des conditions variables au cours du programme.

## Les approches de l'évaluation

L'évaluation peut être utilisée dans une variété de buts. L'objectif visé par une évaluation donnée aura une grande influence sur la manière dont elle est conçue. Deux grands types d'évaluation sont habituellement reconnus : l'évaluation formative et l'évaluation sommative.

L'évaluation formative montre le lien le plus immédiat avec la manière dont les élèves apprennent, ce qui lui vaut parfois la qualification d'« évaluation **au service de** l'apprentissage ». L'évaluation sommative est quant à elle parfois considérée comme une « évaluation **de** l'apprentissage ». Une telle approche minimise

l'impact majeur qu'a l'évaluation sommative sur ce qui est réellement appris en classe. Toute évaluation doit apporter son soutien à un apprentissage approprié. L'évaluation sommative ne consiste pas uniquement en une activité succédant à l'apprentissage. Elle doit au contraire être conçue de manière à faire partie intégrante de l'apprentissage et de l'enseignement.

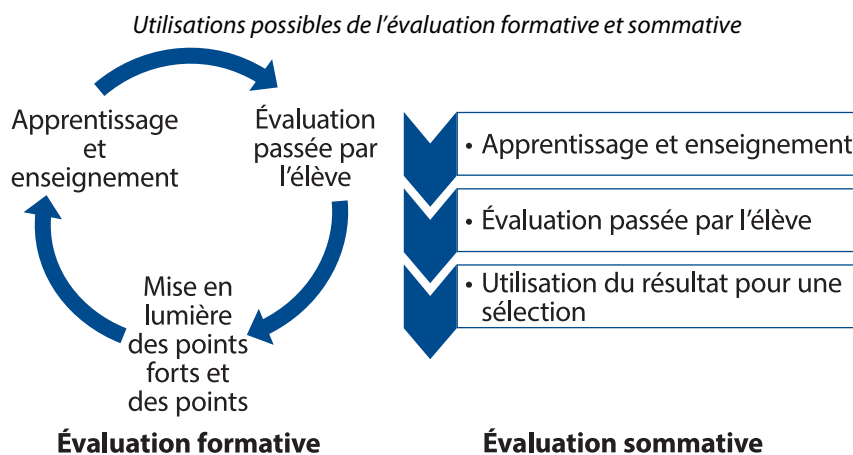
Pour l'évaluation formative, l'objectif principal est de fournir au personnel enseignant et aux élèves des commentaires détaillés quant à la nature des points forts et points faibles des élèves, et de les aider à développer leurs aptitudes. Les évaluations de type interaction directe, comme une discussion entre l'enseignant ou enseignante et l'élève, sont alors particulièrement utiles.

Vygotsky (1962) voit les enseignantes et enseignants comme des soutiens plutôt que comme des personnes qui dirigent l'apprentissage. Le personnel enseignant doit donc utiliser des tâches et des outils d'évaluation qui aident les élèves à travailler dans ce que Vygotsky appelle la « zone proximale de développement ». Il s'agit de l'écart (la gamme d'accomplissements) entre ce que l'élève peut faire en toute autonomie et faire avec l'aide de l'enseignant ou enseignante.

La notion d'« étayage » (ou soutien à l'apprentissage) a vu le jour avec Wood et coll. (1976). Durant le processus d'étayage, l'enseignant ou enseignante fournit l'étayage autour duquel l'élève pourra « construire » son apprentissage, mais cette construction ne peut être effectuée que par l'élève. L'objectif de l'enseignant ou enseignante doit être de préparer des évaluations formatives qui posent aux élèves des défis d'un niveau adéquat et de continuer à adapter ce niveau au fur et à mesure que les élèves progressent.

À l'inverse, l'évaluation sommative se concentre sur une mesure de **ce que** l'élève peut faire. Elle est habituellement utilisée pour montrer la réussite d'un programme de formation ou le degré de préparation au passage à l'étape éducative suivante. Alors que l'évaluation formative s'intéresse aux **raisons** pour lesquelles l'élève fait quelque chose, l'évaluation sommative cherche à savoir si l'élève a **correctement** fait les choses. Même si cela peut sembler moins utile que comprendre les raisons, il est intéressant d'étudier l'objectif particulier de l'évaluation sommative, qui consiste à émettre un jugement sur l'élève, et non à orienter l'enseignement à venir. La nécessité de l'évaluation sommative est analysée de façon plus poussée dans la section intitulée « La description de la réussite des élèves permise par l'évaluation sommative ».

Figure 7



Avec l'évaluation formative, il est plus important de repérer correctement les connaissances, les compétences et la compréhension que les élèves n'ont pas encore développées, que de mesurer avec précision le niveau de chaque accomplissement de l'élève. L'équilibre entre ces priorités porte le nom de « validité », une notion expliquée plus en détail par la suite. Cet équilibre entre l'accomplissement de l'élève et la qualité du retour d'information est inversé avec l'évaluation sommative, dans laquelle les résultats de l'évaluation sont utilisés pour prendre des décisions concernant l'élève, souvent pour une sélection compétitive en vue d'un emploi ou de possibilités d'études, mais aussi pour mieux décider de la poursuite de l'enseignement.

Toute analyse des différents systèmes d'évaluation nationaux révélera rapidement une grande variété de techniques et de méthodes d'évaluation. Tous ces systèmes ont leurs forces et leurs faiblesses en ce qui concerne les techniques, les ressources et le temps, ainsi que leur impact sur le système éducatif national auquel ils sont associés. Même s'il était possible, dans un contexte donné, de recommencer à zéro la conception d'un système d'évaluation, il n'existe pas de « meilleure technique » universelle qui pourrait être adoptée. Au contraire, les choix qui sont faits lors de la conception des systèmes d'évaluation reflètent inévitablement les valeurs et les priorités du contexte social plus vaste dans lequel ils sont faits. Pour approfondir ce sujet, veuillez vous référer à Cresswell (1996) et Broadfoot (1996).

Il convient également de noter que l'évaluation sommative est de plus en plus utilisée comme une mesure de la qualité de l'enseignement, ce qui ajoute une dimension supplémentaire aux raisons pour lesquelles l'évaluation est réalisée : elle se fait davantage au profit du système éducatif qu'à celui de l'élève.

## L'effet de remous et l'apprentissage

Nous avons besoin d'un processus d'évaluation qui soit aussi valide que possible. De plus, un tel processus ne doit pas, en raison de son effet de remous, fausser la qualité de l'enseignement, être trop lent ou absorber une trop grande quantité de nos ressources scolaires, qui sont peu abondantes.

(Peterson, 1971)

Dans cette citation, Alec Peterson, le fondateur de l'IB, reconnaît que la manière de mener l'évaluation à bien peut influencer la manière d'aborder l'apprentissage, et que cela présente un risque. Comme l'indique Surgenor (2010), cela peut avoir des effets positifs et négatifs sur l'expérience d'apprentissage et d'enseignement des élèves.

Dans un article publié en 1971, Snyder propose que les élèves créent leur propre compréhension du programme d'études à partir de messages implicites et explicites sur les éléments qui comptent pour l'évaluation, ce qu'il décrit comme le « programme caché ». Cela peut amener les élèves à comprendre comment réussir l'examen dans une matière, mais pas à comprendre la matière elle-même. Une autre compréhension possible apparaît dans la citation de Gibbs qui suit.

Lorsque j'ai repassé l'examen, j'ai uniquement cherché à le réussir. J'ai obtenu un résultat de 96 % et le type n'a pas compris pourquoi j'avais échoué la première fois. Je lui ai répondu que, cette fois, j'avais uniquement cherché à réussir l'examen et non à comprendre la matière. Étant donné que je ne comprends toujours pas cette matière, j'ai manqué mon objectif d'une certaine façon.

(Gibbs, 1992, p. 101, traduction libre)

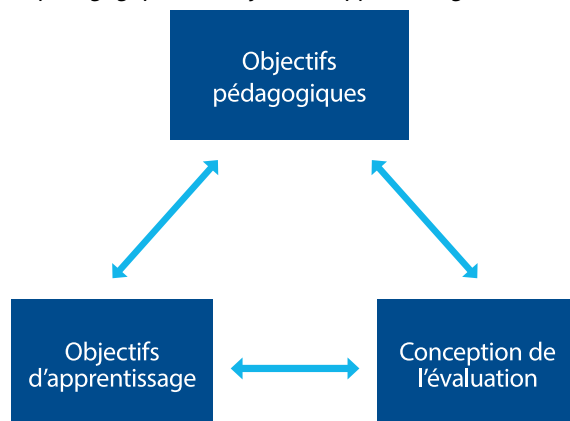
Ce concept de remous est également exprimé dans un adage souvent repris : « Si ce n'est pas évalué, ce ne sera pas enseigné. » Cela souligne l'idée que l'évaluation et l'enseignement ne peuvent pas être considérés comme deux choses indépendantes.

Bien que l'évaluation cherche à encourager les élèves à fournir des preuves de l'atteinte des objectifs pédagogiques, plusieurs raisons peuvent expliquer qu'elle n'y parvienne pas. La plus vraisemblable est que les objectifs d'apprentissage sur lesquels repose l'évaluation ne soient pas le reflet fidèle du but premier de l'éducation. Par exemple, un cours de formation pour athlètes de haut niveau pourrait avoir des objectifs d'apprentissage axés sur la compréhension des règles de l'athlétisme. Un autre problème récurrent est que, tandis que l'évaluation peut couvrir tous les objectifs visés, l'essentiel du test peut se concentrer sur des tâches faciles à évaluer, mais de moindre importance pour les objectifs pédagogiques globaux.

Cette interdépendance entre évaluation et but pédagogique est illustrée par la figure 8, inspirée du modèle de Furst (1958) qui est présenté dans Frith et Macintosh (1984). Tout élément qui ne serait pas relié à un autre impliquerait presque inévitablement une évaluation de mauvaise qualité.

Figure 8

Le lien entre les objectifs pédagogiques, les objectifs d'apprentissage et la conception de l'évaluation



## La validité

- Le terme « validité » désigne une évaluation **adaptée à l'objectif visé**. Il s'agit d'une notion complexe, qui possède de multiples facettes.
- Les évaluations sont utilisées à diverses fins.
- Une évaluation peut être valide pour une fin, mais pas pour une autre. Par exemple, elle peut tester la capacité à orthographier les mots, mais pas l'aisance d'expression.
- Il peut être considéré que ce sont les objectifs associés aux résultats de l'évaluation qui peuvent être valides ou non, plus que les évaluations elles-mêmes.
- La principale préoccupation de l'IB est de savoir si le programme est valide, puis si les éléments de ce programme (par exemple, les différents cours) sont valides et, enfin, si chaque évaluation est valide.

### Définition de la validité

Pourquoi passer des examens ? À quoi servent les examens ? Ces questions se trouvent au cœur du sens donné à une évaluation « adaptée à l'objectif visé » ou « mesurant ce qu'il faut » et bien souvent, les réponses apportées sont en contradiction les unes avec les autres. Par ailleurs, il peut arriver que tous les objectifs pédagogiques ne soient pas évalués par un examen ou que certains soient impossibles à évaluer.

Imaginez par exemple que les quatre possibilités suivantes soient données comme objectifs d'une évaluation en mathématiques :

1. déterminer la compréhension développée par l'élève après le programme ;
2. servir de système de sélection en vue de la poursuite des études ou de l'entrée dans le monde du travail ;
3. donner une indication de la réussite future ;
4. consolider l'enseignement des objectifs pédagogiques du programme.

Malgré leur simplicité, ces objectifs sont difficiles à concilier. Si la réussite future en mathématiques dépend plus de l'analyse mathématique que de la géométrie, l'accent doit-il être mis sur l'analyse mathématique ? Quel est le lien entre cet objectif et le premier, qui consiste à déterminer ce que l'élève sait (en géométrie) ? Que se passe-t-il quand deux de ces objectifs ou plus sont en contradiction avec la manière dont une évaluation doit être conçue ?

La liste des quatre objectifs possibles n'est aucunement définitive ni exhaustive. Pour Newton (2007), les résultats de l'évaluation peuvent être utilisés à de nombreuses fins. Il en donne les exemples qui suivent.

- |  |  |
|--|--|
| <ul style="list-style-type: none"> <li>• Utilisation à des fins d'évaluation sociale</li> <li>• Utilisation à des fins formatives</li> <li>• Utilisation à des fins de contrôle de la progression de l'élève</li> <li>• Utilisation à des fins de transfert</li> <li>• Utilisation à des fins de répartition</li> <li>• Utilisation à des fins de diagnostic</li> <li>• Utilisation à des fins d'orientation</li> <li>• Utilisation à des fins de mesure de la compétence</li> </ul> | <ul style="list-style-type: none"> <li>• Utilisation à des fins de permis</li> <li>• Utilisation à des fins de décisions scolaires</li> <li>• Utilisation à des fins de contrôle de l'établissement</li> <li>• Utilisation à des fins d'affectation des ressources</li> <li>• Utilisation à des fins d'intervention au niveau de l'organisation</li> <li>• Utilisation à des fins d'évaluation du programme</li> <li>• Utilisation à des fins de contrôle du système</li> <li>• Utilisation à des fins de comparabilité</li> </ul> |
|--|--|

- Utilisation à des fins de sélection

En dépit de cet étalage d'utilisations contradictoires, les personnes et les établissements doivent prendre des décisions qui concernent des individus dans des contextes variés. Si les résultats de l'évaluation ne sont pas disponibles, alors d'autres paramètres – moins bien conçus et moins impartiaux – risquent d'être utilisés pour cette prise de décision. Pour souligner ce point, Cresswell (1986) définit le problème de la manière suivante.

À l'évidence, si les autres critères sont moins fiables que les examens, compter davantage sur ces critères signifiera faire reposer la sélection sur des décisions moins fiables.

(Cresswell, 1986, p. 42, traduction libre)

Cette idée de diplôme ou certificat (ou de composant d'un diplôme ou certificat, comme une évaluation) « adapté à l'objectif visé » est englobée dans la notion de validité.

La **validité** est souvent perçue comme un terme incontournable pour désigner la qualité d'une évaluation, mais les universitaires cherchent depuis longtemps à en donner une définition précise (Newton, 2012). Et pour compliquer les choses, il existe une notion plus étroite de la validité, souvent associée au terme « **fiabilité** », qui décrit la mesure dans laquelle une évaluation parvient à mesurer réellement ce qu'elle est censée mesurer. Pour éviter toute confusion dans la présente publication, l'expression « pertinence du construit » a été choisie pour évoquer ce second concept. Veuillez vous référer à la section intitulée « [La pertinence du construit et l'authenticité](#) ».

En fin de compte, déterminer la validité d'une évaluation donnée est une question de jugement. Ce jugement repose sur le recueil systématique des preuves qui aideront à interpréter les notes obtenues aux tests, pour les usages possibles qui leur ont été associés. Les exigences contradictoires de l'évaluation qui ont été énumérées précédemment seront traitées plus en détail par la suite, mais il faudra toujours parvenir à un compromis entre ces exigences. Il convient de se demander si les preuves disponibles suggèrent que l'évaluation est suffisamment adaptée à l'objectif visé pour être utile. Par conséquent, il est recommandé de réfléchir à la validité en s'intéressant à la solidité des preuves – c'est-à-dire à la solidité de l'argument de validité qui est avancé – plutôt que de se contenter d'une simple déclaration établissant la validité ou l'invalidité d'une évaluation. Déterminer la validité de quelque chose impose donc de porter un jugement à partir de preuves. La validité n'est pas une qualité binaire : elle dépend du contexte et du but visé. Newton et Shaw expliquent qu'il faut s'intéresser à l'adéquation de l'objectif d'une évaluation avant de préparer une déclaration concernant sa validité, et que la validation s'entend du processus consistant à développer et à juger la force d'une argumentation sur l'interprétation et l'utilisation des notes obtenues aux tests (Newton et Shaw, 2014, p. 3). La validité n'est donc pas l'apanage de l'évaluation elle-même, mais des déductions qui sont faites des résultats de l'évaluation.

Enfin, la validité n'est pas quelque chose qui est atteint lors de la conception d'une évaluation. Elle se développe au contraire en permanence au cours de son cycle de vie. De même, l'argument de validité n'est pas avancé au début du cycle : il est constamment nourri et précisé au cours du processus de conception d'une évaluation.

## La création d'un argument de validité

La validité n'est pas un concept objectif simple : il s'agit d'un juste milieu entre des questions opposées. Il est impossible de « prouver » la validité. Il faut plutôt construire une argumentation convaincante qui montre pourquoi les décisions prises au sujet d'un cours ou d'une évaluation justifient sa pertinence et son adéquation avec l'objectif visé. Ce n'est donc pas l'évaluation (ni le résultat) qui est valide, mais le but qu'elle poursuit. Newton (2012) pense que l'utilisation d'une procédure d'évaluation donnée dans un but spécifique (c'est-à-dire pour prendre une décision particulière) est valide si son argument d'interprétation est suffisamment puissant.

Pour des raisons similaires, la preuve d'un argument de validité doit émaner naturellement des processus d'élaboration, de contrôle et d'enseignement d'un programme, et rendre compte des discussions et des décisions prises au cours de ces processus. Pour prendre un exemple dans le domaine de l'évaluation, les jugements à l'origine des questions posées dans un examen donné devront prouver que le programme d'études est convenablement couvert.

Il est indispensable que tout argument de validité ait une structure appropriée couvrant tous les aspects considérés comme importants. À cet effet, l'IB a formulé une série de questions pour lesquelles des preuves sont réunies.

## Le maintien de la validité

La validité n'est pas quelque chose qui se décide ou se « prouve » au moment de concevoir un modèle d'évaluation. Elle se développe au contraire avec l'évaluation, au cours de son cycle de vie, et tant que les décisions prises sont basées sur son résultat.

L'IB rassemblera différentes preuves pour démontrer qu'une évaluation donnée est valide (c'est-à-dire plusieurs arguments de validité), à plusieurs moments du cycle de vie d'un cours.

Lorsqu'un cours est élaboré ou révisé, les discussions sur le but du cours et la manière de l'évaluer pour atteindre ce but constitueront la partie centrale de l'argument de validité. Une attention toute particulière sera portée à l'équilibre entre la pertinence du construit et les autres aspects de la validité. À chaque session, la rédaction d'un nouvel examen ou d'une nouvelle tâche d'évaluation fournira des preuves supplémentaires, notamment dans le domaine de la fiabilité, de l'équité, de la comparabilité, de la facilité d'application et de la pertinence du construit. Les retours d'information du personnel enseignant et les résultats des élèves permettront donc de juger dans quelle mesure une évaluation parvient à atteindre son objectif, mais aussi à apporter des preuves de la fiabilité de la notation et de la comparabilité des notes finales. Ces informations seront utilisées pour orienter le développement de la prochaine série d'évaluations.

Enfin, les preuves rassemblées dans toutes les évaluations concernant un cours particulier, et dans tous les cours formant un programme donné, serviront de base aux décisions pour la prochaine révision du cours et du programme.

L'évaluation de la validité d'un test n'est pas une démarche statique et unique, mais un processus continu.

(Sireci, 2007, p. 477, traduction libre)

## Les éléments de la chaîne de validité

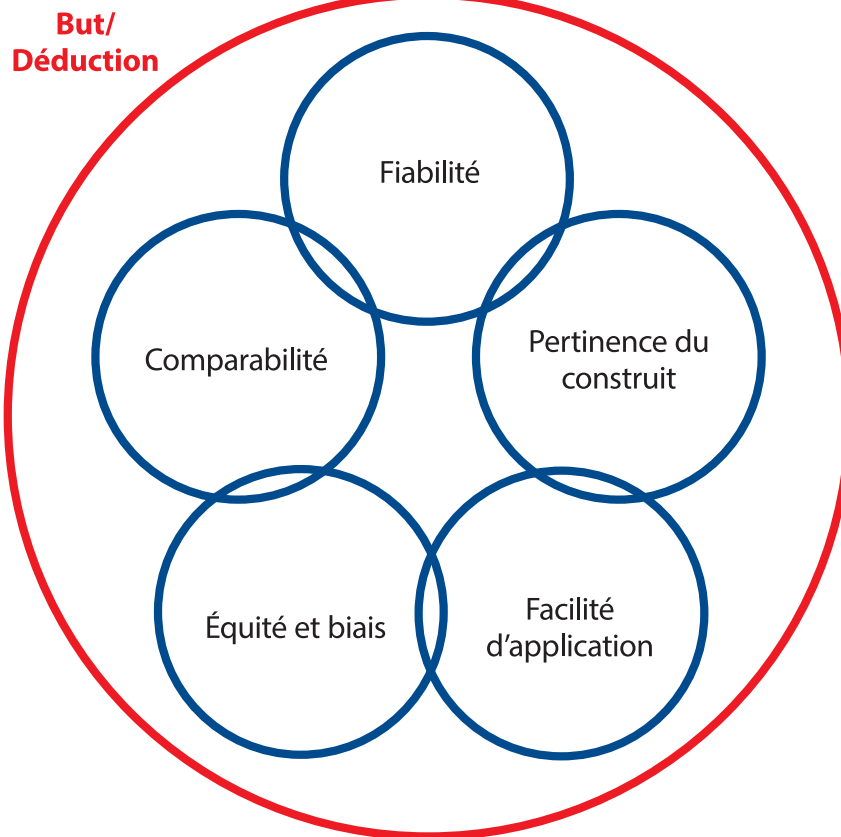
- La validité comporte de nombreux éléments. En imaginant qu'ils forment ensemble une chaîne, si un seul maillon est cassé, alors c'est toute l'évaluation qui est détruite.
- L'IB se concentre généralement sur cinq éléments de la validité : la fiabilité, la pertinence du construit et l'authenticité, la facilité d'application, l'équité et le biais, et la comparabilité.
- Ces cinq éléments entrent souvent en contradiction les uns avec les autres. Pour savoir si une évaluation est valide, il faut s'intéresser au but premier de l'évaluation afin de déterminer l'importance relative de chaque élément.
- L'IB s'attache avant tout à concevoir des programmes, des cours et des évaluations qui prouvent la pertinence du construit.
- Dans le domaine de l'évaluation, cela signifie que l'IB place l'accent sur l'élaboration de tâches et de questions pertinentes, qui mesurent les compétences de pensée de haut niveau, plutôt que sur la création d'exams garantissant une notation fiable.

La validité et la fiabilité sont largement considérées comme des caractéristiques essentielles de tout système d'évaluation, et les enjeux sont d'autant plus élevés que le résultat revêt une grande importance pour l'élève ou l'enseignant ou enseignante. Les caractéristiques sont en réalité multidimensionnelles, et les types de validité et de fiabilité sont nombreux.

L'idée de la chaîne montre la complexité de la **validité** (Crooks et coll., 1996). Chacun des cinq maillons de la chaîne revêt une importance en soi, mais aucun ne suffit à lui seul à rendre l'évaluation valide (c'est-à-dire adaptée à l'objectif poursuivi). Ainsi, une évaluation peut être très fiable, mais désavantager systématiquement un groupe donné. Il est également possible qu'une tâche atteigne précisément l'objectif visé, mais qu'elle mette également à rude épreuve l'endurance des élèves et les ressources de l'établissement en raison de sa durée et de ses exigences. Les sections suivantes donnent plus de précisions sur chacun des éléments de cette chaîne. Même si tous sont nécessaires pour parvenir à la validité, il convient de souligner qu'il existe une certaine tension entre eux. Par exemple, même si un examen est supposé aborder tous les aspects du programme d'études, cela peut facilement mener à des tests trop longs pour les élèves. De la même manière, la contextualisation du test en fonction de chaque culture ou pays où il est passé (équité) peut amener à se demander si le même test est réellement proposé à chaque fois (comparabilité).

Pour qu'une évaluation soit valide, tous les maillons de la chaîne doivent coexister.

Figure 9  
La chaîne de validité






## Trouver un équilibre entre les différents aspects de la validité

Lors de la conception d'une évaluation, il est important de trouver un équilibre entre les aspects contradictoires de la validité en tenant compte du but visé par l'évaluation. Aucune évaluation ne peut satisfaire à la norme la plus élevée de tous les éléments ; c'est pourquoi il faut parvenir à un compromis. Alors que certains de ces éléments sont fixés lors de la conception de l'évaluation, d'autres (notamment la fiabilité et l'équité) évoluent avec l'utilisation de l'évaluation, la notation et l'attribution des notes finales.

Figure 10

Trouver un équilibre entre les priorités contradictoires de plusieurs aspects de la validité : l'équité et la facilité d'application

|   |   |  |
|---|---|--|
|    |    |   |
| <b>Mauvais équilibre</b><br>La charge pesant sur les élèves est excessive (problème de facilité d'application).   | <b>Bon équilibre</b><br>Les modalités d'évaluation des élèves sont raisonnables.  | <b>Mauvais équilibre</b><br>Il y a une trop grande part de chance dans le choix du thème évalué (problème d'équité).   |
| L'évaluation est composée de 50 tâches (de 20 minutes chacune) requérant une réponse développée et traitant tous les aspects du cours.<br><br>Cela représente plus de 16 heures d'évaluation. | L'évaluation comporte 20 questions à réponse brève (de 2 minutes chacune), 5 tâches plus longues (de 10 minutes chacune) et 2 tâches approfondies (de 30 minutes chacune). Chaque question porte sur un aspect différent du cours.<br><br>L'évaluation dure 2 heures 30 en tout, et porte sur un peu plus de la moitié des thèmes du cours. | L'évaluation consiste en une tâche approfondie de 30 minutes, qui mesure à la fois les connaissances et la compréhension d'un aspect du cours (sur les 50 disponibles).<br><br>La note finale de l'élève repose exclusivement sur une tâche. |

Malgré leurs définitions distinctes, il existe des recoupements importants au niveau de l'intervention et de la gestion de ces cinq éléments. Lorsqu'ils sont associés, ils constituent le concept de validité au sens large qui se retrouve dans les évaluations de l'IB.

Enfin, l'IB accorde une priorité absolue à la **validité de construit**, c'est-à-dire que ses évaluations mesurent les caractéristiques et les capacités qu'elles sont censées mesurer. Malgré cela, il convient de rappeler que ce sens des priorités ne peut se faire totalement au détriment des autres aspects de la chaîne de validité.

## La fiabilité

Dans le domaine de l'évaluation, la fiabilité se définit comme la mesure dans laquelle l'élève pourrait obtenir le même résultat à un test si la procédure employée était répétée. Cela n'est pas nécessairement la même chose que l'élève obtenant le « bon » résultat.

Dans leur introduction au concept de fiabilité, Winkley et Cresswell (2011) énumèrent les domaines dans lesquels la fiabilité est souhaitée (notamment la notation des résultats des élèves) et donnent des exemples qui révèlent une liste de menaces potentielles pour la fiabilité.

1. **La fiabilité entre les correcteurs ou correctrices :** Un examinateur ou une examinatrice pourrait faire preuve de plus ou moins d'indulgence qu'une autre personne pour certaines questions, voire montrer un degré d'indulgence différent d'un jour à l'autre.
2. **La variabilité des résultats de l'élève :** les résultats de l'élève à un examen peuvent varier d'un jour à l'autre, notamment si les conditions d'examen sont différentes. Ces paramètres variables peuvent être le moment du déroulement de l'examen (matin ou après-midi), la personne faisant passer le test, la qualité du sommeil la veille de l'examen, le bruit d'une tondeuse à gazon pendant l'examen.
3. **La différence entre les examens :** les questions sont différentes d'un examen à l'autre et risquent de tester des aspects différents de la compréhension de l'élève (les tests portent généralement sur un échantillonnage de thèmes du programme d'études en raison d'un manque de temps pour tester la totalité, et les élèves peuvent choisir de ne pas réviser tous les thèmes).

4. **La comparabilité des résultats d'une année à l'autre** : cette comparabilité garantit une comparaison pertinente entre les résultats des élèves indépendamment des années concernées.
5. **Les différences entre les modalités des examens** : il peut être difficile de garantir la comparabilité dans le temps lorsque des changements sont apportés aux modalités des examens et aux programmes.
6. **Les différents types d'activités d'évaluation** : de nombreuses évaluations sont composées de plusieurs types d'activités. Ces différentes méthodes d'évaluation peuvent entraîner plusieurs sortes de problèmes de fiabilité.
7. **Les différents types de questions** : les élèves pourront obtenir des résultats différents en fonction des types de questions à traiter.

Dans la pratique, l'IB part généralement du principe que l'élève passe l'évaluation une seule fois, et recherche **l'homogénéité** du résultat du processus. La priorité est donc donnée aux points 1, 3, 6 et 7 de la liste.

En tant qu'organisme décernant des titres, l'IB prend des mesures pour accroître le niveau de fiabilité de ses évaluations. La fiabilité de la notation occupe une place centrale dans cette démarche, et la normalisation, le modèle d'assurance de la qualité et la révision de notation visent à garantir que les examinateurs et examinatrices procèdent à la notation selon une norme cohérente.

## La fiabilité de la notation

Le rapide exercice présenté ci-dessous illustre le concept de fiabilité de la notation et son fonctionnement dans la pratique.

La figure 11 rassemble cinq extraits de réponses authentiques d'élèves à une évaluation en Acquisition de langues de l'IB.

Utilisez le barème de notation pour juger de la qualité de langue démontrée par les élèves et consignez la note sur 8 que vous attribueriez à chaque travail.

Après avoir noté chaque travail, consultez le retour d'information sur la qualité de votre notation. Ce retour d'information est fourni plus bas, après les copies des élèves.

### Barème de notation

Avec quel degré d'efficacité et de précision est-ce que l'élève applique les compétences de réflexion ?

| Nombre de points | Descripteurs de niveaux   |
|------------------|---|
| 0                | L'élève n'atteint aucun des niveaux décrits ci-dessous.   |
| 1 – 2            | L'élève indique les effets du projet sur son apprentissage et précise si les objectifs ont été atteints.  |
| 3 – 4            | L'élève résume les effets du projet sur son apprentissage et précise si les objectifs ont été atteints, en fournissant partiellement des preuves à l'appui.   |
| 5 – 6            | L'élève décrit les effets du projet sur son apprentissage et évalue le résultat en fonction des critères de réussite, en fournissant des preuves à l'appui.   |
| 7 – 8            | L'élève explique les effets du projet sur son apprentissage et évalue le résultat en fonction des critères de réussite, en fournissant systématiquement des preuves spécifiques et/ou des exemples détaillés à l'appui. |

Visualisez les copies une par une et attribuez une note à chacune en vous aidant du barème de notation.

Figure 11A

Exercice sur la fiabilité de la notation

Dans notre monde tellement dominé par les réseaux sociaux, il me semble que les gens ne savent plus distinguer le vrai et le faux qui s'affiche devant leurs yeux sur l'écran de leurs smartphones. Malheureusement, la superficialité domine, pas l'information... c'est triste.

Aujourd'hui au collège, i.e. parlais avec Pierre et Louis de l'article sur les extraterrestres qui est devenu "viral" sur Facebook l'autre jour. C'était un article qui essaye de convaincre le public que les aliens ont fait contact avec nous et qu'ils aiment bien le climat de l'Alaska... grosse blague.

Figure 11B

Exercice sur la fiabilité de la notation

La réflexion « Dans les réseaux sociaux, c'est la superficialité qui domine, pas l'information reliée au thème communication et média, car ça concerne le réseau social, et relié aussi à question mondiale car l'information concerne tout le monde.

Pour commencer, je pense que le réseau social aide beaucoup pour faire connaître tout qui entraîne de passer au temps réel, ça aide à donner tout l'information, le danger, l'histoire, etc. Par exemple si quelqu'un qui vit pas en Amérique veut savoir comment la votation est passée, le réseau social a permis à tout le monde de suivre sur l'internet.

Figure 11C

## Exercice sur la fiabilité de la notation

C'est vrai qu'avec Instagram, par exemple, il y a un grand aspect de superficialité, où l'on veut tous créer l'image d'un vie 'parfaite'. C'est-à-dire seulement oublier des photos de nous sur son site. Par exemple, pendant les attentats à Paris et en Florides, il était possible de se marquer 'safe' (en sécurité), et ceci notifie les amis. Du coup, Facebook a utilisé des paramètres pour informer les amis et les familles que la personne qui était près du danger était en sécurité.

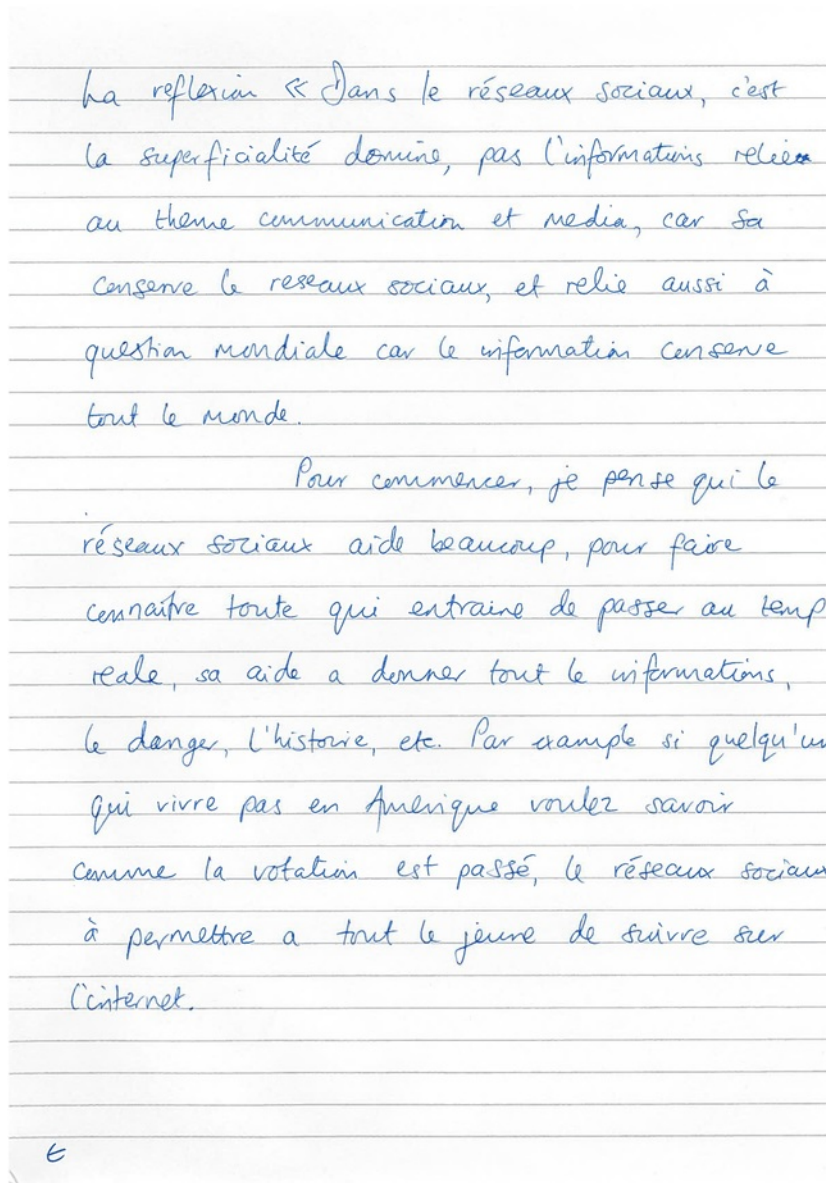
Figure 11D

## Exercice sur la fiabilité de la notation

En revanche, quand on sait a utilisé les réseaux sociaux responsable dans une manière responsable, l'un peu dire que l'individu pourrai se sauver d'être un accro à les réseaux sociaux par l'utilisation <sup>en moderation</sup> modérée. Le savant va réfléchir l'impact des images ou informations il choisit a partager. Parfois les gens veuille la sensation dans leur réseaux sociaux car c'est intéressant pour consumer, alors, des sites, comme «wise cracks» nous fournissent avec des information culturelle et actuelle en forme comédie mais équilibré avec les faits réelle.

Figure 11E

Exercice sur la fiabilité de la notation



Lorsque vous aurez terminé votre notation de chaque copie et attribué une note à chaque élève, vous pourrez lire le retour d'information sur votre notation.

### Retour d'information

Avez-vous remarqué que le travail de l'élève B et de l'élève E était le même, mais avec une écriture différente ? Avez-vous donné la même note à l'élève B et à l'élève E ?

Vous avez peut-être aussi hésité entre les deux notes pour une bande de notation donnée. Dans quelle mesure avez-vous la certitude de prendre la même décision si vous deviez noter de nouveau ces travaux ?

Ce qui est en cause ici n'est pas la qualité du travail de l'élève, mais la façon dont la même personne peut formuler des jugements légèrement différents à chaque fois qu'elle note ce travail. Globalement, votre opinion sur la qualité du travail restera la même à chaque fois, mais la note exacte que vous attribuerez pourra varier légèrement.

Refaites cet exercice dans quelques jours en notant de nouveau chaque copie, et comparez cette deuxième notation à la première. (Essayez de ne pas vous souvenir des notes que vous aviez données initialement.)

Pour obtenir de plus amples informations sur la fiabilité que l'IB confère à la notation, veuillez vous référer à la vidéo intitulée « [Fiabilité : comment garantissons-nous une notation équitable ?](#) ».

## Les résultats cohérents et les « bons » résultats

L'objectif consistant à obtenir un niveau de fiabilité élevé permet de garantir que les élèves auront la même note (juste) quelle que soit la personne qui passe leur travail en revue, et non la « bonne » note. Pour déterminer la qualité d'un travail, l'enseignant ou enseignante s'appuie sur son jugement professionnel, et deux personnes auront souvent du mal à tomber d'accord sur la note à attribuer. La fiabilité intervient lorsque toutes les deux émettent le même jugement (identique à celui de l'examineur superviseur ou examinatrice superviseuse).

Cela présente donc des difficultés particulières lors du traitement des réclamations concernant les résultats. Dans ce cas, l'examineur ou examinatrice doit veiller à attribuer la même note que celle qu'il ou elle aurait attribuée si la copie avait été la première notée, et s'assurer que cette note n'est pas influencée (ni positivement ni négativement) par de nouvelles informations, comme les seuils d'attribution des notes finales ou les conséquences pour l'élève en question.

La mauvaise compréhension de la fiabilité de l'évaluation dans des contextes autres que l'éducation a fait l'objet de nombreuses publications, mais les débats sur les résultats aux examens étant de plus en plus courants, une importance particulière doit être accordée à ce thème.

Comme la littérature semble l'indiquer, les participantes et participants ont estimé que la fiabilité de l'évaluation, et plus précisément l'inexactitude de la mesure, étaient des concepts difficiles à saisir.

(Chamberlain, 2010, p. 3, traduction libre)

Je me dis parfois que les examens ne permettent pas de jauger l'aptitude d'une personne à sa juste valeur. Ai-je tort ? Ils reflètent davantage comment la personne allait à ce moment précis, sa faculté à étudier, etc. Ce n'est pas parce qu'une personne obtient d'assez mauvais résultats aux examens qu'elle ne pourrait exceller par la suite et compter parmi les meilleures de son domaine. (Propos d'un employé du secteur de la santé)

(Chamberlain, 2010, p. 27, traduction libre)

Environ 63 % du personnel enseignant et des élèves ont choisi « Aucun niveau d'erreur ne doit être acceptable ; même le fait qu'une seule personne obtienne une note finale incorrecte est totalement inacceptable ». Mais plus de 50 % ont également choisi « Il existe une différence entre une erreur évitable – comme une coquille dans un examen – et quelque chose d'inévitable comme un manque de cohérence entre deux correcteurs ou correctrices », ce qui suggère une tolérance envers l'erreur. Cette incohérence peut illustrer la faiblesse du lien entre les connaissances sur la fiabilité et les attitudes envers le manque de fiabilité.

(He et coll., 2010, p. 27, traduction libre)

## La pertinence du construit et l'authenticité

Avec quelle précision la chose censée être mesurée est-elle véritablement mesurée ? La mesure précise des compétences et connaissances censées être mesurées est parfois désignée par l'expression « validité de construit ». Cependant, pour éviter toute confusion avec la notion générale de validité, la présente publication utilise uniquement l'expression « **pertinence du construit** ».

L'idée d'évaluation authentique est étroitement liée à la notion de pertinence du construit. L'**authenticité** signifie que les modalités du test correspondent, dans la mesure du possible, aux situations dans lesquelles l'élève pourrait s'attendre à rencontrer le problème dans la vie réelle. Les évaluations qui isolent les tâches de leur contexte, qui simplifient les choses à outrance ou qui sont manifestement artificielles, sont des exemples d'évaluations non authentiques.

Il est facile de trouver des exemples d'évaluations mal conçues, avec un faible degré de pertinence du construit (par exemple, évaluer la capacité à rédiger une lettre en utilisant une épreuve orale). Or souvent, même une méthode reconnue en matière d'élaboration des tâches ne permet pas d'évaluer véritablement l'objet visé. Prenez par exemple la conception de certains examens de littérature traditionnels. Les élèves ont étudié les procédés littéraires qui sont utilisés dans un texte donné, puis l'évaluation leur demande de

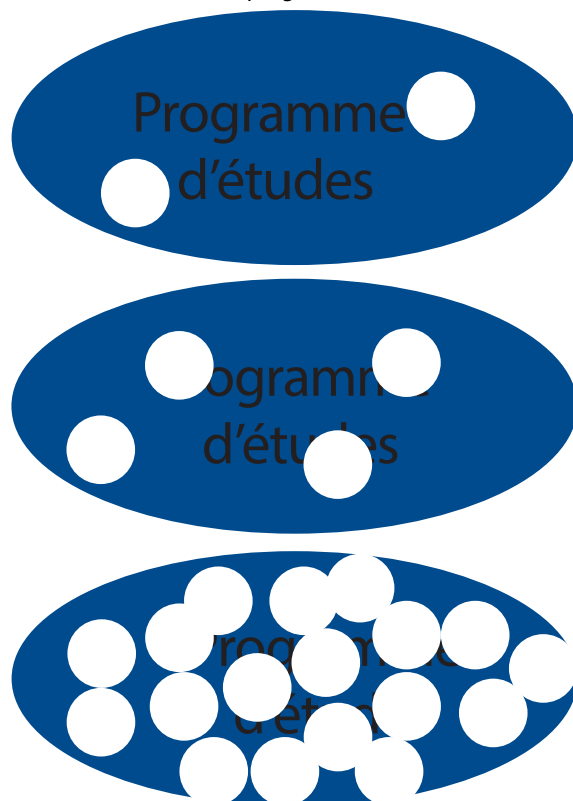
rédigier une dissertation en rapport avec l'un de ces procédés. Si les élèves doivent simplement se souvenir de ce qui a été appris, leur est-il réellement possible de montrer une compréhension de la littérature ?

Bien comprendre ce qu'une tâche d'évaluation donnée souhaite permettre de montrer, et analyser avec soin dans quelle mesure elle parvient à le faire, constitue la meilleure pratique pour s'assurer de l'adéquation du degré de pertinence du construit. Il est tout particulièrement important de réfléchir aux autres compétences dont les élèves ont besoin pour réaliser la tâche. Ainsi, les tâches ouvertes, comme les dissertations ou les projets, nécessitent très souvent un haut niveau de compétence en matière de rédaction pour que les élèves puissent montrer les compétences de recherche ou d'analyse visées par la tâche. Cette idée est reprise dans le concept de conception universelle de l'évaluation, qui s'attache à réduire les obstacles pour les élèves lorsque les difficultés n'ont aucun rapport avec ce que l'évaluation doit évaluer (Dolan et coll., 2013).

La pertinence du construit est étroitement liée au **programme d'études**. Très souvent, l'évaluation interroge l'élève uniquement sur une partie limitée du programme d'études. La question consistant à se demander si ce qui est traité est suffisant relève de la pertinence du construit et du jugement. De même, le choix des types de questions qui seront posées constitue un élément important de la conception de l'évaluation, car chaque type de question ou de tâche peut évaluer différents types de construit.

Figure 12

*Les examens et le programme d'études traité*



## La facilité d'application

La facilité d'application a fait l'objet de peu de recherches et il n'existe pas de définition ni d'approche unique de sa mesure. En termes généraux, la « facilité d'application » désigne l'effort requis pour passer une évaluation, du point de vue de l'élève, de l'établissement scolaire et de l'IB.

Pour les **élèves**, la facilité d'application fait souvent référence à l'effort requis pour réussir une évaluation. Ainsi, un examen de 8 heures est considéré comme une exigence excessive en dessous de 18 ans. Évaluer la facilité d'application pour les élèves signifie également tenir compte de l'heure à laquelle a lieu l'évaluation et de la durée de la période d'examen sur la totalité de la journée des élèves. L'IB doit faire particulièrement attention à la charge cumulée induite par les différentes évaluations de l'IB que les élèves

effectueront dans le cadre de leurs études, ce qui ramène à l'idée qu'il ne faut pas considérer uniquement la validité de chaque **matière**, mais celle des **programmes**.

Il peut être demandé à un **établissement scolaire** de fournir du matériel pour l'évaluation, ce qui aura des répercussions sur la facilité d'application de cette évaluation. Ainsi, une formation professionnelle en ingénierie pourrait exiger que chaque élève assemble un moteur. En fonction du nombre d'élèves, le nombre de moteurs nécessaires pourrait devenir ingérable.

Un autre aspect de la facilité d'application pour les établissements scolaires concerne la manière dont les travaux des élèves sont transmis à l'IB en vue de la notation. L'obligation d'enregistrer ou de filmer un exposé impose davantage de contraintes à l'établissement que l'envoi d'un travail écrit.

Enfin, l'IB prête également une dimension à la facilité d'application à son propre niveau, en raison de la quantité de travaux d'élèves envoyés pour l'évaluation qui sont passés en revue par l'organisation. Ainsi, assister à une production théâtrale de trois heures peut constituer la meilleure preuve de l'aptitude de l'élève, mais cette pratique n'est pas envisageable si la tâche doit être évaluée en externe. Prévoir un large éventail de questions facultatives pourrait aussi avoir des répercussions sur la facilité d'application. En effet, il ne serait possible de parvenir à une norme commune entre ces questions qu'à l'issue d'un travail long et minutieux.

Il est souvent difficile de concilier la **facilité d'application**, la **fiabilité** et la **pertinence du construit**. Toute augmentation de la quantité de l'évaluation passée (ou de la durée d'un examen donné) fournirait davantage de preuves de la compréhension qu'à l'élève du programme d'études. Cela augmenterait la vraisemblance d'un lissage entre la générosité et la sévérité des décisions en matière de notation, ce qui améliorerait la fiabilité. Cependant, une telle augmentation conduirait également à une évaluation mesurant l'aptitude des élèves à maintenir leur rendement pendant un examen long au lieu des objectifs de l'évaluation, ce qui réduirait la pertinence du construit et la facilité d'application pour les élèves.

L'IB procède à des contrôles stricts de la facilité d'application de l'évaluation, notamment pour ce qui est de la quantité totale de l'évaluation dans chaque cours.

## L'équité et le biais

Un test est biaisé s'il confère un avantage ou désavantage involontaire à l'élève. Voici quelques exemples d'examens biaisés :

- un examen d'histoire entièrement rédigé en latin ;
- un examen de mathématiques dans lequel toutes les questions se rapportent au marquage de points dans un match de cricket ;
- un examen d'arts pour lequel le chevalet de peintre est fixé à deux mètres du sol.

Dans chaque situation, les tâches peuvent considérablement **défavoriser** des élèves en particulier (les jeunes de petite taille pour le dernier exemple). Dans la réalité, les exemples de biais sont généralement plus subtils que ceux donnés précédemment, mais, à moins d'y faire véritablement attention, le biais peut provoquer des différences importantes au niveau des résultats des élèves. Placer les questions dans un contexte totalement exempt de biais se révèle particulièrement difficile, car les situations familières à des élèves peuvent être inconnues d'autres élèves, ce qui se vérifie d'autant plus avec la dimension internationale de l'IB.

Le « biais » peut être défini comme une différence dans le résultat d'un processus d'évaluation qui n'est pas due à une réelle différence dans l'aptitude ou l'accomplissement mesuré. Le biais peut venir des **conditions** de l'évaluation, de la **notation** de l'évaluation (dans ce cas, le biais devient un problème de fiabilité de la notation) ou directement des **tâches** d'évaluation.

### Le biais venant des conditions de l'évaluation

Les exemples donnés au début de cette section comportent un biais au niveau des conditions, parfois improbables comme le chevalet placé trop haut pour des élèves en particulier. De nombreuses conditions sont susceptibles d'entraîner un biais pour une évaluation, et plus précisément pour un examen. Beaucoup concernent l'horaire de l'examen. Un biais serait effectivement induit avec des examens prévus alors que les

températures ou le niveau de pollution sont extrêmes dans certaines parties du monde, ou à un moment où des élèves ne peuvent pas se concentrer ni se préparer convenablement. Cela pose un défi majeur à l'IB, étant donné que les exigences contradictoires de certains pays et régions (ou territoires) impliquent qu'aucune date n'est idéale pour au moins quelques établissements.

L'agencement de la salle d'examen peut lui aussi constituer un biais fréquent. Prenez le cas de deux places occupées par des élèves, l'une directement exposée au soleil et la seconde, située à l'ombre. Le manque de cohérence dans l'application du règlement de l'examen peut fournir un autre exemple : un respect scrupuleux des délais impartis peut parfois laisser la place à une certaine souplesse. Pour pallier cette situation, l'IB s'appuie sur des règles claires et cohérentes rassemblées dans les *Procédures d'évaluation* spécifiques à chaque programme, et considère tout manquement à ces règles comme une mauvaise administration.

## Le biais venant de la notation

Le biais venant de la notation est bien plus inconscient que délibéré. Par nature, l'esprit humain utilise des raccourcis pour faciliter la prise de décision, et il en résulte souvent des biais inconscients. Même si l'IB a le devoir de limiter ce biais, aucune faute ne lui est imputée.

Lorsque le biais se manifeste lors de la notation, il peut être dû à un certain nombre de raisons, telles que l'attitude personnelle à l'égard de la lisibilité (ou d'une autre caractéristique) de l'écriture de l'élève (voir par exemple, Hughes et coll., 1983), un traitement préférentiel en fonction du genre de l'élève (lorsque celui-ci est connu au préalable ou deviné par l'examineur ou l'examinatrice) et une attention excessive portée à des facteurs comme le formatage, la ponctuation et l'orthographe, qui ne sont pas très pertinents dans certains contextes d'évaluation. De tels problèmes peuvent être résolus grâce à la formation des examinateurs et examinatrices, et à la vérification de leur travail.

Le biais inconscient fondé sur des facteurs tels que le genre, la nationalité ou l'établissement est également bien documenté. Pour le minimiser, l'IB cherche à rendre tous les travaux envoyés pour l'évaluation anonymes avant qu'ils soient notés.

Un autre domaine de biais bien étudié est connu sous le nom d'« effet de halo ». Cet effet peut être observé lorsqu'un examinateur ou une examinatrice développe une opinion positive sur l'élève en raison de la bonne qualité de ses premières réponses et que le bénéfice du doute devient disproportionné pour les questions suivantes.

Il peut arriver que des examinateurs ou examinatrices déclarent ne pas pouvoir noter équitablement et sans biais une réponse concernant un thème précis. Le cas échéant, l'IB dispose de systèmes permettant de faire évaluer ce travail par une personne à l'abri de tout biais et capable de tenir compte des seules valeurs scolaires du travail présenté.

## Le biais associé aux tâches d'évaluation

Le biais s'applique si la différence observée dans les résultats est sans lien avec l'objet mesuré. Ainsi, utiliser un problème mathématique s'inscrivant dans un exemple culturel spécifique pourrait faire obstacle à la compréhension et à l'accomplissement (qui figurent parmi les modalités de l'évaluation) si l'élève ne dispose pas des connaissances requises sur la tradition culturelle en question (qui ne figurent pas parmi les modalités de l'évaluation).

Lors de l'élaboration des tests psychométriques, tout élément de question pour lequel une réponse peu habituelle est donnée durant le test préliminaire, ou pour lequel les différents sous-groupes d'élèves donnent une réponse très différente (ce que l'on appelle « fonctionnement différentiel des éléments »), est considéré comme biaisé et il est supprimé du test. Les sous-groupes d'élèves peuvent être formés en fonction du genre, de l'appartenance ethnique, de la classe sociale ou de la compétence linguistique, mais aussi en fonction de n'importe quelle caractéristique n'ayant aucun rapport avec le construit faisant l'objet du test.

Au moment de décider si une évaluation donnée ou un certain élément de l'évaluation sont biaisés, il faut soigneusement établir de quelle façon la tâche en question peut être reliée de manière explicite au construit sous-jacent et quels sont les facteurs de biais possibles. Baser cette décision sur des considérations purement statistiques risque de faire tomber dans le piège de la confusion entre une

évaluation biaisée et une évaluation dont l'objet établit une discrimination entre les groupes. Goldstein (1996) et Humphreys (1986) ont suggéré qu'il était utile de faire une distinction entre la « différence », qui est un fait déterminé de manière objective, et le « biais », qui est un jugement porté sur la pertinence de la différence.

Les six possibilités avancées par Black (1999) se retrouvent parmi les cas les plus communs où une question peut être injuste à cause de ses effets sur les élèves.

1. Le contexte dans lequel la question est posée peut favoriser les élèves qui le connaissent bien. Par exemple, les références culturelles américaines favorisent les élèves vivant aux États-Unis et peuvent défavoriser les élèves vivant ailleurs.
2. Les sujets de dissertation portant sur les relations humaines peuvent favoriser les élèves évoluant dans des contextes culturels ou familiaux qui encouragent l'expression des sentiments.
3. Les questions à choix multiple peuvent avantager les garçons.
4. Les travaux réalisés dans le cadre des cours ou les projets peuvent avantager les filles.
5. Une question utilisant le langage ou les conventions associés à un milieu socio-économique en particulier peut désavantager les élèves venant d'un autre milieu.
6. Certaines questions ne sont compréhensibles que dans certains contextes culturels. C'est notamment le cas d'une question portant sur le fait que les personnes âgées vivent seules, car ce concept courant dans certains contextes culturels est complètement étranger à d'autres.

Par ailleurs, bien qu'il existe des preuves du biais fondé sur le genre, aux États-Unis notamment, il est impossible pour l'heure de déterminer avec certitude quels sont les aspects de cette question qui conduisent à ce bilan.

Au moment de concevoir les évaluations, la façon dont il faut tenir compte des différences entre les élèves passant une évaluation donnée n'apparaît pas toujours clairement. Les évaluations doivent être conçues de façon à ce que l'incidence globale du biais soit réduite. Pour ce faire, un éventail de tâches et de types de questions sera utilisé. Tout stéréotype lié à la culture ou au genre (qu'il soit explicite ou non) doit être évité. Le contenu de chaque question doit être minutieusement examiné afin d'éviter les catégories les plus évidentes de questions injustes. De plus, les tests préliminaires, qui consistent à donner ces questions aux échantillons d'élèves des différents groupes de la population scolaire, peuvent révéler des cas cachés. Toutefois, si l'on exclut tous les types de questions biaisées et tous les scénarios possibles, les personnes chargées de la conception des évaluations et de la rédaction des questions se retrouvent devant un choix limité. Les contraintes qui en résultent auront un impact négatif sur la validité de l'évaluation. À part éviter les pièges évidents et superflus, la solution la plus raisonnable semble être d'adopter une approche équilibrée lors de la conception de l'évaluation, en utilisant pour cela divers types de tâches d'évaluation et de présentations.

Une autre question se pose : celle du nombre de groupes différents d'élèves d'une même population scolaire qui doivent être pris en considération. Par exemple, faut-il tenir compte des élèves ayant des styles d'apprentissage différents ? Comme Hieronymus et Hoover (1986) l'ont indiqué, si les différences en matière d'intérêt et de motivation sont considérées comme des facteurs de biais, l'on peut alors considérer que toutes les tâches ou méthodes d'évaluation comportent un certain biais. Par exemple, certains extraits utilisés pour les examens de langue intéresseront sûrement des élèves en particulier plus que d'autres. L'équité dans le contexte de l'évaluation, qui comprend notamment le soin d'éviter le biais, est un problème majeur, surtout dans certains pays et régions (ou territoires) où tout biais qui peut être démontré dans un outil d'évaluation peut entraîner des poursuites judiciaires. Toutefois, la preuve du biais, par opposition à la différence d'accomplissement, est souvent une question d'interprétation, étroitement liée au contexte social dans lequel l'évaluation a lieu.

Dans la conclusion de *A fair test? Assessment, achievement and equity*, Gipps et Murphy (1994) déclarent que le test impartial n'existe pas et ne peut exister, car la situation est trop compliquée et la notion trop simpliste. Cela ne veut toutefois pas dire que les personnes chargées de la conception des évaluations et de la rédaction des questions ne doivent pas faire tout ce qui est en leur pouvoir pour réduire les effets du biais et de l'injustice. Gipps et Murphy soutiennent également que les concepteurs et conceptrices d'évaluations doivent avoir pour but l'égalité des chances et l'égalité d'accès à l'évaluation plutôt que l'égalité des

résultats, qui est fabriquée de toute pièce en manipulant chaque élément du test en fonction des statistiques obtenues pour sa réponse. Ces personnes se demandent, par exemple, dans quelle mesure l'introduction de questions à choix multiple dans les examens d'anglais afin d'améliorer les résultats des garçons serait justifiable, puisqu'elle fausserait la validité de l'évaluation selon les définitions communément acceptées de ce terme.

Il est largement reconnu que le manque d'impartialité dans le processus d'évaluation n'est qu'un des facteurs qui contribuent à l'iniquité dans le domaine de l'éducation, et peut-être même l'un des moins importants. Il existe bien d'autres sources d'iniquité dans le domaine de l'éducation qui ont une incidence considérable sur l'accomplissement des élèves, notamment :

- les différences de qualité de l'enseignement au sein d'un même établissement scolaire ;
- les différences de ressources dans les différents établissements scolaires et les différentes zones géographiques ;
- les différences de soutien social et familial apporté aux élèves individuellement.

N'importe laquelle de ces sources pourrait compromettre considérablement la réussite scolaire de chaque élève de telle façon que le processus d'évaluation, si impartial soit-il, ne pourrait rétablir l'équilibre. Par exemple, Smith et Tomlinson (1989) ont découvert que l'efficacité de l'enseignement dans un établissement scolaire était un facteur bien plus important pour expliquer les différences dans les résultats d'examen que l'appartenance ethnique des élèves. Cela montre que les tentatives qui sont faites pour adapter les outils d'évaluation afin d'empêcher que les résultats des divers groupes ne soient différents peuvent parfois être inappropriées.

Ce genre de considération a été utilisé pour justifier l'évaluation des aptitudes plutôt que celle de l'accomplissement, mais l'on a compris qu'il était impossible d'évaluer exclusivement l'aptitude, la capacité ou le potentiel, sans tenir compte du contexte social et du vécu scolaire. Il est également impossible d'examiner l'accomplissement scolaire indépendamment du contexte social et culturel. Le concept de réussite scolaire est défini et mesuré en fonction des normes d'une partie précise et délimitée de la société, quelle qu'elle soit.

### **La suppression du biais et des obstacles à l'évaluation**

Un autre aspect du biais auquel il faut remédier est la possibilité qu'une tâche d'évaluation établisse une discrimination injuste contre des élèves. Il est possible d'y parvenir en veillant à ce que les conditions de l'évaluation permettent la mise en œuvre des aménagements nécessaires pour supprimer les obstacles à l'évaluation. Les élèves dans leur totalité pourront ainsi montrer leur accomplissement scolaire sur un pied d'égalité avec leurs pairs.

Cette idée est abordée plus en détail dans la section intitulée « [Le caractère prioritaire de l'équité universelle](#) ».

### **La reconnaissance du biais**

Il ne faut pas perdre de vue que le biais peut être aussi bien positif que négatif. Si une tâche est très familière à un groupe d'élèves ou qu'il leur est plus facile qu'à d'autres de la réaliser, il s'agit bien d'un biais. Une évaluation équitable a pour objectif de donner la même chance à l'ensemble des élèves.

Cette publication explique de quelle façon un biais peut être introduit dans la conception des évaluations ou dans le processus de notation. Bien qu'il soit très important d'être proactif en abordant le biais dans ces phases du cycle d'évaluation, une simple réflexion sur le biais potentiel ne suffit pas. Il faut plutôt chercher des preuves du biais dans les résultats de l'élève, en analysant les résultats obtenus pour une question en particulier par rapport à ceux obtenus à l'examen dans son ensemble.

Il est également important de fonder les décisions relatives au biais sur des preuves, et non sur des stéréotypes.

## La comparabilité

La comparabilité constitue l'un des aspects les plus complexes de la validité. Les résultats aux évaluations sont régulièrement utilisés pour comparer les élèves à des fins de sélection. Lorsque deux élèves passent le même examen en même temps, il est possible d'affirmer avec certitude que l'élève ayant obtenu une note finale de 7 a mieux réussi ce jour-là que l'élève ayant obtenu une note finale de 4. Cependant, des comparaisons plus complexes sont souvent faites. Voici quelques exemples de comparaisons courantes :

- deux élèves ayant obtenu une note finale de 6 en histoire, mais ayant répondu à des questions différentes ou choisi des options différentes ;
- deux élèves, ayant obtenu une note finale de 5 en littérature espagnole en mai 2025 dans le premier cas et ayant obtenu une note finale de 5 dans cette matière en novembre 2023 dans le second cas ;
- deux élèves, ayant obtenu une note finale de 4 en physique dans le premier cas et ayant obtenu une note finale de 4 en chimie dans le second cas ;
- deux élèves, ayant obtenu une note finale de 4 en mathématiques dans le premier cas et ayant obtenu une note finale de 4 en géographie dans le second cas ;
- une élève vivant dans une certaine zone géographique ayant obtenu une note finale de 3 en informatique et une note finale de 6 en littérature chinoise, et un élève vivant dans une autre zone géographique ayant obtenu une note finale de 5 en littérature japonaise et une note finale de 4 en biologie ;
- deux élèves de 15 ans, ayant obtenu le certificat du PEI dans le premier cas et ayant passé les examens d'un autre organisme décernant des titres dans le second cas ;
- deux élèves ayant obtenu une note finale de 6 en indonésien B au niveau moyen (NM) dans le premier cas et ayant obtenu une note finale de 5 en indonésien B au niveau supérieur (NS) dans le second cas.

La comparabilité s'interroge sur la possibilité de mettre sur le même plan les résultats obtenus à deux évaluations. Cela semble particulièrement difficile entre les matières, car l'IB évalue des choses différentes et se demande ensuite si elles sont de valeur équivalente.

Le concept de la validité appliquée dans un but précis plutôt que comme la caractéristique d'une évaluation abstraite est particulièrement pertinent ici. Les résultats d'une élève en musique sont une meilleure indication de sa préparation à devenir musicienne professionnelle que ses notes en arts visuels, mais ces deux résultats peuvent avoir la même valeur pour déterminer sa préparation à étudier l'histoire.

Le problème de la comparabilité est rendu encore plus complexe par le fait que chaque individu a ses propres points forts et points faibles, ce qui explique que les élèves trouvent plus facile de réussir dans certaines matières que dans d'autres. L'IB a donc un grand défi à relever puisqu'il existe des disparités considérables au niveau des cohortes passant chaque examen, et ce, d'autant plus lorsqu'un choix de cours est proposé.

L'IB s'efforce de préserver les trois principes suivants concernant la comparabilité.

- La qualité du travail associée aux différentes notes finales dans une matière ou une discipline est comparable entre les différentes sessions d'une même année et d'une année à l'autre.
- Les notes finales ont la même signification dans toutes les matières ou disciplines, de sorte que les différents parcours possibles pour obtenir le diplôme correspondant au programme (par exemple, le certificat du PEI ou le diplôme de l'IB) soient comparables.
- Bien que l'IB maintienne un engagement en faveur de l'évaluation des compétences de pensée de haut niveau, il veille scrupuleusement à ce qu'une comparaison reste possible entre les notes finales obtenues dans son système et celles obtenues dans d'autres diplômes.

## La mesure de la comparabilité

Il existe de nombreuses façons de mesurer dans quelle proportion deux évaluations sont comparables, et ce thème a fait l'objet de nombreux articles de recherche. Dans leur revue de littérature sur la comparabilité entre les matières, Coe et coll. (2008) ont réparti les méthodes de comparaison des difficultés des examens

dans plusieurs matières en deux grandes catégories : les méthodes statistiques et les méthodes fondées sur une appréciation.

Les méthodes statistiques s’attachent à comparer les résultats obtenus par les élèves aux évaluations et à chercher des tendances. Elles reposent sur l’idée que si deux évaluations sont comparables alors, avec un échantillon aléatoire suffisamment grand d’élèves passant ces évaluations, les résultats devraient globalement être les mêmes aux deux évaluations.

À l’inverse, les méthodes fondées sur une appréciation font appel à des spécialistes des matières pour examiner un certain nombre d’évaluations et donner un avis mûrement réfléchi sur leurs difficultés respectives. Divers outils et techniques de recherche sont utilisés pour garantir que ces spécialistes comparent ce qui est comparable.

Ces deux approches présentent de graves défauts conceptuels. Coe et coll. (2008) ont ainsi formulé six critiques générales concernant ces techniques.

Tableau 4

**Critiques des méthodes statistiques et des méthodes fondées sur une appréciation adoptées pour comparer des évaluations**

| Critiques des méthodes statistiques  | Critiques des méthodes fondées sur une appréciation  |
|--|--|
| <ul style="list-style-type: none"> <li>• Les études sont susceptibles de mesurer des facteurs autres que la difficulté, comme l’enseignement ou la motivation.</li> <li>• Caractère multidimensionnel : il peut n’y avoir aucun point commun entre les matières.</li> <li>• Problème de représentativité : les statistiques reposent-elles sur un groupe d’élèves intrinsèquement biaisé ?</li> <li>• Différences entre les sous-groupes : si différents groupes d’élèves sont exposés à des niveaux de difficulté différents, cela ne remet-il pas en question les conclusions de la comparabilité relative ?</li> <li>• Manque de cohérence entre les méthodes : est-il possible de considérer qu’une méthode est « correcte » ?</li> <li>• Problèmes liés à l’obligation d’égalité : quels en seront les effets sur les élèves qui passent les évaluations ?</li> </ul> | <ul style="list-style-type: none"> <li>• Étendue des critères : pour pouvoir être appliqués à différentes matières, les critères doivent être très généraux et par conséquent, imprécis.</li> <li>• Notation des réponses en fonction du niveau de difficulté des questions : les examinateurs et examinatrices ont tendance à accorder davantage de points aux bonnes réponses apportées aux questions faciles qu’aux réponses plus faibles apportées aux questions plus difficiles.</li> <li>• Notation des différents types de travaux : les examinateurs et examinatrices ont beaucoup de mal à comparer des tâches différentes (par exemple, les réponses brèves et les dissertations).</li> <li>• Appui sur des comparaisons statistiques : ces méthodes sont influencées par les connaissances en matière de schéma d’évolution probable des résultats d’élèves types.</li> <li>• Interprétation et contexte : les évaluations fondées sur une appréciation de la capacité des élèves ou de l’équité doivent tenir compte des différences structurelles, comme celles qui existent entre des examens finaux uniques et une série d’évaluations modulaires.</li> <li>• Regroupement d’appréciations : la plupart des évaluations mesurent plusieurs critères qui doivent être équilibrés.</li> </ul> |

Les personnes qui défendent l’une des méthodes sont souvent très critiques à l’égard de l’autre méthode, et d’autres soutiennent que toutes les approches actuelles sont fondamentalement erronées. Pour étudier

ce sujet, Coe et coll. (2008) ont appliqué cinq méthodes de mesure de la comparabilité entre les matières aux examens britanniques (GCSE et A-level). Ils en ont conclu qu'il existait un niveau de cohérence relativement élevé entre les mesures statistiques et fondées sur une appréciation de la comparabilité, et que les différences qui les séparaient étaient bien moins importantes que les différences entre les matières. Ils ont également constaté que les difficultés dans chaque matière restaient stables d'une année à l'autre.

L'IB assure la comparabilité entre les années et les options grâce à une méthode passant par l'appréciation de l'examineur ou examinatrice, l'analyse statistique et le retour d'information de l'enseignant ou enseignante (voir la section « [L'attribution des notes finales et la totalisation](#) »). L'IB étudie également la comparabilité sur le plan des matières ou des disciplines à l'aide d'approches statistiques et d'avis de spécialistes.

## L'approche de la validité adoptée par l'IB

- L'IB accorde plus d'importance à l'authenticité et à la pertinence du construit de l'évaluation qu'à la recherche de la fiabilité maximale.
- L'IB croit en une éducation globale et équilibrée. Sa priorité est de construire une argumentation solide qui montre la validité de chaque programme. Cette validité prévaut sur celle des cours et options pris individuellement.

La validité consiste en une recherche complexe et multidimensionnelle d'un équilibre entre diverses exigences importantes et contradictoires. Il n'existe pas une seule bonne manière de trouver un équilibre entre les divers éléments de la validité : le choix du point d'équilibre résulte d'un jugement fondé sur les valeurs de l'organisation élaborant les évaluations.

Pour l'IB, il est fondamental d'évaluer **ce qui est important** et de le faire de manière authentique en proposant des évaluations qui **tiennent compte du monde réel**. La notion de « pertinence du construit » englobe le premier point. En d'autres termes, les évaluations de l'IB portent sur les éléments qui comptent vraiment dans chaque matière, et non sur ce qu'il est facile de noter. Le second point fait référence à la notion d'authenticité : les évaluations de l'IB comportent des tâches riches de sens qui présentent des situations que les élèves peuvent rencontrer dans le monde réel, et non des tâches artificielles et forcées.

Ces objectifs sont cependant atteints au détriment de certains autres aspects de la validité, en particulier la **fiabilité** de l'évaluation. La notation des tâches authentiques et riches de sens s'accompagne généralement d'un degré important de subjectivité, ce qui implique d'accepter de plus grands écarts entre les notes des examinateurs et examinatrices qu'avec, par exemple, des évaluations à choix multiple qui n'admettent qu'une bonne réponse et peuvent être notées en toute objectivité. De plus, ces tâches ont également une incidence sur la **facilité d'application** de l'évaluation. Ces types d'évaluations, qui ont la faveur de l'IB, sont en effet plus complexes et nécessitent plus de temps, que ce soit pour leur conception, leur utilisation ou leur correction. Ils demandent également un plus grand investissement de la part des élèves. Cela induit en effet une charge de travail plus grande, facile à observer en comparant, par exemple, le travail fourni lors de la réalisation de tâches nécessitant un travail de recherche sérieux au travail fourni pour se concentrer uniquement sur la réussite aux examens.

Bien que l'IB admette que d'autres organisations puissent choisir des priorités différentes pour trouver un équilibre dans la validité de l'évaluation, il est quant à lui convaincu que sa position est adaptée et justifiable pour son système de vérification externe des évaluations.

L'IB a pour ambition de proposer des programmes d'études qui vont plus loin que les autres, en développant chez les jeunes la curiosité intellectuelle, les connaissances, la sensibilité et la détermination nécessaires pour réussir. Comme le souligne sa déclaration de mission, l'IB espère que grâce à ses programmes éducatifs, les élèves seront capables de bâtir un monde meilleur, dans un esprit d'entente mutuelle et de respect interculturel. Par ailleurs, chacun des programmes de l'IB contribue au développement des aptitudes du profil de la communauté d'apprentissage de l'IB. Ce profil met en avant dix aptitudes auxquelles l'IB tient et qu'il espère voir se développer chez les élèves.

Du point de vue de l'évaluation, cela signifie que les buts de l'IB sont définis au niveau du programme, et non au niveau des différentes matières ou disciplines. De ce fait, la question de la validité de l'évaluation

doit être posée au niveau du programme et tenir compte des règles qui régissent l'attribution du certificat ou du diplôme correspondant au programme, et pas simplement de chaque note finale.

Cela ne signifie pas pour autant qu'il n'est pas important de tenir compte de chaque cours, ni de chaque tâche d'évaluation. Au contraire, certains aspects de la validité n'ont en réalité de sens qu'à ce niveau. Cependant, construire un argument de validité au niveau global implique que l'IB réfléchisse à l'ensemble du programme d'études suivi par l'élève.

La plupart des principes de l'évaluation exposés dans la présente publication s'appliquent à l'ensemble du système éducatif de l'IB, mais une partie précise (« [Section C : les processus spécifiques aux programmes de l'IB](#) ») se penche sur chaque programme tour à tour et examine les particularités de chacun.

## Les avantages de l'évaluation numérique pour la validité

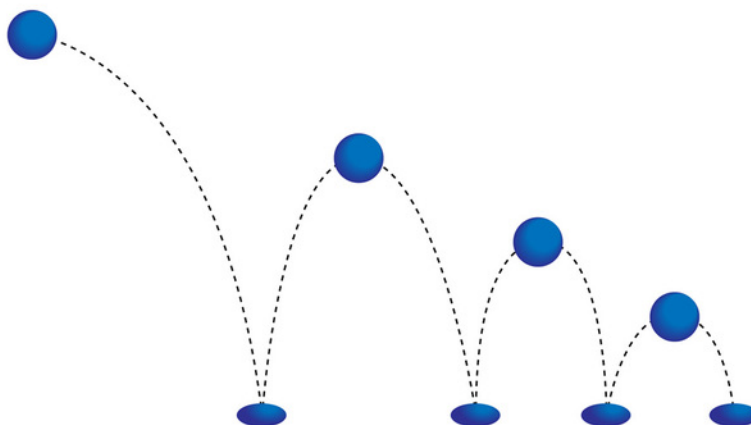
- Aujourd'hui, le monde n'utilise plus uniquement les supports papier : les appareils informatiques ont pris une grande place dans tous les aspects de la vie.
- Les examens sur papier ne permettent pas l'utilisation d'images en mouvement ni une interaction intéressante avec les élèves.
- Grâce à la polyvalence des appareils informatiques, les élèves peuvent apporter directement les modifications visuelles et auditives nécessaires pour mieux prendre connaissance des questions, et éviter ainsi d'avoir à demander des épreuves modifiées plusieurs mois avant les examens.
- Contrairement aux idées reçues, l'évaluation numérique offre davantage de protection contre la mauvaise conduite des élèves et la mauvaise administration des établissements.

Les courriels, les textos et les médias sociaux sont devenus des moyens de communication bien plus courants que les lettres, et dans bien des métiers, utiliser un appareil informatique est une pratique quotidienne. Si l'IB se donne pour objectif de proposer des évaluations authentiques, alors il paraît logique qu'il intègre l'utilisation de la technologie et des appareils dans les évaluations.

Depuis toujours, l'évaluation est très limitée quant au spectre qu'elle peut couvrir par l'intermédiaire des tests traditionnels (par exemple, réalisés avec du papier et un crayon). Avec les examens sur papier, les élèves doivent se contenter de réagir à un stimulus basique ou de répondre à une question simple, sans avoir aucune possibilité de manipulation ni d'interaction. À son niveau le plus élémentaire, l'évaluation numérique permet d'inclure des stimuli audio et vidéo, lesquels sont directement accessibles aux élèves qui n'ont plus à attendre que la lecture soit lancée pour l'ensemble de la classe. Dans le cadre du Programme d'éducation intermédiaire (PEI), l'évaluation numérique permet également de remplacer une image et une longue description du mouvement d'un objet par des schémas animés et des simulations.

Figure 13

*Gain de clarté apporté à une question d'évaluation électronique par l'utilisation d'une image en mouvement*



La technologie faisant évoluer les tâches d'évaluation, il devient possible d'évaluer la manière dont l'élève aborde un problème, réagit à de nouvelles informations ou interagit avec une simulation. Avec les systèmes de tutorat intelligents, l'évaluation numérique devrait permettre de réagir aux pistes lancées par les élèves. Cela recréera les possibilités offertes par les évaluations organisées par les enseignantes et enseignants, comme les épreuves orales, en évitant les problèmes engendrés par le fait que le personnel enseignant ne soit pas le même pour l'intégralité des élèves. Par conséquent, un atout majeur de l'évaluation numérique est qu'elle permet à l'IB de concevoir de nouvelles tâches pour évaluer ce qu'il souhaite véritablement évaluer, sans être limité par les capacités d'évaluation d'un examen sur papier.

Un autre avantage notable de la validité est qu'elle limite le biais, ce qui est particulièrement appréciable lorsque des aménagements à des fins d'accès et d'inclusion sont requis. L'IB traite en effet régulièrement des demandes d'impression des examens dans des polices de caractères différentes, sur du papier d'une autre couleur, ou encore en braille.

L'évaluation numérique ne permet pas totalement de supprimer ces obstacles, mais la technologie offre la possibilité aux élèves de choisir la taille de la police et la couleur qui sont les plus adaptées à leurs besoins ou qui ont leur préférence. Et les avantages en matière d'accessibilité ne s'arrêtent pas là : les élèves peuvent utiliser des logiciels de lecteurs d'écran, ce qui constitue un atout de taille. Si l'appareil informatique ne peut pas répondre à tous les besoins en matière d'accès, l'évaluation numérique est accessible à un éventail plus large d'élèves, et ce, sans recourir à une aide supplémentaire, même si celle-ci reste envisageable si nécessaire.

La sécurité des examens figure parmi les préoccupations les plus fréquentes au sujet de l'évaluation numérique, et elle présente également une grande menace pour la validité. Les personnes réticentes à l'évaluation numérique soutiennent souvent que ce système est moins sûr que les examens sur papier. En réalité, les évaluations numériques sont à l'abri de certains risques auxquels sont exposés les examens sur papier, mais elles en courent d'autres.

L'un des avantages les plus intéressants des examens numériques est qu'ils sont proposés en toute sécurité et qu'ils ne sont accessibles que pendant le laps de temps défini par l'IB, les utilisateurs et utilisatrices ne pouvant accéder au contenu des examens qu'après autorisation de leur établissement scolaire. Ce système est donc préférable à l'envoi des examens sur papier, qui peuvent être lus par n'importe qui et imposent donc aux établissements de les conserver en lieu sûr pendant plusieurs jours avant la tenue de l'examen.

L'IB admet que l'évaluation numérique pose des difficultés supplémentaires le jour de l'examen, étant donné que les élèves ont accès à des appareils connectés à Internet. Pour réduire les risques d'utilisation de l'appareil informatique pour consulter du matériel non autorisé, les examens numériques de l'IB se déroulent en mode verrouillage. Pendant toute la durée de l'examen, l'accès à tout contenu autre que l'examen lui-même est bloqué afin de renforcer la sécurité et de soutenir l'intégrité intellectuelle. De plus, il est possible d'enregistrer la manière dont les élèves répondent aux questions (au moyen d'un fichier journal par exemple), en plus de leurs réponses définitives, et donc de détecter tout comportement assimilable à une mauvaise conduite.

## Les normes

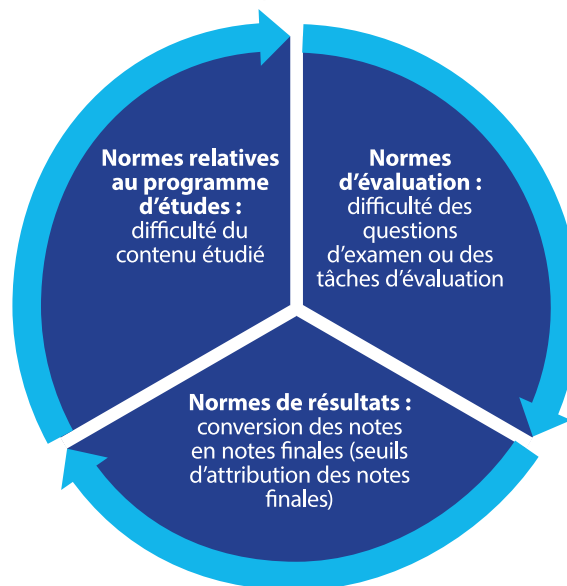
- Toute définition des normes comporte trois volets : le **programme d'études**, l'**évaluation** et les **résultats**.
- Dans le cadre de l'évaluation formative, l'accent est habituellement mis sur les normes relatives au programme d'études et à l'évaluation.
- L'IB suit une **approche comparée** (qui portait auparavant le nom d'« approche critériée limitée »), c'est-à-dire qu'il utilise de manière équilibrée des critères et une comparaison avec les résultats obtenus par les élèves les années précédentes pour établir la norme.
- Il est tout aussi important de maintenir les normes que de les établir.

## Les trois aspects des normes

Dans le domaine de l'évaluation, le terme « normes » fait généralement référence à la difficulté des tâches conçues pour les élèves. Les normes font partie intégrante de la comparabilité. Trois dimensions peuvent généralement être données aux normes.

Figure 14

Les trois aspects des normes



Si chaque aspect peut être modifié pour faire évoluer la difficulté générale de la matière, il ne faut pas perdre de vue que les modifications auront différents effets et seront assorties de différents délais.

L'IB modifie généralement un **programme d'études** lors du cycle de révision du programme, même si des modifications sont possibles à d'autres moments. À des fins d'équité, l'IB laisse au personnel enseignant le temps d'ajuster sa pédagogie pour qu'elle tienne compte des changements. Pour le Programme du diplôme et le Programme à orientation professionnelle (POP), cela se fera dans l'idéal sur deux ans (la durée du programme). Plus de temps pourra être nécessaire pour le Programme d'éducation intermédiaire (PEI), étant donné qu'il dure cinq ans, même si l'évaluation ne porte en principe que sur les deux dernières années.

Les normes d'**évaluation** dépendent du cycle d'élaboration des examens, qui s'étale généralement sur 12 à 18 mois. Contrairement aux normes relatives au programme d'études, les normes d'évaluation varieront légèrement d'une année à l'autre, car il est impossible de mettre au point deux séries de questions présentant le même niveau de difficulté. En effet, même si les questions sont reprises de l'examen précédent, le niveau de difficulté sera moindre, car ces questions utilisées pour la seconde fois seront, dans une certaine mesure, familières aux élèves. Il peut arriver qu'une décision soit prise de changer délibérément les normes d'évaluation d'un examen si les examens précédents n'ont pas produit les résultats escomptés. En fonction du degré de changement nécessaire, il est possible que la modification concerne un examen déjà en cours d'élaboration. Là encore, l'équité impose à l'IB d'avertir les enseignantes et enseignants de ce changement mais, selon son ampleur, il se peut que cet avertissement se fasse dans des délais bien plus courts que pour les changements relatifs au programme d'études.

Les normes de **résultats** (c'est-à-dire les seuils d'attribution des notes finales) sont ajustées chaque année pour suivre les modifications apportées aux normes d'évaluation. Par exemple, si l'examen est plus difficile et que les élèves obtiennent majoritairement des notes plus basses, alors les seuils d'attribution des notes finales doivent être baissés en conséquence, afin que les élèves présentant un travail de même qualité obtiennent une note finale identique à celle qui leur aurait été attribuée l'année précédente. À l'inverse des deux types de normes précédents, ce sont ici surtout les examinateurs et examinatrices, et non les enseignantes et enseignants, qui doivent comprendre les normes de résultats, afin de pouvoir les maintenir à chaque session. Une norme de résultats peut être modifiée lors de l'établissement des seuils d'attribution des notes finales. Si l'IB a l'intention de le faire, il avertit les établissements scolaires suffisamment à l'avance pour qu'ils puissent gérer et adapter en conséquence les attentes à l'égard des élèves. Dans un tel cas, il est primordial que les examinateurs et examinatrices comprennent cette nouvelle norme pour pouvoir l'appliquer lors des sessions ultérieures.

Au moment d'établir la norme globale pour une évaluation, il est important de trouver un équilibre entre ces trois définitions. Bien qu'il soit possible de poser des questions extrêmement difficiles sur un contenu simple ou de définir des normes de résultats extrêmement élevées pour une série de questions simples, il en résulte généralement une pertinence du construit très réduite. Par exemple, si le maximum de points est exigé à un test pour qu'une note finale de 7 puisse être attribuée, l'élève doit alors faire preuve de très hauts niveaux de précision. Il serait peu probable d'obtenir une note finale de 7 avec une excellente compréhension de la matière, mais peu de soins apportés au travail ou des faiblesses affichées au niveau des compétences d'expression écrite. S'agit-il là du but visé par l'évaluation ?

La définition de la norme s'applique à la fois à l'évaluation **formative** et à l'évaluation **sommative**. Dans le cas de l'évaluation formative cependant, l'établissement des normes de résultats est habituellement intégré dans le jugement de l'enseignant ou enseignante qui décide du retour d'information à faire. En règle générale, bien plus d'attention est accordée aux normes relatives au programme d'études et aux normes d'évaluation dans l'évaluation formative, afin de garantir qu'un test donné est adapté à l'élève et qu'il apportera des renseignements utiles pour orienter l'enseignement.

## L'approche normative et l'approche critériée

- L'« approche normative » implique de classer les résultats des élèves et de faire correspondre un pourcentage fixe d'élèves à chaque note finale.
- Avec l'« approche critériée », la norme de résultats est établie à partir d'une description de ce qui devrait apparaître dans le travail accompli par l'élève.
- L'IB suit une approche comparée (également connue sous le nom d'« approche critériée limitée »), c'est-à-dire qu'il utilise de manière équilibrée des critères et une comparaison avec les résultats obtenus par les élèves les années précédentes pour établir la norme.

Les approches normative et critériée désignent deux façons d'établir et de maintenir les normes de résultats des évaluations.

## L'approche normative

L'approche normative est généralement associée aux tests standardisés. Il s'agit d'essayer le test en le faisant passer à un échantillon type d'élèves, puis de se servir des résultats (qui, par définition, devraient représenter une courbe de distribution normale ou en forme de cloche) comme d'une échelle de référence pour produire une note pour toute future cohorte qui passera ce même test. Le processus qui consiste à produire une distribution standard des notes à partir de l'essai initial constitue l'« étalonnage ».

Cette description technique de l'approche normative n'implique pas nécessairement d'appliquer une distribution fixe à chaque ensemble de résultats au test ; la distribution fixe n'est utilisée que pour l'étalonnage initial. La distribution des notes qui seront obtenues par les futures cohortes peut être différente de cette distribution normale.

Dans la pratique, le terme « approche normative » désigne souvent le processus consistant à attribuer un rang aux élèves en fonction des résultats obtenus et à établir le pourcentage d'élèves recevant chaque note finale. Par exemple, la meilleure note finale sera attribuée aux 15 % d'élèves en tête.

## L'approche critériée

Glaser (1963) a été le premier à proposer la notion d'évaluation critériée. Elle représentait un changement important dans la manière d'établir les normes de résultats, en mettant l'accent sur l'évaluation des accomplissements de l'élève dans un domaine comportemental bien défini (Popham, 1978).

Dans l'approche critériée, les résultats de l'élève sont comparés à une description prédéfinie des attentes correspondant à chaque note finale. Des spécialistes de la matière ou de l'évaluation se chargent en principe de cette comparaison, en faisant appel à leur jugement professionnel.

Cette approche présente une limite, car il est très difficile de proposer des descriptions sans équivoque qui ont la même signification pour l'ensemble des spécialistes susceptibles d'émettre les jugements. En effet, il a été avancé qu'aucun critère, quel que soit le niveau de précision de sa formulation, ne permettait une interprétation sans équivoque (William, 1993). Le résultat d'un test critérié traditionnel est la démonstration ou non de la maîtrise du domaine pertinent.

Dans la pratique, chacune de ces deux approches comporte des inconvénients majeurs. Une approche normative stricte impose de prouver avec force que le test actuel présente le même niveau de difficulté que le test initial, alors que l'approche critériée est soumise à l'effet de Good et Cresswell (1988), selon lequel le jugement des spécialistes ne tient pas compte avec précision des exigences des questions.

## L'approche comparée

L'IB suit une « approche comparée » (Newton, 2011), portant précédemment le nom d'« approche critériée limitée » (Baird et coll., 2000), qui repose sur des critères et reconnaît aussi l'effet de Good et Cresswell (1988). D'un point de vue concret pour l'IB, cette approche demande aux examinateurs spécialisés et examinatrices spécialisées d'établir un intervalle réduit dans lequel situer les seuils d'attribution des notes finales en partant des critères (descripteurs des notes finales). Une comparaison est ensuite faite avec les seuils calculés pour correspondre aux résultats des années précédentes. Les notes finales sont fixées lorsque ces deux seuils coïncident, mais en cas de divergence, d'autres discussions sont nécessaires pour déterminer comment harmoniser les preuves contradictoires.

Enfin, il convient de retenir que les tests critériés et normatifs diffèrent plus au niveau de l'analyse et de l'interprétation des réponses des élèves qu'au niveau des questions elles-mêmes.

## Le maintien des normes

- Une fois les normes appropriées établies, l'IB doit s'assurer qu'elles sont appliquées chaque année.
- Les normes relatives au programme d'études sont revues lors de la révision du programme.
- Les normes d'évaluation et les normes de résultats sont maintenues en associant jugement professionnel et preuves statistiques.

La comparabilité est un aspect fondamental de la validité. Dans le cadre du maintien des normes, cela signifie que l'IB doit veiller à ce que les mêmes normes soient appliquées chaque année. Les normes de l'IB reposent sur le sens des notes finales, et non des notes. La différence entre ces deux types de notes est expliquée dans la section « [La distinction entre les notes et les notes finales](#) ».

Les **normes relatives au programme d'études** sont les normes les plus faciles à maintenir, étant donné qu'elles restent les mêmes d'une session à l'autre. Les évaluations doivent être un reflet fidèle de l'ensemble du programme d'études, ce qui est vérifié lors de la rédaction des examens. Des changements peuvent toutefois être apportés aux normes relatives au programme d'études en raison de facteurs extérieurs. Les compétences en informatique illustrent parfaitement cette idée. Les appareils évoluent et prennent une place grandissante dans la vie quotidienne. De ce fait, les connaissances et la compréhension qui étaient autrefois réservées aux spécialistes en raison de leur complexité, se sont banalisées. Des modifications ont donc dû être apportées à la norme relative au programme d'études portant sur ce thème, même si le contenu reste inchangé.

Pour résoudre les problèmes de ce type, mais aussi pour garder le contenu à jour, l'IB s'appuie sur un cycle de révision des programmes. Il est prévu que les normes relatives au programme d'études soient modifiées à l'issue de ces révisions, et l'IB doit alors trouver un équilibre au niveau des normes d'évaluation et de résultats afin de ne pas désavantager les élèves.

Concernant les examens de l'IB, les **normes d'évaluation** sont modifiées à chaque session, car il est impossible que deux examens présentent exactement les mêmes exigences. Dans certains systèmes éducatifs, toute une série de tests préliminaires sont organisés pour déterminer la difficulté de chaque question, et les examens sont soigneusement conçus afin de garantir un niveau de difficulté adapté. L'IB a fait le choix de ne pas procéder à ce type de tests préliminaires, en raison du risque de divulgation des questions avant que les élèves ne passent les examens. La conception d'examens cohérents et équilibrés repose donc sur le jugement professionnel et l'expérience des personnes chargées de la rédaction et de la vérification des examens.

L'IB ajuste également les **normes de résultats** à partir des réponses des élèves pour veiller au maintien de la norme globale. Les deux principales méthodes de maintien des normes de résultats sont examinées dans la section précédente, et le processus d'attribution des notes finales de l'IB a pour but de maintenir ces normes.

## La description de la réussite des élèves permise par l'évaluation sommative

- Même si les notes finales ne sont qu'un aperçu très simplifié des accomplissements des élèves, elles permettent aux parties prenantes (en particulier les universités, les structures d'emploi et les divers établissements scolaires) de disposer d'une base pour décider de leur sélection.
- Si l'IB ne fournissait que des informations plus complexes et globales, ses parties prenantes auraient à simplifier ces preuves pour prendre des décisions éclairées en vue d'une sélection. Elles le feraient certainement avec une attention moins aigüe que celle apportée par l'IB lors de l'établissement des seuils d'attribution des notes finales.
- Dans les évaluations de l'IB, le recours au jugement professionnel est important pour atteindre un résultat riche de sens pour les élèves. L'IB reconnaît qu'il doit étayer ce jugement par des preuves objectives et sans biais pour garantir l'équité.

### Les effets des notes finales

Réfléchissez aux connaissances, aux compétences et à l'expérience qui font l'excellence des chefs. Leurs connaissances portent sur les combinaisons d'ingrédients et de saveurs qui leur permettront de composer un plat parfaitement équilibré. Leurs compétences interviennent pour le choix et la préparation des ingrédients, la cuisson ainsi que le dressage et la présentation. Et leur expérience concerne l'utilisation des techniques pour atteindre la perfection, l'appréciation d'un plat merveilleusement cuisiné, et la présentation d'un ensemble d'ingrédients et de plats pour procurer une immense satisfaction à table.

Réduisez maintenant toute cette complexité en une note finale unique sur 7 pour déterminer qui mettra sur la première place du podium. Le résultat est presque vide de sens. Même augmenter l'échelle et choisir une note finale sur 100, voire sur 1 000, serait peu utile. Cela permettrait d'avoir un plus grand champ de différenciation, mais le problème fondamental, à savoir comparer des ensembles de compétences différents, ne serait pas résolu.

Ce problème est précisément celui que pose l'évaluation. Comment représenter au mieux dans un résultat unique toute la complexité des connaissances, de la compréhension et des capacités des élèves ? Même s'il était possible de faire un recueil et un résumé précis de toutes les informations concernant les élèves, cela ne signifierait pas pour autant que l'IB pourrait leur attribuer une note finale reflétant parfaitement leurs talents.

Une autre possibilité, souvent suggérée, serait de ne pas attribuer de note finale du tout. L'IB pourrait peut-être fournir pour chaque élève une description personnalisée de ce qui a été bien réussi et des aspects à renforcer. Une telle approche respecte les principes d'un apprentissage et d'un enseignement de qualité, mais elle présente l'inconvénient majeur de compliquer considérablement les comparaisons entre élèves.

L'évaluation de l'IB a pour but d'aider les élèves à indiquer leurs accomplissements afin de poursuivre leurs études ou d'intégrer le monde du travail. Cela signifie que les établissements d'accueil (généralement les universités pour le Programme du diplôme ou le POP) doivent pratiquer une forme de sélection pour choisir les élèves à admettre.

À l'évidence, si les autres critères sont moins fiables que les examens, compter davantage sur ces critères signifiera faire reposer la sélection sur des décisions moins fiables.

(Cresswell, 1986, p. 42, traduction libre)

Si une sélection doit avoir lieu, l'IB est tenu d'apporter son soutien aux élèves en rendant ces décisions en matière de sélection aussi justes et pertinentes que possible. Si l'IB ne remet aux établissements d'accueil que des bilans descriptifs pour chaque élève, ils auront à trouver une autre méthode de comparaison, qui

sera certainement moins fiable et uniforme que les résultats aux examens exprimés sous forme de notes finales.

L'intention n'est pas de laisser entendre que les évaluations sommatives sont un moyen idéal ou particulièrement bon pour procéder à une telle sélection, mais elles sont plus équitables que bon nombre d'autres possibilités. De plus, lorsqu'elles sont associées à d'autres indicateurs pertinents (comme les moyennes pondérées cumulatives, les activités extrascolaires, les dissertations personnelles et les lettres de recommandation), elles peuvent apporter des renseignements utiles à la prise de décision. Et plus important encore peut-être, l'IB s'efforce constamment de rendre ces évaluations les plus équitables, utiles et fiables possibles.

Figure 15

*Une distinction devrait-elle être faite entre les deux élèves de chaque groupe ?*



Dans l'exemple de la figure 15, la différence entre les deux élèves du premier groupe est celle qui est faite entre les notes finales de 5 et 6 dans les descripteurs généraux des notes finales du Programme du diplôme. Si un choix était à faire, il serait probablement juste de choisir entre ces deux élèves, qui obtiendraient sans doute le même résultat en repassant le test. Dans le second groupe cependant, la différence risque de ne pas être significative, et les deux élèves ont probablement la même aptitude.

Cet exemple prouve que toutes les différences ne sont pas significatives, et que l'utilisation des notes finales donne une indication des aspects servant de base à l'IB pour distinguer deux élèves. Si l'élève se situe à la limite entre deux notes finales (que ce soit juste au-dessus ou juste au-dessous du seuil d'attribution), il est certainement vrai que n'importe laquelle des deux notes finales est juste. Mais pour la plupart des élèves, une différence entre les notes finales représente un écart d'accomplissement global important.

Ce raisonnement amène à s'interroger sur le nombre de notes finales à employer. Avec seulement deux notes finales (réussite ou échec), la plupart des élèves auraient une note juste, mais les conséquences seraient sérieuses pour qui se situe à la limite. À l'inverse, avec 20 notes finales, bien plus d'élèves se situeraient à une limite ou près d'une limite, mais les conséquences d'un rattachement à la mauvaise catégorie seraient moins graves pour chaque élève. Ce concept a fait l'objet d'une étude plus poussée par Cresswell (1986).

En général, l'IB a retenu **sept notes finales** pour représenter le nombre de catégories importantes découlant de ses évaluations et pour trouver un juste équilibre entre le nombre d'élèves se situant à une limite et les conséquences de l'attribution d'une note finale inappropriée.

## L'importance du jugement professionnel

Les compétences de pensée complexes et de haut niveau sur lesquelles se concentrent les évaluations de l'IB ne se prêtent pas facilement à une notation simple et exempte de jugement. Les réponses des élèves seront probablement très variées, tout en étant également valables et correctes. La recherche semble indiquer que les connaissances et les compétences complexes ne devraient pas être décomposées en petits éléments distincts pour être enseignées, et ce principe vaut également pour la notation de ces réponses. Lorsque les barèmes de notation sont élaborés, les spécialistes de matière de l'IB doivent donc donner des orientations très précises quant à la manière de noter le travail de l'ensemble des élèves, et ce, de façon homogène.

L'IB doit ainsi accorder une grande importance au jugement professionnel des examinateurs et examinatrices, et surtout à l'expérience professionnelle des examinateurs superviseurs et examinatrices superviseuses qui établissent et expliquent les normes de notation. Tout cela constitue un grand défi pour la fiabilité du système d'évaluation, mais ce défi doit être relevé dans le contexte de la précision des résultats.

## La notation des évaluations de l'IB

- Le terme « notation » désigne le processus consistant à évaluer dans quelle mesure l'élève a bien réalisé la tâche confiée.
- La notation peut se concentrer sur chacun des détails d'une réponse ou au contraire porter un jugement plus global.
- Dans le cadre de l'évaluation formative, la notation peut ne pas produire de résultat numérique et être uniquement descriptive.

### Définition de la notation

Les termes « notation » et « méthodes de notation » utilisés dans cette section désignent les évaluations sommatives validées par l'IB pour le PEI, le Programme du diplôme et le POP. Si l'évaluation formative est utilisée dans tous les programmes, cela se fait généralement en classe et l'IB ne valide pas les évaluations de ce type au moyen d'une notation par des examinateurs ou examinatrices.

La notation ne décrit pas dans quelle mesure l'élève a réussi. Le niveau atteint dépend de plusieurs facteurs, dont le degré de difficulté des questions et la note de passage correspondant à chaque niveau. La notation consiste à comparer la réponse de l'élève à ce qui est attendu d'une réponse parfaite.

Il est possible de procéder à la notation de multiples façons, en fonction de la nature de la tâche et du type de notation approprié au vu de l'outil d'évaluation utilisé.

### Les méthodes de notation

L'IB s'appuie sur toute une gamme d'outils d'évaluation et les exigences associées à chaque outil aident à déterminer quel type de notation convient le mieux.

Il arrive que la notation soit très **objective** : la réponse de l'élève est soit correcte, soit incorrecte. C'est souvent le cas lorsque seuls quelques mots-clés sont à donner ou lorsque l'élève doit faire un choix parmi plusieurs réponses possibles. L'autonotation désigne le processus qui consiste à utiliser la technologie pour évaluer le travail des élèves à l'aide d'un barème de notation prédéfini qui considère objectivement une réponse comme correcte ou incorrecte. L'IB fait appel à l'autonotation pour les questions à choix multiple, pour lesquelles les réponses sont indubitablement correctes ou incorrectes, l'exactitude d'une réponse pouvant être assurée à 100 % sans avoir à attendre le jugement d'un examinateur ou d'une examinatrice. Lorsque la gamme de réponses attendues est étroite, mais qu'un jugement est nécessaire, l'IB ne remplace pas la notation humaine par une notation assistée par l'intelligence artificielle.

La notation peut aussi être bien plus **subjective**. L'intervention d'un examinateur ou d'une examinatrice peut s'avérer nécessaire lorsqu'il s'agit de juger si l'élève a produit une réponse acceptable ou de choisir, parmi plusieurs énoncés (appelés « bandes de notation »), celui qui décrit le mieux la mesure dans laquelle une réponse donnée correspond à la réponse « parfaite ». Le tableau 5 illustre le type de bandes de notation utilisées par l'IB.

Tableau 5

#### Exemple de bandes de notation de l'IB

| Nombre de points | Descripteurs de niveaux                                    |
|------------------|--|
| 0                | Le travail n'atteint aucun des niveaux décrits ci-dessous. |

| Nombre de points | Descripteurs de niveaux   |
|------------------|---|
| 1 – 3            | <ul style="list-style-type: none"> <li>• La réponse montre une compréhension limitée de la question. Le vocabulaire spécifique à la matière n'est pas utilisé ou il est systématiquement utilisé de façon inappropriée.</li> <li>• La réponse explique très peu la question. Les idées sont superficielles et manquent souvent de clarté.</li> <li>• La réponse est descriptive. Toute analyse présente est superficielle ou incohérente. Les différents points de vue sont peu voire pas du tout discutés. Lorsqu'elle est présente, la conclusion est très superficielle ou n'est pas cohérente avec le reste de la réponse.</li> </ul>       |
| 4 – 6            | <ul style="list-style-type: none"> <li>• La réponse montre une compréhension élémentaire de la question. Le vocabulaire spécifique à la matière est utilisé, mais souvent de façon inappropriée.</li> <li>• L'explication de la question est sommaire et insuffisamment développée. Les idées sont souvent peu précises ou vagues, et il est souvent difficile de comprendre ce que la réponse essaie d'exprimer.</li> <li>• L'analyse est limitée et la réponse est dans l'ensemble plus descriptive qu'analytique. Les différents points de vue sont discutés, mais de façon limitée. Une conclusion simpliste est fournie.</li> </ul>        |
| 7 – 9            | <ul style="list-style-type: none"> <li>• La réponse montre une certaine compréhension de la question. Le vocabulaire spécifique à la matière est utilisé, parfois de façon appropriée.</li> <li>• La réponse explique de façon satisfaisante la question, mais l'explication n'est parfois pas assez claire et développée. Des idées pertinentes sont présentées, mais elles ne sont pas assez précises et détaillées.</li> <li>• La réponse présente une analyse, mais cette dernière n'est pas assez développée. Des points de vue différents sont quelque peu discutés. Une conclusion est fournie.</li> </ul>                               |
| 10 – 12          | <ul style="list-style-type: none"> <li>• La réponse montre une bonne compréhension de la question. Le vocabulaire spécifique à la matière est utilisé, la plupart du temps de façon appropriée.</li> <li>• La réponse explique clairement la question, mais l'explication n'est pas assez développée. Les idées présentées sont pertinentes et correctes, mais elles ne sont pas assez détaillées.</li> <li>• La réponse présente une analyse critique, mais cette dernière n'est pas assez développée. Différents points de vue sont discutés. La réponse mène à une conclusion qui est en adéquation avec les arguments présentés.</li> </ul> |
| 13 – 15          | <ul style="list-style-type: none"> <li>• La réponse montre une très bonne compréhension de la question. Le vocabulaire spécifique à la matière est utilisé avec exactitude et précision.</li> <li>• L'explication de la question est claire et suffisamment développée. Les idées présentées sont pertinentes, correctes et détaillées.</li> <li>• La réponse présente une analyse critique bien développée. Différents points de vue sont discutés de façon critique. La réponse mène à une conclusion raisonnée et clairement indiquée, qui est en adéquation avec les arguments présentés.</li> </ul>  |

Une autre analyse concerne la façon dont la notation est réalisée, étant donné qu'elle peut différencier plusieurs aspects du travail, qui portent souvent le nom de « **critères** ». Par exemple, une dissertation pourrait être évaluée selon quatre critères distincts : la qualité de la grammaire, l'exactitude des faits essentiels, la structure de la dissertation et la qualité de la conclusion. Cette approche impose de veiller à ce que les critères soient indépendants les uns des autres, car il ne serait pas juste d'attribuer des points à plusieurs reprises pour le même aspect d'une réponse.

Le contraire de la **notation à l'aide de critères** est le **jugement reposant sur l'impression générale**. Dans ce cas, l'examineur ou l'examinatrice trouve un compromis entre les différents aspects d'une réponse idéale et émet un jugement final qui reflète le travail dans son ensemble.

La principale difficulté qui se pose à la fiabilité de la notation vient de la nature de la tâche et de l'étendue des réponses attendues des élèves. Il serait en effet plus compliqué de noter avec fiabilité une tâche de large portée pour laquelle une grande étendue de réponses serait attendue qu'une tâche restreinte qui ne recevrait qu'une étendue de réponses limitée.

Pour en savoir plus sur les différentes approches de la notation, veuillez vous référer à la section consacrée à la [notation](#).

## La notation de l'évaluation formative

Tous les programmes de l'IB s'appuient sur l'évaluation formative. Il s'agit toutefois du **seul** type d'évaluation utilisé dans le cadre du Programme primaire (PP). Avec l'évaluation formative, la notation n'a pas à être uniquement numérique. Il est possible de comparer un travail à la réponse « parfaite » et de fournir des commentaires descriptifs sur les similitudes et les différences entre les deux. Les critères de réussite peuvent être portés à la connaissance des élèves pour leur indiquer précisément à quoi correspond un « bon » travail. Malgré la difficulté posée par toute comparaison entre deux réponses afin de déterminer celle qui est la meilleure, l'évaluation formative s'attache à fournir un retour d'information pour soutenir l'apprentissage.

Le personnel enseignant peut recourir à l'évaluation formative en classe afin de déceler les lacunes dans les connaissances, la compréhension ou les compétences des élèves. Le cas échéant, la fiabilité est de moindre importance. Les notes et les notes finales sont superflues. Le résultat de l'évaluation peut au contraire consister en une liste de compétences ou de thèmes sur lesquels l'élève doit travailler. Dans une telle situation, la pertinence du construit est primordiale, car l'élève doit recevoir un retour d'information sur les éléments qui présentent une importance pour le programme ou qui comptent pour lui permettre de montrer ses connaissances.

## Le renforcement d'un état d'esprit éthique

### Définition d'un état d'esprit éthique

Adopter un état d'esprit éthique signifie s'engager à agir de façon **responsable** et **authentique**, en suivant des principes d'**honnêteté**, de **respect** et d'**intégrité**. Dans le contexte scolaire, l'intégrité intellectuelle s'en trouve favorisée, car l'importance de produire un travail de façon légitime et honnête est constamment mise en avant. Cet état d'esprit va au-delà des règles et des politiques en ancrant la prise de décision et le comportement éthiques dans la culture de la communauté d'apprentissage. Il invite les élèves à accorder de la valeur à leur processus d'apprentissage, à faire des choix éclairés et à contribuer à rendre l'environnement scolaire équitable.

### Pourquoi créer un état d'esprit éthique ?

Il est essentiel de développer un état d'esprit éthique pour garantir la crédibilité et la validité des accomplissements scolaires. Il est le gage de l'intégrité des évaluations en permettant aux accomplissements des élèves d'être le reflet fidèle de leurs véritables capacités. Encourager un comportement éthique signifie aussi préparer les élèves à la vie en dehors de la salle de classe, en inculquant des valeurs fondamentales pour rencontrer la réussite dans l'enseignement supérieur et dans le milieu professionnel. Lorsque les principes éthiques sont ignorés, la confiance s'érode au sein de la communauté scolaire et la crédibilité de l'établissement s'en trouve entachée. Les établissements scolaires qui cherchent à renforcer un état d'esprit éthique chez leurs élèves les amènent naturellement à devenir des citoyennes et citoyens faisant preuve d'intégrité et agissant avec équité et respect dans tous les aspects de la vie.

### Comment un état d'esprit éthique est-il créé ?

Pour développer un état d'esprit éthique, il faut adopter une approche globale à l'échelle de la communauté. Les établissements scolaires doivent ancrer l'intégrité intellectuelle dans leur culture, en commençant par formuler des attentes claires, adaptées au stade de développement des élèves. Le personnel enseignant joue un rôle déterminant en adoptant un comportement éthique, en élaborant des stratégies préventives et en inscrivant des discussions autour de l'intégrité dans les programmes d'études. Les coordonnateurs, coordonnatrices et membres de la direction doivent veiller à ce que les politiques soient bien définies, constamment appliquées et promues par la formation de toutes les parties prenantes. Le dialogue avec les parents et les personnes exerçant une tutelle doit rester ouvert afin de faciliter l'uniformité des attentes et de corroborer l'importance d'adopter des pratiques éthiques à la maison.

Pour donner encore plus de force à ces principes, il ne faut pas hésiter à organiser des ateliers régulièrement, à donner des exemples parlants de défis éthiques et à favoriser le dialogue autour d'applications concrètes de l'intégrité. C'est en mettant en lumière les avantages de l'intégrité, et non en se concentrant exclusivement sur les sanctions, que les établissements scolaires créeront un environnement favorable au développement des compétences et de la compréhension nécessaires pour prendre des décisions éthiques.

Développer un état d'esprit éthique n'est pas un simple but mais une pratique continue, qui passe par le partage des responsabilités, la cohérence des informations transmises et l'investissement indéfectible de l'ensemble de la communauté scolaire. Pour de plus amples informations, veuillez consulter la ressource imprimable intitulée « [Le développement d'un état d'esprit éthique](#) ».

## L'utilisation éthique de l'intelligence artificielle dans l'évaluation

Pour savoir comment l'intelligence artificielle (IA) peut être combinée à l'évaluation, deux grands points sont à prendre en compte.

- Quelle est l'incidence de l'IA sur l'objet évalué ?
- Quel est le rôle de l'IA dans le processus suivi pour porter des jugements sur les travaux des élèves ?

L'IA étant aujourd'hui devenue une réalité concrète, l'attitude la plus responsable consiste à l'adopter et à aider les élèves à se servir de cette technologie de façon éclairée et éthique.

### « Qu'entend-on vraiment lorsque l'on indique qu'un travail est le nôtre ? »

Cette question se trouve au cœur même de la première difficulté liée à l'IA, à savoir déterminer son influence sur l'objet évalué. Lorsqu'il a été question d'introduire l'IA dans l'éducation, il a été craint au départ que les élèves s'en servent pour « tricher », en adoptant ce système pour faire réaliser le travail à leur place. Il a même été prétendu à plusieurs reprises que l'IA pouvait produire de meilleurs résultats que les élèves à certaines évaluations. Ces préoccupations ramènent aux questions les plus fondamentales au sujet de l'évaluation : quel est le but visé par la question considérée ? Quelles preuves sont recherchées dans le travail de l'élève ? Quelques secondes suffisent pour trouver sur Internet des faits tels que des dates historiques, mais ce moyen n'est pas toujours le plus approprié pour trouver une réponse lorsque le facteur temps est décisif, comme par exemple l'atterrissage d'un avion.

Ce constat vaut également pour les autres usages de l'IA. Parfois le contenu compte plus que la présentation. Le cas échéant, noter le style de communication présenterait peu d'intérêt, et l'IA pourrait constituer un soutien pour que l'élève présente ses idées sous une forme plus lisible. Lors de la rédaction d'un poème au contraire, le choix de chaque mot est primordial. Aucune partie de la version définitive d'un poème ne peut donc être composée à l'aide de l'IA, mais cet usage peut être autorisé pour constituer un thésaurus par exemple.

La question qu'il faut avant tout se poser pour intégrer l'IA dans le système éducatif est la suivante : qu'est-ce que les élèves ont besoin d'apprendre pour pouvoir faire un usage responsable et éthique de l'IA au quotidien ? Nul doute ne fait que cette question évoluera au fil du temps. Pour l'heure, il est possible d'y répondre en apprenant aux élèves à :

- porter un regard critique sur les informations reçues et avoir conscience des biais potentiels ou des conclusions trompeuses ;
- vérifier les sources et les références pour s'assurer de leur bonne utilisation ;
- comprendre ce qui relève de leur travail (ou pas) et en faire une présentation honnête.

### « Bonjour, je serai votre outil d'assistance virtuelle pour l'évaluation de cet examen. »

L'IA va progressivement changer la façon d'évaluer les élèves. Grâce aux progrès réalisés en matière d'évaluation numérique, les tâches deviennent bien plus interactives et l'introduction de l'IA permet d'aller encore plus loin : on peut ainsi envisager que les élèves expliquent leurs idées à des avatars numériques et interagissent avec eux. Cette technologie peut élargir le champ des approches de l'évaluation, en permettant par exemple de recourir plus souvent aux entretiens, une pratique ingérable actuellement en raison du nombre d'examineurs et examinatrices que cela mobiliserait.

Cela fait longtemps que l'évaluation s'appuie sur la notation automatique des réponses objectives (pour les questions à choix multiple ou de type « glisser-déposer » notamment), mais les progrès en matière d'apprentissage automatique et d'IA ont également accru le potentiel de cette notation automatique qui peut désormais s'appliquer à des questions plus subjectives et à des réponses plus complexes. Dans un avenir proche, toute note à enjeux finaux élevés devra être fixée par une personne humaine, mais l'IA se

montre très prometteuse pour améliorer la qualité de la notation et repérer tout travail devant être passé en revue par une seconde personne.

La notation par l'IA pourrait également soutenir l'évaluation formative utilisée en classe et offrir des possibilités d'évaluation adaptative, en se servant des réponses précédentes des élèves pour adapter les questions à leurs capacités. Si cette approche était réservée aux questions à choix multiple jusqu'à présent, les progrès réalisés en matière d'IA ont élargi l'éventail des types de questions pris en charge, ce qui permet de personnaliser davantage les examens pour les élèves.

Lorsque l'on réfléchit aux effets d'une nouvelle technologie, quelle qu'elle soit, il faut commencer par s'assurer d'une part, que le but de l'évaluation reste pertinent par rapport aux objectifs d'apprentissage visés et d'autre part, que les évaluations elles-mêmes conservent leur validité par rapport à ce but.

Souvenez-vous : il peut être utile que les élèves utilisent l'IA pour **soutenir leur apprentissage**, mais il n'est pas acceptable que les élèves l'utilisent **au lieu d'apprendre** ou pour remettre un travail qui n'est pas le leur.

## Lectures complémentaires

- *Points clés pour expliquer l'utilisation des outils d'intelligence artificielle (IA) (fiche d'informations)*
- *13 scénarios d'utilisation de l'intelligence artificielle par les élèves (version longue)*

## Définition des pratiques

- Les pratiques décrites dans cette section s'appliquent aux évaluations sommatives dans le cadre du Programme d'éducation intermédiaire (PEI), du Programme du diplôme et du Programme à orientation professionnelle (POP).
- Elles ne s'appliquent pas au Programme primaire (PP), dans lequel le Baccalauréat International (IB) n'organise aucune évaluation sommative, ni à l'évaluation formative organisée par l'établissement scolaire.

Cette section présente les pratiques employées par l'IB pour produire les résultats des élèves passant les évaluations qui sont notées ou révisées en externe. Elle n'aborde pas les évaluations qui ne sont pas de la responsabilité de l'IB, telles que les évaluations conçues par le personnel enseignant pour être utilisées dans le PP.

Un **principe** indique **pourquoi** quelque chose est fait, et une **pratique** décrit **comment** cela est mené à bien. Cette section explique donc les pratiques de haut niveau auxquelles l'IB fait appel pour garantir la validité des résultats de son évaluation.

Elle traite ensuite amplement les **procédures** de l'IB, qui décrivent les différentes étapes à suivre pour appliquer chaque pratique. Lorsque ces procédures concernent les établissements scolaires, elles sont expliquées en détail dans les *Procédures d'évaluation* spécifiques à chaque programme.

## Ce que les évaluations de l'IB mesurent et le rôle des acquis antérieurs

Les évaluations sommatives visent à mesurer les résultats des élèves en s'intéressant à leurs **connaissances, compréhension et compétences** à la fin de leurs études pour le PEI, le Programme du diplôme et le POP. De l'explication ci-dessus découlent notamment les points suivants.

- Les résultats de l'évaluation sommative reflètent la performance de l'élève à un moment donné. L'évaluation sommative ne mesure pas le potentiel de l'élève ni ses progrès en matière d'apprentissage.
- Les évaluations de l'IB sont conçues de façon à minimiser les effets d'un résultat d'une évaluation imputables à la contre-performance de l'élève sur une question particulière ou un jour particulier. Il est généralement possible d'y parvenir en faisant passer à l'élève différents examens, ce qui lui donne ainsi plusieurs occasions de montrer ses capacités. Toutefois, le nombre d'examens doit rester gérable et ne pas constituer une charge excessive pour l'élève.
- L'évaluation sommative reflète les connaissances, la compréhension et les compétences d'un ou une élève à titre individuel et non d'un groupe d'élèves.
- L'IB ne prend pas en compte les acquis antérieurs lors de l'attribution des notes finales des évaluations. Cela signifie que l'IB ne tient pas compte des diplômes, des notes finales, ni des accomplissements de l'élève avant de commencer le programme de l'IB concerné. Les guides pédagogiques de l'IB précisent si les acquis antérieurs sont pertinents pour chaque matière ou domaine d'apprentissage.

## La transmission des résultats des élèves

- L'objectif affiché par la déclaration de mission de l'IB est d'aider à former des jeunes capables de bâtir un monde meilleur. Une évaluation réussie devra donc soutenir cet objectif.
- Les notes finales de l'IB ont une signification, et les seuils d'attribution des notes finales sont établis de façon à en tenir compte.
- Même si les notes finales ne sont qu'un aperçu très simplifié des accomplissements des élèves, elles permettent aussi aux parties prenantes (en particulier les universités, les structures d'emploi et les divers établissements scolaires) de disposer d'une base pour décider de leur sélection.
- Si l'IB ne fournissait que des informations plus complexes et globales, d'autres auraient à simplifier ces preuves pour prendre des décisions éclairées en vue d'une sélection, ce qui serait certainement fait avec une attention moins aigüe que celle apportée par l'IB lors de l'établissement des seuils d'attribution des notes finales.
- Les accomplissements des élèves ne se réduisent pas aux résultats obtenus aux examens, et même en fonctionnant avec des notes finales, un résultat décevant pour une personne pourra être un grand accomplissement pour une autre.

Le Baccalauréat International (IB) a pour but de développer chez les jeunes la curiosité intellectuelle, les connaissances et la sensibilité nécessaires pour contribuer à bâtir un monde meilleur et plus paisible, dans un esprit d'entente mutuelle et de respect interculturel.

(Déclaration de mission de l'IB, 2024)

## La signification des notes finales de l'IB

Les résultats obtenus à l'issue des évaluations de l'IB portent le nom de « notes finales ». Ces notes finales illustrent la qualité du travail démontré par l'élève dans ses réponses.

Chaque note finale est associée à un descripteur, qui est publié par l'IB. Étant donné que ces descripteurs reflètent la qualité du travail que l'IB attend des élèves en fonction de leur âge, ils sont différents pour le PEI, le Programme du diplôme et le POP.

Figure 16

Exemples de descripteurs des notes finales pour le Programme du diplôme

## Descripteurs des notes finales du groupe 3 (Individus et sociétés)

### Individus et sociétés

#### Note finale 7

L'élève met en œuvre des compétences de pensée critique qui témoignent de sa sensibilité par rapport aux concepts, de sa perspicacité, de ses connaissances et de sa compréhension. Il s'avère extrêmement apte à apporter des réponses pleinement développées, structurées de manière logique et cohérente, et illustrées à l'aide d'exemples appropriés. Il emploie avec précision la terminologie spécifique à la matière et montre une bonne connaissance de la documentation relative à la matière. Il est capable d'analyser et d'évaluer des preuves, ainsi que de synthétiser des connaissances et des concepts. Il a conscience des différents points de vue ainsi que des préjugés subjectifs et idéologiques. Il est capable de tirer des conclusions raisonnables, même si celles-ci sont parfois approximatives. Son travail fait systématiquement état d'une réflexion critique et approfondie. Il est capable d'analyser et d'évaluer des données ou de résoudre des problèmes de manière extrêmement compétente.

#### Note finale 6

L'élève fait preuve d'une connaissance et d'une compréhension précises. Ses réponses sont cohérentes, structurées de manière logique et bien développées. Il emploie systématiquement la terminologie appropriée. Il est capable d'analyser, d'évaluer et de synthétiser des connaissances et des concepts. Il connaît les recherches, les théories et les questions pertinentes et a conscience des différents points de vue et contextes dans lesquels elles s'inscrivent. Son travail fait systématiquement état d'une réflexion critique. Il est capable d'analyser et d'évaluer des données ou de résoudre des problèmes de manière compétente.

#### Note finale 5

L'élève fait preuve d'une connaissance et d'une compréhension solides de la matière et a recours à la terminologie spécifique à la matière. Ses réponses sont cohérentes et structurées de manière logique, mais ne sont pas pleinement développées. Il est capable d'apporter des réponses pertinentes en essayant d'intégrer des connaissances et des concepts. Il a tendance à produire un travail plus descriptif qu'évaluatif même s'il fait preuve d'une certaine capacité à présenter et à développer des points de vue opposés. Son travail fait parfois état d'une réflexion critique. Il est capable d'analyser et d'évaluer des données ou de résoudre des problèmes.

Les descripteurs généraux des notes finales doivent être identiques pour toutes les matières d'un même groupe de matières, mais l'IB les place souvent dans des contextes spécifiques aux matières pour les rendre plus compréhensibles dans chacune. Il convient de comprendre que la qualité du travail ne change pas en fonction du contexte de la matière. Ainsi, une note finale de 4 en langue doit avoir la même signification qu'une note finale de 4 dans une matière scientifique ou artistique. Cette idée se retrouve dans l'approche adoptée par l'IB dans ses programmes, où toutes les notes finales sont d'importance équivalente. Elle fait cependant l'objet de nombreuses discussions parmi les spécialistes de l'éducation qui se posent les questions suivantes : comment peut-on comparer un accomplissement réalisé dans deux matières différentes ? Est-il même utile de faire une telle comparaison ? Ce concept est examiné de manière plus approfondie dans la section « [La comparabilité](#) ».

Figure 17

Comment ces deux travaux peuvent-ils être comparés ?



9. (a)  $x = e^{3y+1}$

en prenant le logarithme népérien de chaque membre de l'équation

$$(f^{-1}(x)) = \frac{1}{3} (\ln x - 1)$$

(b) les coordonnées de Q sont (1,0)

$$\frac{dy}{dx} = \frac{1}{x}$$

at Q,  $\frac{dy}{dx}$

$$y=x-1$$

(c) soit A l'aire demandée

$$A = \int_1^e 1 dx - \int_1^e \ln x dx$$

utilisez l'intégration par parties pour trouver  $\int \ln x dx$

$$= \left[ \frac{x^2}{2} - x \right]_1^e - [x \ln x - x]_1^e$$

$$= \frac{e^2}{2} - e - \frac{1}{2} \left( \frac{e^2 - 2e - 1}{2} \right)$$

Lorsque l'on s'interroge sur la validité, l'on peut défendre la comparabilité entre les notes finales de l'IB dans les différentes matières en s'appuyant sur l'objectif qui leur a été donné. En effet, les notes finales de l'IB ont pour but de permettre aux parties prenantes de comparer les niveaux atteints par les élèves. Par conséquent, il est judicieux de suivre des méthodes statistiques et qualitatives pour parvenir à la parité des significations données aux notes finales.

## La distinction entre les notes et les notes finales

Notes et notes finales sont deux choses différentes.

Figure 18

Comparaison imagée entre les notes et les notes finales



On s'attendrait à ce que les élèves de bon niveau aient terminé la plus grande partie de la tâche.

On s'attendrait à ce que les élèves de bon niveau aient terminé une partie de la tâche seulement.



La différence entre les notes et les notes finales peut être expliquée à l'aide de diverses images. Dans la figure 18 par exemple, la distance parcourue pourrait représenter les notes, la distance servant à mesurer jusqu'où la personne est allée. Mais pour comprendre à quel accomplissement correspond la distance parcourue, il faut tenir compte de la nature du terrain emprunté, une donnée qui est prise en compte dans l'établissement des notes finales.

- Une **note** est le reflet de la **proportion** de la tâche réalisée par l'élève.
- Une **note finale** tient compte de la **difficulté** de la tâche, afin de quantifier l'accomplissement traduit par la note de l'élève.

Analysez les deux exemples fournis dans la figure 19. Dans le premier cas, on s'attendrait à ce que presque toute la tâche soit réussie à 16 ans pour avoir une « bonne » note finale. Dans le second cas, les attentes seraient bien moins élevées pour attribuer une « bonne » note finale.

Figure 19

Tâches d'évaluation présentant des niveaux de difficulté variables



On s'attendrait à ce que les élèves de bon niveau aient terminé la plus grande partie de la tâche.

On s'attendrait à ce que les élèves de bon niveau aient terminé une partie de la tâche seulement.

Ce raisonnement amène à l'un des défis posés par l'organisation d'évaluations de qualité. Les élèves doivent avoir la possibilité de montrer leur plein potentiel, ce qui est souvent incompatible avec des tâches trop simples. À l'inverse, si les tâches sont trop difficiles, il est probable qu'un grand nombre d'élèves parviennent à peine à les commencer, d'où une utilité très limitée des notes finales pour faire une distinction entre les individus.

L'utilisation de tâches trop simples peut aussi conduire à mesurer l'exactitude plutôt que la compréhension. De ce fait, les élèves ayant commis des erreurs mineures ne pourraient pas obtenir les meilleures notes finales, malgré une bonne compréhension des thèmes.

Pour mieux comprendre la différence entre les notes et les notes finales, veuillez vous référer à la vidéo intitulée « Distinction entre notes et notes finales : comment prendre en compte la difficulté de l'évaluation et pas seulement ce que l'élève a réussi ? ».

## Une session d'examens réussie

Figure 20

*Différence de perspective sur une session d'examens réussie*



Je connaissais les réponses à toutes les questions.

Nous avons pu sélectionner les élèves de bon niveau pour nos cours.



Nos élèves ont mieux réussi que les élèves des établissements voisins.

Aucune erreur n'a été commise dans les questions et dans la notation.

La forme donnée à la réussite dépend essentiellement du point de vue adopté. L'IB met tout autant l'accent sur le haut niveau de pertinence du construit que sur la facilité d'utilisation. Pour être considérée comme réussie, une session d'examens doit réunir les conditions suivantes.

- L'évaluation doit donner à chaque élève des chances égales de montrer ses capacités.
- L'expérience doit être la plus simple possible pour les établissements scolaires comme pour les élèves.
- Toutes les parties prenantes doivent garder confiance dans les résultats (notes finales) publiés par l'IB.

## Les notes finales et les accomplissements

Comme l'indique la déclaration de mission de l'IB, le système éducatif de l'IB vise bien plus que la simple attribution d'une série de notes finales. Cela transparaît dans le profil de la communauté d'apprentissage de l'IB et le but est expliqué dans la publication *En quoi consiste le système éducatif de l'IB ?* (2019).

Les résultats de l'évaluation ne peuvent porter que sur un nombre restreint d'aspects de la déclaration de mission. Compte tenu du système de compensation utilisé dans les examens de l'IB, quel pourcentage des points attribués dans l'évaluation de mathématiques devrait être associé à l'altruisme ou à l'audace, en supposant qu'il soit raisonnable d'attribuer une valeur numérique à ces attitudes essentielles ? Les évaluations de l'IB reposent sur le principe qu'elles doivent avoir un effet de remous positif qui facilitera l'atteinte des objectifs de l'IB.

Figure 21

Une image globale de l'accomplissement de l'élève



Lorsque l'on rend compte des résultats des évaluations d'un programme de l'IB, il est important de ne pas s'arrêter aux notes finales pour illustrer l'ampleur de l'accomplissement de l'élève.

L'IB ne consigne que le résultat final, sans donner aucune indication sur les difficultés rencontrées par l'élève pour parvenir à ce résultat ou pour atteindre son plein potentiel. L'IB prend toute la mesure de ces deux éléments, mais il estime qu'il ne peut les mesurer efficacement avec l'évaluation sommative. Une telle évaluation relève donc de la responsabilité de l'établissement scolaire, car il a une vision globale du travail de l'élève.

Même si certaines méthodes permettent de calculer la « valeur ajoutée » de l'accomplissement de l'élève ou les « notes finales prévues en se basant sur l'accomplissement antérieur », l'IB recommande de faire preuve de prudence si elles sont utilisées pour avoir une idée de la réussite de l'élève. Ces mesures reposent en effet sur un profil moyen, alors que chaque élève présente un mélange unique de caractéristiques et de qualités personnelles, qui devraient être convenablement mises en valeur par les personnes ayant eu l'occasion d'apprendre à ses côtés pendant toute la durée d'un programme.

## Les notes finales prévues

Comme indiqué dans les *Normes de mise en œuvre des programmes et applications concrètes* (2020) : « L'établissement scolaire communique [...] les notes finales de l'IB prévues à l'IB dans le cadre de l'évaluation validée par l'IB, de manière à valoriser l'intégrité intellectuelle et à refléter le plus fidèlement possible le niveau des élèves. »

Étant donné que les résultats des élèves peuvent varier tout au long du programme d'études, il peut être difficile de prévoir les notes finales. Les notes finales prévues remplissent plusieurs fonctions capitales. Elles constituent des preuves qui sont examinées lors des réunions de délibérations, pour aider l'IB à mieux comprendre l'ensemble des points forts de la cohorte et la distribution des notes finales pour chaque matière. Elles aident aussi à repérer les écarts entre les notes finales prévues et les notes réelles, plusieurs contrôles pouvant être effectués au besoin. Du point de vue des établissements scolaires, les notes finales prévues aident à établir et à gérer les attentes pour les élèves, les parents et les personnes exerçant une tutelle.

Même si les membres du personnel enseignant ne seront vraisemblablement pas en mesure d'attribuer avec une exactitude absolue les notes finales prévues de l'ensemble de leurs élèves, l'objectif est de parvenir à un résultat aussi précis que possible. En effet, plus les notes finales prévues seront exactes, plus elles deviendront utiles et constructives dans le processus.

Une note finale prévue représente aussi une donnée importante pour l'évaluation : elle est prise en compte au niveau de la matière, pour établir les seuils d'attribution des notes finales, et au niveau de l'établissement scolaire, à condition que son exactitude soit avérée. L'exactitude des notes finales prévues est donc fondamentale pour pouvoir les considérer comme des données fiables.

Les notes finales prévues sont aussi :

- des données mises à la disposition des établissements, qui sont libres de les communiquer aux élèves ;
- une indication de la capacité à suivre des options après le programme, ce qui prépare les élèves à la réussite ;
- des notes données sous forme de notes finales (allant de 1 à 7 ou de A à E) et non de notes intermédiaires.

Bien prévoir les notes finales impose de comprendre ce qu'elles ne sont pas. Les notes finales prévues ne sont pas :

- un mécanisme permettant de compenser les circonstances défavorables ;
- un outil disciplinaire ;
- le résultat d'une équivalence reposant sur les seuils d'attribution des notes finales d'une session antérieure ;
- la même chose que les notes finales prévues à d'autres fins, comme pour intégrer un établissement d'enseignement supérieur ;
- un mécanisme visant à rassurer les élèves, leurs parents ou les personnes exerçant une tutelle ;
- une amplification ni une minoration des capacités.

Pour obtenir de plus amples informations sur les notes finales prévues et découvrir des stratégies permettant de gagner en exactitude, veuillez vous référer à la ressource imprimable intitulée « [Notes finales prévues : guide à l'intention du personnel enseignant](#) » (PDF) ainsi qu'aux conseils disponibles sur le Centre de ressources pédagogiques comme indiqué ci-après.

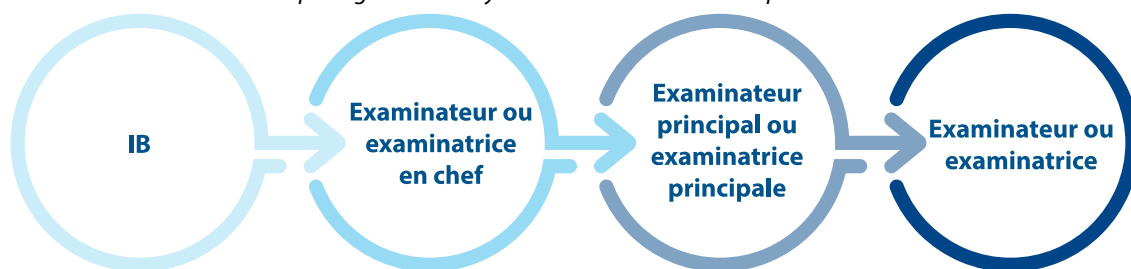
- **Ressources du PEI > Évaluation électronique du PEI > Ressources pour toutes les sessions > *Prévision des notes finales de l'IB***
- **Ressources du Programme du diplôme > Évaluation > Ressources pour toutes les sessions > *Prévision des notes finales de l'IB***
- **Ressources du POP > Évaluation > Ressources pour toutes les sessions > *Prévision des notes finales de l'IB***

## Les rôles et les responsabilités entourant le processus d'évaluation

- Chaque rôle occupé dans le processus d'évaluation est assorti de responsabilités particulières et d'un ensemble de compétences.
- Il peut arriver que la même personne remplisse plusieurs fonctions à différentes étapes du cycle d'évaluation.
- Certaines responsabilités incombent à l'IB et d'autres sont conférées à des spécialistes de la communauté de l'IB. Dans le dernier cas de figure, il revient à l'IB de procéder à la validation finale et de veiller au maintien de la qualité de tous les aspects des évaluations.

Figure 22

*Les protagonistes du cycle d'évaluation et leurs responsabilités*



### **IB**

- Responsabilité de tous les aspects de l'évaluation
- Responsabilité des processus liés à l'évaluation, tels que le recrutement des examinateurs et examinatrices, le contrôle de la qualité de leur travail et la publication des résultats
- Prise de décisions pour toute question relative à la mauvaise conduite des élèves, à la mauvaise administration des établissements, aux aménagements à des fins d'accès et d'inclusion, et aux circonstances défavorables
- Acceptation ou contestation des seuils d'attribution des notes finales recommandés par l'examineur ou examinatrice en chef

### **Examineur principal ou examinatrice principale**

- Responsabilité d'une composante donnée
- Décision finale concernant la note à attribuer aux réponses des élèves dans cette composante
- Vérification auprès de l'ensemble des examinateurs et examinatrices de leur bonne compréhension de la norme de notation définie pour la composante
- Aide apportée à l'examineur ou examinatrice en chef pour établir les seuils d'attribution des notes finales pour la composante

### **Examineur ou examinatrice en chef**

- Vue d'ensemble de toutes les composantes de sa matière (ou de son groupe de matières)
- Garantie de la cohérence des normes entre les différentes composantes, y compris pour la rédaction des examens
- Clôture de tous les débats portant sur l'évaluation
- Recommandation des seuils d'attribution des notes finales à l'IB

### **Examineur ou examinatrice**

- Notation du travail des élèves en respectant la norme établie par l'examineur principal ou examinatrice principale

## Les examinateurs principaux ou en chef et examinatrices principales ou en chef

L'examinateur principal ou examinatrice principale :

- gère une composante en particulier, comme l'évaluation interne ;
- établit les normes de notation pour chaque session en veillant à la cohérence entre les sessions, et explique ces normes aux examinateurs et examinatrices lors de discussions sur la normalisation et de l'application des modèles d'assurance de la qualité ;
- aide l'examinateur ou examinatrice en chef à faire des recommandations concernant les seuils d'attribution des notes finales pour la composante à sa charge ;
- participe généralement à la création du contenu des évaluations, sauf en cas de conflit d'intérêts ;
- intervient à la demande de l'IB à chaque session d'examens en tant que spécialiste de l'éducation externe et apporte généralement sa contribution à plusieurs sessions.

L'examinateur ou examinatrice en chef est responsable du maintien de la qualité de plusieurs composantes connexes. Dans le Programme du diplôme et le POP, cette personne est chargée de toute une matière, alors que dans le cadre du PEI (qui ne compte généralement qu'une composante par discipline), elle est chargée d'un groupe de matières (par exemple, Sciences ou Acquisition de langues). Cette personne joue le rôle de spécialiste de l'éducation dans son domaine.

L'examinateur ou examinatrice en chef :

- veille à la qualité des composantes connexes pour toute une matière (Programme du diplôme ou POP) ou pour tout un groupe de matières (PEI) ;
- joue le rôle de spécialiste de matière de l'IB, en veillant au respect des normes et à la cohérence ;
- aide à régler les conflits ;
- dirige le processus d'attribution des notes finales et recommande des seuils d'attribution des notes finales à l'IB ;
- exerce souvent aussi la fonction d'examinateur principal ou examinatrice principale pour une composante précise ;
- intervient pour le compte de l'IB pour une période donnée.

Dans le cas des matières présentées par un petit nombre d'élèves, l'IB attribue la fonction d'examinateur ou examinatrice responsable à un examinateur ou une examinatrice. Ses responsabilités sont les mêmes que celles des examinateurs ou examinatrices en chef, sa matière comptant généralement moins d'examineurs et examinatrices toutes fonctions confondues. Bien que cela ne soit pas abordé dans la présente publication, il est également demandé aux examinateurs et examinatrices en chef de collaborer avec l'IB pour discuter de ses programmes d'études et évaluations, et les améliorer lors d'activités de révision des programmes d'études.

### Les autres examinateurs et examinatrices

Les examinateurs et examinatrices ont pour responsabilité de noter le travail envoyé pour l'évaluation en respectant la norme établie par l'examinateur principal ou examinatrice principale. Ces personnes doivent prouver qu'elles ont compris cette norme et qu'elles l'appliquent au moyen du modèle d'assurance de la qualité.

Les examinateurs et examinatrices proposent leur candidature à l'IB pour corriger les travaux d'une session. Il s'agira impérativement de spécialistes de la matière visée, qui occuperont généralement des fonctions d'enseignement de cette matière. Une expérience pédagogique auprès d'élèves du groupe d'âge concerné est aussi requise. L'IB vérifie l'adéquation des profils au moyen d'un processus de candidature rigoureux, qui impose notamment de désigner des personnes pouvant fournir des recommandations. En règle générale, les examinateurs et examinatrices peuvent noter uniquement une composante réalisée dans le cadre des cours et une composante de l'examen à chaque session. Les périodes de notation de ces deux composantes ne se chevaucheront pas. La rémunération est fonction des copies d'examen notées « en

conditions réelles ». Les réponses de validation et de contrôle (voir la section « [Le modèle d'assurance de la qualité](#) ») ne sont pas prises en compte, car elles prouvent uniquement à l'IB la bonne compréhension de la norme de notation.

Les chefs d'équipe sont des examinateurs et examinatrices de grande expérience à qui l'IB demande d'aider d'autres examinateurs et examinatrices à comprendre la norme de notation. Les chefs d'équipe apportent leur soutien aux examinateurs et examinatrices tout au long du processus de validation et leur font un retour d'information si leur notation des réponses de contrôle indique un éloignement trop important de la norme définitive. Pour chaque matière, le nombre de chefs d'équipe dépendra du nombre global d'examineurs et examinatrices, lequel est fixé en fonction de la taille de la cohorte. Dans le cas d'une matière présentée par un petit nombre d'élèves, l'examineur principal ou examinatrice principale apporte ce soutien à l'ensemble des examinateurs et examinatrices.

Les membres de l'équipe de normalisation sont habituellement des examinateurs et examinatrices ayant de l'expérience qui prennent part aux discussions qualitatives qui débouchent sur le développement d'un modèle d'assurance de la qualité. Ces personnes discutent de l'adoption de la norme par l'examineur principal ou examinatrice principale afin de veiller à ce qu'elle puisse être convenablement appliquée et respectée.

Pour de plus amples informations sur le rôle des examinateurs et examinatrices, veuillez vous référer à la vidéo intitulée « [À la rencontre des examinateurs et des examinatrices : qui sont ces personnes et que font-elles ?](#) ».

### **Les responsabilités du personnel de l'IB**

Les responsables de matière et les analystes des opérations d'évaluation font partie des personnes clés qui garantissent la fiabilité de l'évaluation du point de vue de l'IB.

Au sein de l'IB, plusieurs équipes gèrent les demandes d'épreuves modifiées, d'aménagement à des fins d'accès et d'inclusion ou en cas de circonstances défavorables ainsi que les problèmes d'intégrité intellectuelle, le contrôle de la qualité du travail des examinateurs et examinatrices, et les facteurs de révision de notation pour les établissements.

Le directeur ou la directrice de l'évaluation travaille avec l'assistance de la ou du responsable en chef du développement et de la logistique de l'évaluation, et de la ou du responsable en chef des principes et pratiques de l'évaluation. Il lui revient d'étudier les seuils d'attribution des notes finales recommandés par l'examineur ou examinatrice en chef à l'issue de la réunion de délibérations pour chaque matière, et de les accepter ou de demander à l'examineur ou examinatrice en chef de revoir ses recommandations.

Enfin, il incombe à l'IB de prendre toutes les décisions qui concernent le cycle d'évaluation. Et s'il fait appel à des spécialistes externes pour l'aider à produire des évaluations équitables et de qualité, il conserve la responsabilité finale des décisions prises.

### **Les rôles se rapportant à la rédaction des examens**

La personne chargée de la création de contenu d'évaluations a un rôle bien défini au sein du processus de rédaction des examens, qui peut englober des activités de rédaction, de révision et de vérification. Comme indiqué ci-après, plusieurs compétences fondamentales sont de mise pour pouvoir créer du contenu d'évaluations.

- Créativité : il s'agit de créer des questions intéressantes et alignées sur le programme d'études.
- Sensibilité culturelle : il faut veiller à ce que les questions soient exemptes de biais.
- Compétences de communication : il est essentiel de rédiger des questions claires et non ambiguës traduisant et évaluant les compétences visées, pour assurer une notation fiable.

La production de l'évaluation implique plusieurs autres rôles, qui se rapportent notamment aux activités qui suivent.

- Conception technique : mise en forme des questions pour les examens
- Traduction : conversion de l'évaluation dans les langues requises

- Révision externe : retour d'information sur la longueur et la clarté de l'évaluation en passant les examens comme des élèves

Pour de plus amples informations sur la manière dont les examens sont rédigés, veuillez vous référer à la section « [La production de l'évaluation](#) ».

## L'organisation hiérarchique des examinateurs et examinatrices

- L'examineur principal ou examinatrice principale prend la décision finale quant à la note à attribuer dans une composante donnée. La norme ainsi établie devient la norme définitive, qui doit être respectée par le reste des examinateurs et examinatrices.
- L'IB part du principe que les personnes ayant aidé l'examineur principal ou examinatrice principale à établir la norme ont une meilleure compréhension de cette norme que les personnes qui en ont eu connaissance avec le modèle d'assurance de la qualité ou auprès de leur chef d'équipe.
- L'adjectif « superviseur/superviseuse » fait référence à la proximité avec l'examineur principal ou examinatrice principale dans la hiérarchie. Cela n'a aucun rapport avec la durée de l'expérience dans la correction d'examens ni avec une expérience d'enseignement combinée.

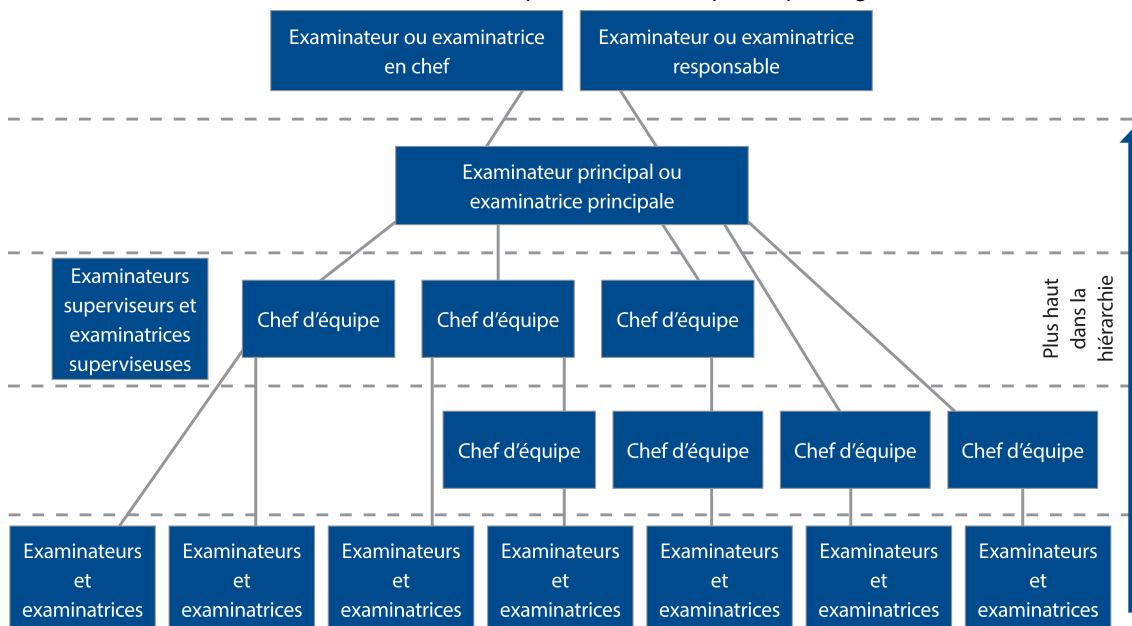
Le principe qui sous-tend la notation de l'IB est que l'examineur principal ou examinatrice principale décide en définitive de la note correcte à attribuer au travail de l'élève. Même si l'IB reconnaît que les définitions d'une note « correcte » peuvent être très différentes et tout aussi raisonnables les unes que les autres, il n'est pas juste que la note diffère en fonction de la personne qui corrige le travail. C'est pourquoi l'examineur principal ou examinatrice principale établit la norme et demande à chaque examinateur et examinatrice de s'y conformer.

L'examineur principal ou examinatrice principale reçoit l'aide d'une équipe d'examineurs superviseurs et examinatrices superviseuses pour établir la norme. Étant donné que ces examinateurs et examinatrices ont l'occasion de discuter de la norme avec l'examineur principal ou examinatrice principale, l'IB considère que leur compréhension n'en sera que meilleure. La conformité de la notation de l'ensemble des examinateurs et examinatrices avec la norme établie par l'examineur principal ou examinatrice principale est cependant consignée à chaque session, pour donner une bonne indication de la fiabilité de la correction des examens.

Il existe une hiérarchie parmi les examinateurs et examinatrices. Pour une matière comptant un grand nombre d'élèves, cette hiérarchie est structurée comme le montre la figure 23.

Figure 23

Hiérarchie des examinateurs et examinatrices pour une matière qui compte un grand nombre d'élèves



Cette hiérarchie n'est pas fixe. Si les données issues des réponses de validation ou de contrôle montrent qu'un examinateur ou une examinatrice a une meilleure compréhension de la norme de l'examineur principal ou examinatrice principale qu'une personne placée plus haut dans la hiérarchie, alors cet élément sera pris en considération. Être chef d'équipe implique de pouvoir expliquer la norme et pas uniquement de l'appliquer ; c'est pourquoi la fiabilité n'est pas le seul critère à prendre en compte pour désigner les chefs d'équipe.

Dans l'idéal, nul ne devrait avoir à prendre une décision en se référant aux priorités données par la hiérarchie, car toutes les notations se font selon la même norme. Dans la réalité, cependant, il s'agit d'une méthode fondamentale pour déterminer quelles notes l'IB doit utiliser. Si la situation est particulièrement difficile ou controversée, l'IB soumet la copie d'examen à l'examineur principal ou examinatrice principale, qui décide de la note qui convient.

## L'intégrité de l'évaluation

- Les évaluations de l'IB ne peuvent être équitables que si l'ensemble des élèves bénéficient d'une égalité des chances en matière de réussite.
- Les cas de mauvaise administration des établissements et les actions des élèves qui enfreignent la politique d'intégrité intellectuelle défavorisent les élèves qui respectent les règles. C'est pourquoi l'IB fait tout son possible pour empêcher ces comportements.
- Les règles de l'IB, énoncées dans la section « Règlement général » des *Procédures d'évaluation* spécifiques à chaque programme et dans la *Politique d'intégrité intellectuelle*, ont été établies pour limiter les possibilités de mauvaise conduite des élèves et de mauvaise administration des établissements. Toutefois, seul l'établissement scolaire peut créer une culture de l'apprentissage refusant et dénonçant la mauvaise conduite.
- Même si certaines formes d'évaluation se prêtent moins que d'autres à la mauvaise conduite, l'IB garde pour principe que la pertinence du construit (c'est-à-dire le fait d'évaluer ce que l'on souhaite réellement évaluer) doit rester la priorité absolue lors de la conception des évaluations.
- Toute suspicion de mauvaise conduite des élèves ou de mauvaise administration des établissements doit être signalée à l'IB.

## Définition de l'intégrité intellectuelle

L'intégrité intellectuelle est un principe directeur de l'éducation ainsi qu'un choix d'agir de manière responsable garantissant à qui la respecte que les autres lui accorderont leur confiance en tant que personne. Ce principe directeur constitue la base de toute prise de décision et de tout comportement éthiques dans l'apprentissage, l'enseignement et l'évaluation ainsi que dans la production de travaux scolaires légitimes, authentiques et honnêtes. L'intégrité intellectuelle ne doit pas être considérée comme une série de règles strictes à imposer, mais prendre la forme d'une culture positive au sein de l'établissement scolaire et de la communauté élargie, y compris chez les parents et les personnes exerçant une tutelle, ce qui conduira naturellement à des résultats d'évaluation plus équitables pour tous et toutes.

Pour être valide, une évaluation doit donner une idée précise de l'accomplissement de l'élève par rapport au reste de la cohorte qui a passé cette évaluation. C'est pourquoi l'IB met tout en œuvre pour avoir des approches cohérentes de la notation, de l'attribution des notes finales et de l'élimination des biais de ses examens. Les règles et règlements établis à cet effet sont un autre aspect de la création de cette égalité des chances et de cette garantie de l'équité.

La *Politique d'intégrité intellectuelle* définit la mauvaise conduite comme un comportement délibéré ou fortuit susceptible de procurer un avantage déloyal à l'élève qui l'adopte ou à toute autre personne dans une ou plusieurs composantes d'évaluation. Une telle activité affecte non seulement l'élève qui l'entreprend, mais encore toutes les personnes qui passent l'évaluation en même temps, car sa validité s'en trouve réduite. L'IB prend donc la mauvaise conduite très au sérieux, et les mesures de prévention comme ses conséquences sont expliquées en détail dans la *Politique d'intégrité intellectuelle*.

La conception de l'évaluation joue un rôle capital pour empêcher la mauvaise conduite, car certains types d'évaluations sont plus faciles à contrôler que d'autres. Par exemple, il est plus difficile de demander de l'aide à quelqu'un pour un examen que pour un travail évalué en interne. Même si ces éléments sont pris en considération pour définir les approches de l'évaluation, l'IB estime que la pertinence du construit ne doit pas être sacrifiée pour empêcher toutes les possibilités de mauvaise conduite.

La facilité d'application est un autre aspect essentiel de la validité, et l'IB doit aussi en tenir compte lorsqu'il s'efforce d'empêcher les possibilités de mauvaise conduite. La préparation des salles d'examen, notamment

pour l'évaluation numérique, peut présenter un vrai défi pour les établissements scolaires, et cela ne doit pas être négligé. L'IB s'efforce de penser aux établissements lorsqu'il fixe ses différentes règles, et il tient beaucoup à ce que les coordonnateurs, coordonnatrices et chefs d'établissement lui fassent part des pratiques jugées positives ou au contraire compliquées.

Pour en savoir plus sur l'intégrité intellectuelle, et sur le sens d'une action éthique, veuillez vous référer à la vidéo intitulée « [L'intégrité intellectuelle et l'IB : faire les bons choix](#) ».

## Les conflits d'intérêts

Au sein de l'IB, l'accès au contenu des examens est rigoureusement contrôlé, et l'organisation gère activement tous les liens avec les élèves passant les examens qui pourraient créer un conflit d'intérêts. Si un tel lien est avéré, alors les responsabilités des membres du personnel sont réorganisées.

L'IB part du principe qu'aucun examinateur ou examinatrice ne peut noter le travail de ses propres élèves ni le travail d'élèves se trouvant dans des établissements occasionnant un conflit d'intérêts (par exemple, un établissement dans lequel une activité de tutorat est exercée par un examinateur ou une examinatrice dans une autre matière ou un établissement récemment quitté). Dans l'éventualité extrêmement rare où cela est inévitable pour des raisons qui échappent au contrôle de l'IB (par exemple, l'impossibilité de trouver une autre personne qualifiée pour noter selon la norme requise), une seconde personne indépendante révisé la notation. L'IB devra alors s'assurer que rien n'indique qu'une norme différente a été appliquée aux élèves de l'examineur ou examinatrice de départ.

## Comment gérer la mauvaise administration des établissements et la mauvaise conduite des élèves

Pour chaque programme, la section « Règlement général » des *Procédures d'évaluation* correspondantes et les documents intitulés *Déroulement des examens* et *Déroulement des examens sur ordinateur du Programme d'éducation intermédiaire de l'IB* rassemblent les règles et les directives pour minimiser les risques de mauvaise administration. Pour les examens au format papier, le livret intitulé *Stockage sécurisé du matériel confidentiel de l'IB destiné aux examens* contient des instructions concernant le stockage du matériel confidentiel de l'IB dans les locaux de l'établissement ainsi que les mesures devant être prises par les établissements en vue de se conformer à cette politique.

Les examens doivent être surveillés conformément aux instructions données dans les documents *Déroulement des examens* (pour le Programme du diplôme et le POP) et *Déroulement des examens sur ordinateur du Programme d'éducation intermédiaire de l'IB* (pour le PEI). Les surveillantes et surveillants doivent faire preuve de vigilance pendant toute la durée de l'examen pour s'assurer qu'aucune mauvaise conduite n'est adoptée.

Pour s'assurer de l'observation de toutes ces règles, l'IB procède à des inspections non annoncées des dispositions d'examens dans les établissements scolaires durant les sessions d'évaluation. Il s'agit de vérifications ponctuelles de la mise en œuvre par les établissements des processus de l'IB. L'IB compte cependant fortement sur les chefs d'établissement et sur les coordonnateurs et coordonnatrices des programmes pour empêcher les éventualités de mauvaise administration, car ces personnes sont les mieux placées pour garantir que des normes de qualité sont maintenues dans leur établissement jour après jour. Il leur revient en outre de s'assurer que la culture de leur établissement encourage le personnel enseignant et les élèves à adopter les meilleures pratiques qui soient et l'intégrité la plus rigoureuse possible.

Outre le respect des normes d'intégrité intellectuelle pendant les examens, il est important de veiller à ce que le travail effectué en classe (et à la maison) soit bien celui des élèves et qu'aucun plagiat ne soit commis.

De nombreux sites Web et outils d'intelligence artificielle (IA) proposent d'« aider » les élèves pour leur travail. La meilleure défense contre ce type de mauvaise conduite consiste à compter sur l'enseignant ou enseignante qui aura travaillé auprès de l'élève et pourra déterminer à quels moments le travail remis n'est pas le reflet du niveau habituel. C'est pourquoi il est demandé à l'élève et à l'enseignant ou enseignante de confirmer que le travail est bien celui de l'élève. Il ne suffit pas de compter sur un logiciel de détection du

plagiat : les enseignantes et enseignants doivent en parallèle travailler auprès de leurs élèves pendant la rédaction de leurs travaux d'évaluation interne, afin d'en vérifier l'authenticité. Une fois qu'un travail lui est parvenu, l'IB n'est nullement obligé d'accepter une autre version en cas de plagiat avéré.

Aucun niveau de plagiat n'est accepté : toutes les citations empruntées à autrui et toutes les utilisations de logiciels d'IA doivent être correctement référencées, dans le respect des explications données dans les publications intitulées *Savoir citer et référencer ses sources* et *Politique d'intégrité intellectuelle*. Pour détecter les cas présumés de mauvaise conduite des élèves, l'IB utilise un certain nombre de logiciels. Si des preuves sont acquises, une enquête formelle est lancée.

Toutes les écoles du monde de l'IB sont tenues de mettre en place et d'entretenir une culture de l'intégrité intellectuelle, et des infractions répétées entraîneront le retrait de l'autorisation de proposer le programme de l'IB.

## Les difficultés posées par les examens internationaux et les zones horaires

L'IB compte des établissements à travers le monde, ce qui le place face à des défis particuliers, car contrairement à ce qui se produit avec la plupart des systèmes nationaux, il ne peut pas faire passer les examens à l'ensemble des élèves en même temps. Ainsi, une partie des élèves finit les examens avant que l'autre ne les ait commencés. De ce fait, l'IB demande aux élèves et au personnel enseignant de faire preuve d'une intégrité absolue pour ne pas profiter de la situation. Lorsque le grand nombre d'élèves justifie une telle mesure, des « zones d'examen » différentes sont prévues afin de réduire le temps entre la fin d'un examen pour les élèves d'une zone horaire et le début de ce même examen dans une autre zone horaire. Il peut toutefois exister de grandes différences entre les heures où l'examen débute.

Pour limiter les risques pouvant découler de ce décalage, l'IB a mis en place un certain nombre de mesures, dont un contrôle minutieux des sites Web et des médias sociaux sur lesquels le contenu de l'examen pourrait être diffusé. La règle qui interdit de faire sortir le contenu de l'examen de la salle vise à minimiser les risques de partage des questions en ligne. Pour contrecarrer le partage d'informations strictement interdit, les règles en matière de surveillance indiquent que si un seul examen est passé au cours de la matinée ou de l'après-midi et que sa durée est inférieure à deux heures, les élèves doivent faire l'objet d'une surveillance minimale de deux heures à compter du début de l'examen.

L'IB veille également à ce que les questions d'examen soient formulées de façon à évaluer la compréhension plutôt que le rappel des connaissances, ce qui signifie qu'il y a peu d'intérêt à connaître les questions à l'avance. Enfin, l'IB interdit aux élèves de discuter du contenu de l'examen pendant les 24 heures qui suivent sa tenue, ce qui supprime toute possibilité de « discussion innocente ».

L'IB cherche constamment à faire preuve d'innovation pour réduire les risques de tricherie dus aux différentes zones horaires.

## Les avantages de l'évaluation numérique

L'évaluation numérique présente un grand nombre d'avantages en matière de gestion de la mauvaise conduite. Non seulement la durée d'accès au contenu de l'examen est très limitée, mais encore l'IB est capable de contrôler à quel moment se produit cet accès et qui en est à l'origine.

En outre, toutes les modifications que l'établissement apporte à l'examen, comme l'octroi de pauses ou de temps supplémentaire, sont enregistrées et des justifications peuvent être demandées, au besoin. Cela signifie que l'IB peut déléguer davantage de responsabilités aux établissements scolaires pour apporter des modifications raisonnables, tout en garantissant l'égalité des pratiques.

L'évaluation numérique permet également d'avoir une compréhension plus poussée de l'expérience de l'examen vécue par les élèves, par l'intermédiaire de métadonnées qui peuvent servir de preuves à l'appui d'une enquête lorsqu'un incident relevant de la mauvaise conduite est rapporté. Par ailleurs, le format électronique des réponses numériques permet une vérification à grande échelle de l'existence de réponses similaires, chose qu'il est impossible d'envisager avec des copies manuscrites.

L'IB reconnaît que l'évaluation numérique sera exposée à de nouvelles formes de mauvaise conduite, notamment à des tentatives de piratage, et l'organisation dispose d'une série d'outils et de processus pour y faire face.

## Les ressources

Le personnel enseignant, les établissements scolaires et les élèves figurent à la meilleure place pour s'élever contre les cas de mauvaise conduite et y mettre un terme, en créant une culture les refusant et en traitant les cas éventuels avec vigilance. Cet aspect est abordé plus en détail dans la section intitulée « [Le renforcement d'un état d'esprit éthique](#) ». Des documents peuvent aider les établissements à créer une culture de l'intégrité intellectuelle, notamment la *Politique d'intégrité intellectuelle*, et sont disponibles sur la page « [Intégrité intellectuelle](#) » du site Web de l'IB.

Si vous avez besoin d'un soutien supplémentaire, si vous avez des doutes quant au traitement de certains cas de mauvaise conduite par votre établissement ou si vous pensez à de nouveaux problèmes se rapportant à l'intégrité intellectuelle, veuillez vous adresser au service [L'IB vous répond](#).

## Le caractère prioritaire de l'équité universelle

- Pour rester valide, l'évaluation formative et sommative doit tenir compte de la variabilité des situations d'apprentissage et refuser toute discrimination. Cette variabilité est intégrée aux évaluations de l'IB dès les premières étapes de leur conception.
- Le meilleur moyen de garantir l'équité est de concevoir une évaluation accessible à tous et toutes. Les évaluations de l'IB suivent les principes de la conception universelle, qui suppriment les obstacles susceptibles de nuire à leur validité. Les établissements scolaires doivent se conformer aux instructions fixées par l'IB pour garantir que les examens écrits, oraux et numériques se déroulent dans des conditions qui ne confèrent aucun avantage ni aucun désavantage aux élèves.
- La *Politique d'accès et d'inclusion* fait état des situations requérant des mesures complémentaires ou des aménagements pour rendre l'évaluation accessible, au moyen d'épreuves modifiées par exemple.
- En cas de circonstances imprévues susceptibles d'entraver l'accès à l'évaluation ou de nuire à sa validité, il est possible de mettre en place certaines mesures pour atténuer les effets négatifs de ces événements, tel qu'indiqué dans la *Politique en matière de circonstances défavorables*.
- Pour préparer les élèves convenablement aux évaluations sommatives et/ou aux évaluations de l'IB, il doit exister une véritable harmonisation entre l'apprentissage, l'enseignement et l'évaluation formative. Cette harmonisation s'avère particulièrement importante en cas d'aménagements à des fins d'accès et d'inclusion, qui sont octroyés à la condition d'être mis en place lors de l'apprentissage, de l'enseignement et de l'évaluation (en adéquation avec le mode de travail habituel de l'élève). Il est par ailleurs fondamental de rechercher cette harmonisation quel que soit le support de l'évaluation, numérique notamment.
- Le but de ces aménagements est de garantir l'équité pour l'ensemble des élèves. Lorsque l'IB doit prendre une décision, il doit tenir compte de ce qui est équitable pour l'intégralité des élèves.

Chaque élève doit pouvoir démontrer ses capacités dans des conditions d'évaluation aussi équitables que possible. Certaines conditions ordinaires d'évaluation peuvent désavantager une partie des élèves en les empêchant de montrer leur niveau d'accomplissement. Des situations ou circonstances défavorables échappant au contrôle des élèves peuvent également avoir une incidence sur leur travail. Ces deux paramètres doivent être pris en compte pour mettre en place des mesures garantissant l'équité.

Le meilleur moyen de garantir l'équité d'une évaluation est de faire en sorte que l'intégralité des élèves passe la même évaluation dans les mêmes conditions. Pour ce faire, les évaluations sont développées en portant une attention toute particulière aux diverses exigences qui peuvent s'imposer aux élèves. Pour tenir compte de la variabilité des situations d'apprentissage et satisfaire les besoins de l'ensemble des élèves, la méthode la plus efficace consiste à concevoir les évaluations en suivant les principes de la conception universelle de l'apprentissage et de la conception universelle de l'évaluation. Une fois les évaluations conçues selon ces principes, il s'agit de garantir l'accessibilité au moyen d'aménagements à des fins d'accès et d'inclusion, tel qu'indiqué dans la *Politique d'accès et d'inclusion* de l'IB.

Les aménagements à des fins d'accès et d'inclusion visent à supprimer les obstacles et à garantir l'accès à l'ensemble des élèves. Ces obstacles peuvent survenir lorsque l'évaluation ne tient pas convenablement compte de la neurodiversité, des incapacités, des difficultés ou des diverses situations socioémotionnelles.

Tous les ajustements raisonnables seront considérés pour chaque élève présentant des besoins spécifiques. Pour plus de renseignements, veuillez consulter la *Politique d'accès et d'inclusion*.

Les circonstances défavorables désignent des situations qui échappent au contrôle de l'élève et qui risquent d'avoir des effets négatifs sur son travail. La *Politique en matière de circonstances défavorables* explique comment aborder ces effets et les atténuer. Des aménagements à des fins d'accès et d'inclusion

peuvent être mis en place pour supprimer les obstacles et garantir que l'ensemble des élèves ont accès à l'évaluation. Ces aménagements ne confèrent aucun avantage aux élèves qui se heurtent à ces obstacles. Bien au contraire, ils sont la garantie d'un accès équitable en permettant à l'ensemble des élèves de montrer leurs compétences, leurs connaissances et leurs capacités sur un pied d'égalité avec leurs pairs.

## Les principes sous-tendant les aménagements à des fins d'accès et d'inclusion

Les principes qui sous-tendent les aménagements à des fins d'accès et d'inclusion sont définis dans la *Politique d'accès et d'inclusion*. L'évaluation doit tenir compte du stade de développement des jeunes, qui se reflète dans les objectifs globaux de chaque programme. Les aménagements à des fins d'accès et d'inclusion doivent être pris en considération dans tous les programmes, à la fois pour l'évaluation formative et pour l'évaluation sommative, et tout au long de l'expérience d'apprentissage et d'enseignement.

L'IB doit garantir qu'une note finale attribuée dans une matière reflète fidèlement le niveau d'accomplissement de l'élève. À cette fin, les mêmes normes d'évaluation s'appliquent à l'ensemble des élèves, que des aménagements à des fins d'accès et d'inclusion soient requis ou non.

Les aménagements à des fins d'accès et d'inclusion, y compris les ajustements raisonnables, sont des mesures prises en amont des examens pour que l'évaluation soit à la portée de l'élève en question. Ils ne peuvent pas être demandés rétrospectivement, que ce soit pour les examens oraux, écrits ou numériques.

Les aménagements demandés pour l'élève ne doivent pas lui donner un avantage dans une composante d'évaluation.

Les aménagements à des fins d'accès et d'inclusion sont destinés aux élèves qui possèdent l'aptitude à satisfaire à toutes les modalités d'évaluation.

Lorsque des aménagements à des fins d'accès et d'inclusion sont requis, l'établissement doit y pourvoir sur la base des critères de l'IB clairement définis dans sa *Politique d'accès et d'inclusion*. Les aménagements de l'évaluation demandés doivent être en adéquation avec le mode de travail habituel de l'élève. Il incombe à l'établissement de choisir les aménagements qui répondent le mieux aux besoins de l'élève tout au long du programme, en garantissant une transition fluide vers les évaluations sommatives. Les demandes d'aménagements qui ne correspondent pas aux pratiques habituelles de l'élève ne pourront être acceptées que dans des cas très exceptionnels et particuliers.

L'IB est attaché à une philosophie éducationnelle fondée sur la sensibilité internationale. La politique de l'IB résulte de l'analyse des pratiques en vigueur dans différents pays et régions (ou territoires) pour offrir les mêmes chances aux élèves ayant des besoins en matière d'aménagement de la procédure d'évaluation.

Toute demande d'aménagements à des fins d'accès et d'inclusion sera jugée au cas par cas. Une autorisation antérieure d'aménagements, accordée par l'IB ou par un autre organisme décernant des diplômes, n'influencera pas la décision relative à l'autorisation des aménagements demandés.

L'IB traite toutes les informations relatives aux élèves avec confidentialité. Si nécessaire, ces informations seront uniquement partagées avec le personnel de l'IB habilité et les membres du comité d'attribution des notes finales, qui ont pour instruction de respecter cette confidentialité.

Lorsqu'un établissement ne respecte pas les conditions définies par l'IB dans l'octroi des aménagements à des fins d'accès et d'inclusion, ou lorsqu'il en accorde sans l'aval de l'IB, l'élève risque de ne pas recevoir de note finale de l'IB pour la matière et le niveau concernés.

S'il est possible de prouver que le faible niveau de l'élève dans sa langue d'usage est dû à un besoin en matière de soutien à l'apprentissage identifié, des aménagements à des fins d'accès et d'inclusion peuvent être accordés.

La situation de l'élève ou les conditions défavorables qui l'affectent ne doivent pas être portées à la connaissance des examinateurs et examinatrices.

Pour le travail évalué en interne, le personnel enseignant ne doit apporter aucun ajustement dans sa notation.

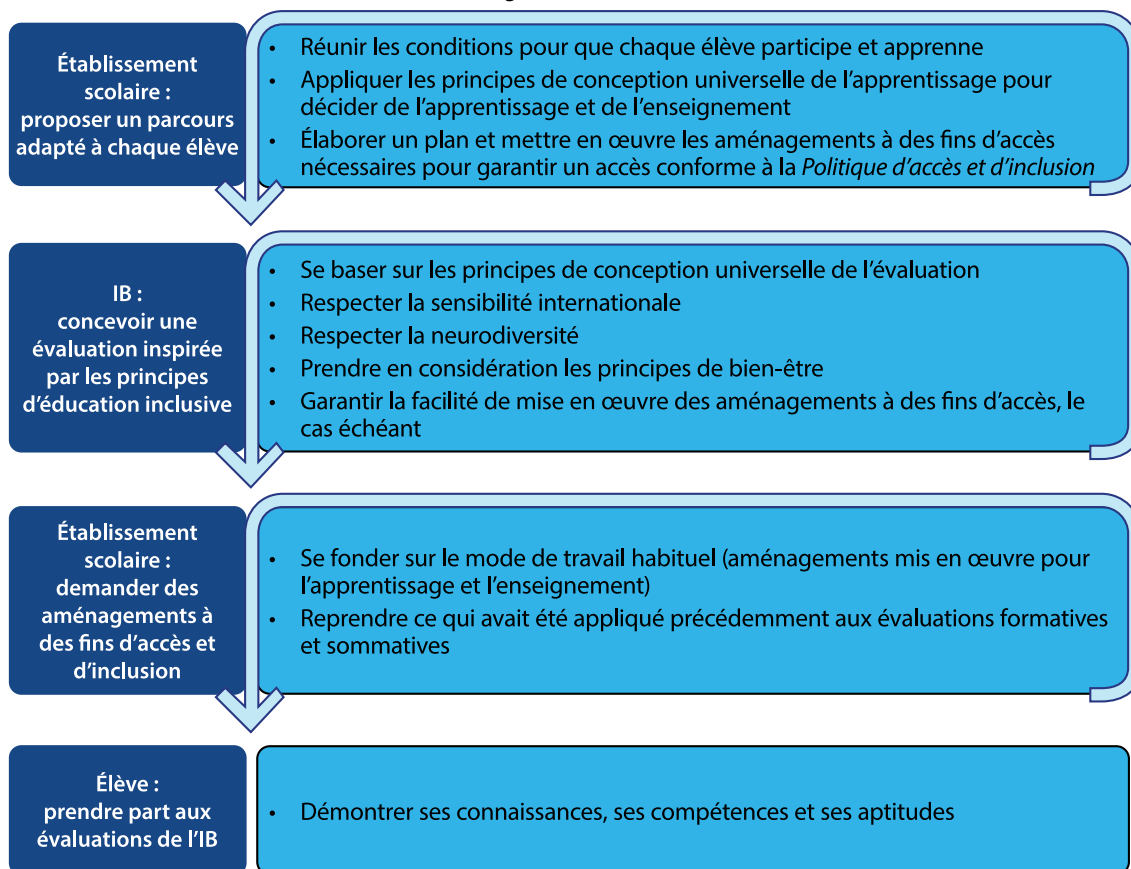
La liste des aménagements à des fins d'accès et d'inclusion disponibles est régulièrement mise à jour. D'autres aménagements suggérés par un établissement pourront être examinés, à condition de pouvoir être mis à la disposition de l'ensemble des élèves ayant les mêmes besoins.

Lorsque le comportement de l'élève et/ou la nature des aménagements à des fins d'accès et d'inclusion autorisés risquent de gêner les autres élèves pendant un examen, l'élève en question doit passer l'épreuve dans une pièce séparée et sous surveillance conformément aux réglementations régissant le déroulement des examens du programme.

Le service L'IB vous répond doit être informé dans les meilleurs délais de toute difficulté liée à la nature des aménagements à des fins d'accès et d'inclusion ou de toute difficulté imprévue éprouvée par l'élève pendant les examens.

Figure 24

#### Mise en œuvre des aménagements à des fins d'accès et d'inclusion



## Les dispenses d'évaluation

Les dispenses ne sont normalement admises pour aucune composante d'évaluation. Cependant, lorsqu'une composante d'évaluation, ou une partie d'une composante, demande une fonction physiologique que l'élève ne possède pas, une dispense peut être autorisée. Toutes les possibilités d'ajustement raisonnable doivent avoir été étudiées avant de déposer une demande de dispense pour une composante d'évaluation. Une autorisation de dispense ne sera accordée que pour des motifs suffisants. L'incapacité physique de l'élève à réaliser les fonctions demandées pour la composante doit être clairement documentée, et ce, de façon exhaustive.

Pour connaître le détail des principes et des processus liés aux dispenses d'évaluation, veuillez vous reporter à la liste des aménagements à des fins d'accès et d'inclusion figurant dans la *Politique d'accès et d'inclusion*.

## Les possibilités d'aménagements à des fins d'accès et d'inclusion avec l'évaluation numérique

L'évaluation numérique permet à l'élève d'avoir un plus grand contrôle sur la manière dont l'évaluation est présentée. Les appareils informatiques proposent tout un éventail de polices de caractères, de tailles de texte et de couleurs pour répondre aux préférences et aux besoins individuels, et ce type d'ajustement peut facilement être mis à la portée de l'ensemble des élèves. S'il est vrai que l'adoption des principes de conception universelle améliore l'accessibilité, il n'en demeure pas moins que certains aménagements resteront nécessaires pour généraliser cet accès. C'est en raison des aménagements permis que les fonctionnalités numériques constituent une option digne d'intérêt. Les aménagements possibles sont énumérés dans la [Politique d'accès et d'inclusion](#).

À l'instar des examens sur papier, les aménagements à des fins d'accès et d'inclusion octroyés pour les examens numériques ont pour objectif de supprimer les obstacles. Pour garantir l'équité, une cohérence doit être assurée au niveau des aménagements à des fins d'accès autorisés pour les examens numériques et au format papier.

## Les circonstances défavorables

Les circonstances défavorables ou imprévues sont celles échappant au contrôle de l'élève et risquant d'avoir des effets négatifs sur ses résultats. Ce sont, par exemple, une maladie ou blessure temporaire, un stress grave, des circonstances familiales particulièrement éprouvantes, un deuil ou des événements pouvant menacer la santé ou la sécurité de l'élève. Les circonstances défavorables peuvent également inclure des événements qui affectent la communauté scolaire dans son ensemble, comme des troubles civils ou une catastrophe naturelle.

Les circonstances défavorables n'incluent pas les insuffisances dues au fait de l'établissement scolaire. Il incombe à l'établissement de s'assurer que l'ensemble des élèves respectent les exigences du programme et les modalités d'évaluation. Il lui revient aussi de gérer les problèmes liés au personnel enseignant.

Pour connaître le détail de ce qui est inclus dans la catégorie des circonstances défavorables et ce qui en est exclu, ainsi que le détail des ajustements proposés, veuillez consulter la [Politique en matière de circonstances défavorables](#).

## La conception universelle de l'évaluation

Les sections précédentes s'intéressaient à la gestion des situations nécessitant des aménagements à des fins d'accès et d'inclusion pour garantir un accès équitable aux évaluations de l'IB. La meilleure solution reste toutefois, avant toute chose, de concevoir des évaluations ne dressant pas d'obstacle à la participation. La conception universelle de l'évaluation insiste sur la prise en considération de l'accès, de l'inclusion, de l'équité, de la diversité culturelle et de la sensibilité. Conformément à cette approche, les évaluations de l'IB sont pensées de sorte à supprimer un grand nombre de barrières du modèle d'évaluation global, et ce, au cours du processus de conception, de développement et de mise en œuvre.

En créant dès le départ des évaluations plus inclusives et plus pertinentes avec le construit, les risques de devoir engager ou mettre en œuvre des aménagements ultérieurement s'en voient considérablement réduits, ce qui profite à un grand nombre d'élèves.

La conception universelle de l'évaluation est un aspect de la notion plus large de conception universelle de l'apprentissage, qui vise à créer des environnements d'apprentissage accessibles à l'ensemble des élèves. Cette approche s'appuie sur une vision élargie de la diversité, qui englobe la neurodiversité, les préférences, les capacités, les troubles, les centres d'intérêt, les cultures, les langues et les profils, entre autres.

Les principes de la conception universelle de l'apprentissage sont applicables à tous les domaines de l'éducation, notamment la conception du programme d'études et de l'évaluation, le développement du programme d'études et de l'évaluation ainsi que la gestion de l'établissement scolaire. Pour en savoir plus

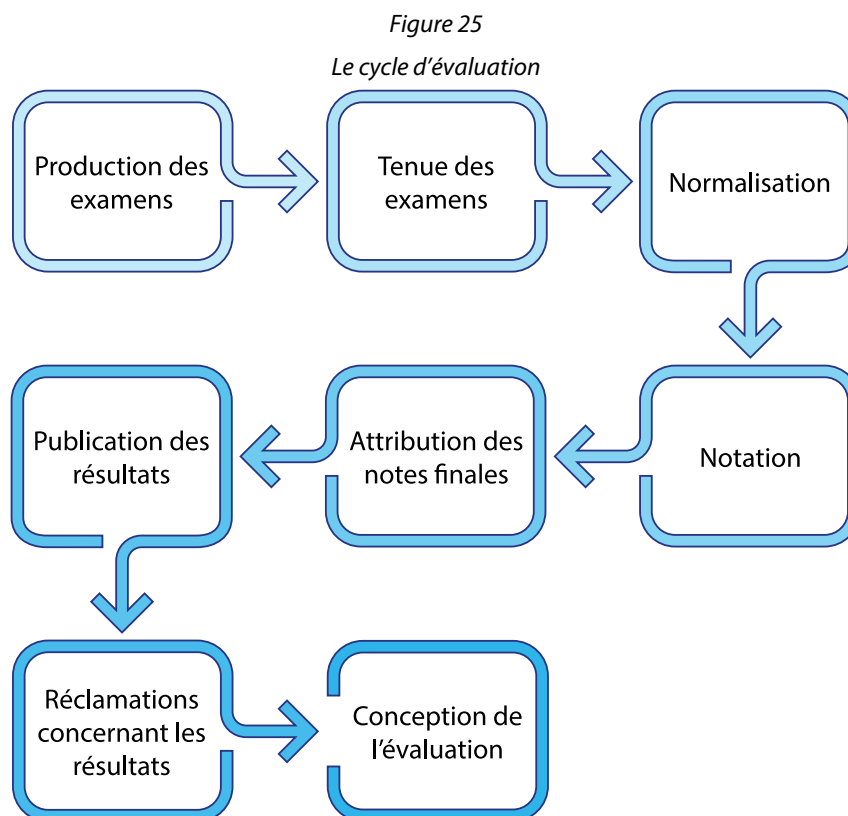
sur la conception universelle de l'apprentissage dans le contexte de l'IB, veuillez vous référer à l'ouvrage de Rao et coll. (2016).

## Le cycle de vie de l'évaluation

- Le processus de création des évaluations doit être imaginé comme un cycle continu dans lequel chaque étape dépend de la précédente et influence la suivante.
- Il faut en moyenne 18 mois pour créer un examen (et le barème de notation correspondant) ; c'est pourquoi l'IB crée des examens pour différentes sessions en parallèle.

L'évaluation s'inscrit dans un modèle cyclique qui compte plusieurs étapes : création, organisation de sessions pour les élèves, notation réalisée par des examinateurs et examinatrices, et publication des résultats. La clé de voûte de ce processus est que l'IB tire des enseignements de chaque session pour améliorer la qualité de l'évaluation pour la session suivante.

La figure 25 n'est qu'une illustration possible de ce processus. Même si de nombreuses étapes pourraient être divisées ou regroupées différemment, ce schéma explique de façon adaptée le cycle de vie d'une évaluation. Un résumé de chaque étape et un lien vers la section correspondante sont donnés dans le tableau 6.



*Tableau 6*  
*Les étapes du cycle de vie d'une évaluation*

| Étape du cycle d'évaluation   | Description  |
|-------------------------------|--|
| Attribution des notes finales | Étape au cours de laquelle les examinateurs superviseurs et examinatrices superviseuses décident de la manière de convertir des notes (qui sont propres à l'examen) en notes finales (qui conservent la même signification quel que soit l'examen) |

| Étape du cycle d'évaluation           | Description   |
|---------------------------------------|---|
|                                       | concerné). Parmi les activités qui font suite à l'attribution des notes finales, il convient de citer le contrôle de la distribution des notes finales données par les examinateurs et examinatrices ainsi que l'identification des élèves susceptibles de recevoir une note finale incorrecte en se fondant sur les preuves d'un problème particulier. |
| Conception de l'évaluation            | Processus guidé par les résultats des élèves et l'interprétation des questions d'examen. L'IB s'appuie sur ces éléments pour améliorer ce qui doit être évalué et ajuster l'approche adoptée, notamment le nombre de tâches et leur type.   |
| Normalisation                         | Processus au cours duquel les examinateurs superviseurs et examinatrices superviseuses expliquent à leurs équipes comment noter les travaux des élèves et sélectionnent les copies comportant les notes définitives qui seront utilisées pour la notation lors des processus d'assurance de la qualité.   |
| Notation                              | Activité au cours de laquelle les examinateurs et examinatrices étudient le travail de chaque élève et décident de la note à attribuer, en suivant impérativement les instructions données par l'examineur principal ou examinatrice principale, la qualité de leur correction étant garantie par des contrôles réguliers.                              |
| Production des examens                | Processus de création de chaque examen, qui couvre un spectre très large : choix des thèmes abordés dans les questions, rédaction et révision des questions, traduction dans d'autres langues, mise en forme pour parvenir à la présentation souhaitée et, enfin, contrôles de la qualité.  |
| Publication des résultats             | Diffusion des résultats aux établissements scolaires et aux élèves.   |
| Réclamations concernant les résultats | Possibilité offerte aux établissements scolaires d'indiquer à l'IB toute erreur suspectée dans le processus de notation des examens et de demander une vérification du travail de l'élève en question.  |
| Tenue des examens                     | Étape au cours de laquelle les élèves passent les évaluations dans les établissements scolaires.  |

Pour en savoir plus sur le cycle de vie de l'évaluation, veuillez vous référer à la vidéo intitulée « [Le cycle de vie d'un examen : quelles sont les étapes nécessaires à la création d'un examen \(de l'élaboration à la publication des notes finales\) ?](#) ».

## Les effets de l'évaluation numérique sur le cycle d'évaluation

Le principe du cycle d'évaluation n'est nullement altéré par l'introduction de l'évaluation numérique. Cela permet au contraire à l'IB de gagner du temps dans certaines étapes du cycle (par exemple, fin de l'expédition par voie postale des examens et des copies au format papier aux centres de numérisation). Il reste toutefois nécessaire de respecter les différentes étapes qui composent le cycle d'évaluation.

## La production de l'évaluation

- Au cours de sa production et avant de pouvoir être passée par les élèves, une évaluation doit passer par plusieurs étapes de finalisation. En voici quelques exemples :
  - rédaction de l'évaluation et validation du contenu ;
  - élaboration et relecture ;
  - facilité d'utilisation ;
  - traduction ;
  - transmission de l'évaluation aux établissements scolaires.
- Toutes les modifications requises par les élèves (aménagement à des fins d'accès et d'inclusion) sont également examinées lors du processus de finalisation.
- Les principes du cycle de production de l'évaluation demeurent inchangés, que les examens soient sur papier ou au format numérique.

## De bonnes composantes d'évaluation

Outre les qualités caractéristiques d'une bonne évaluation qui sont décrites dans la « Section A : les principes de l'évaluation », l'IB attend des équipes chargées du développement des programmes d'études et de la rédaction qu'elles gardent à l'esprit un certain nombre de points essentiels.

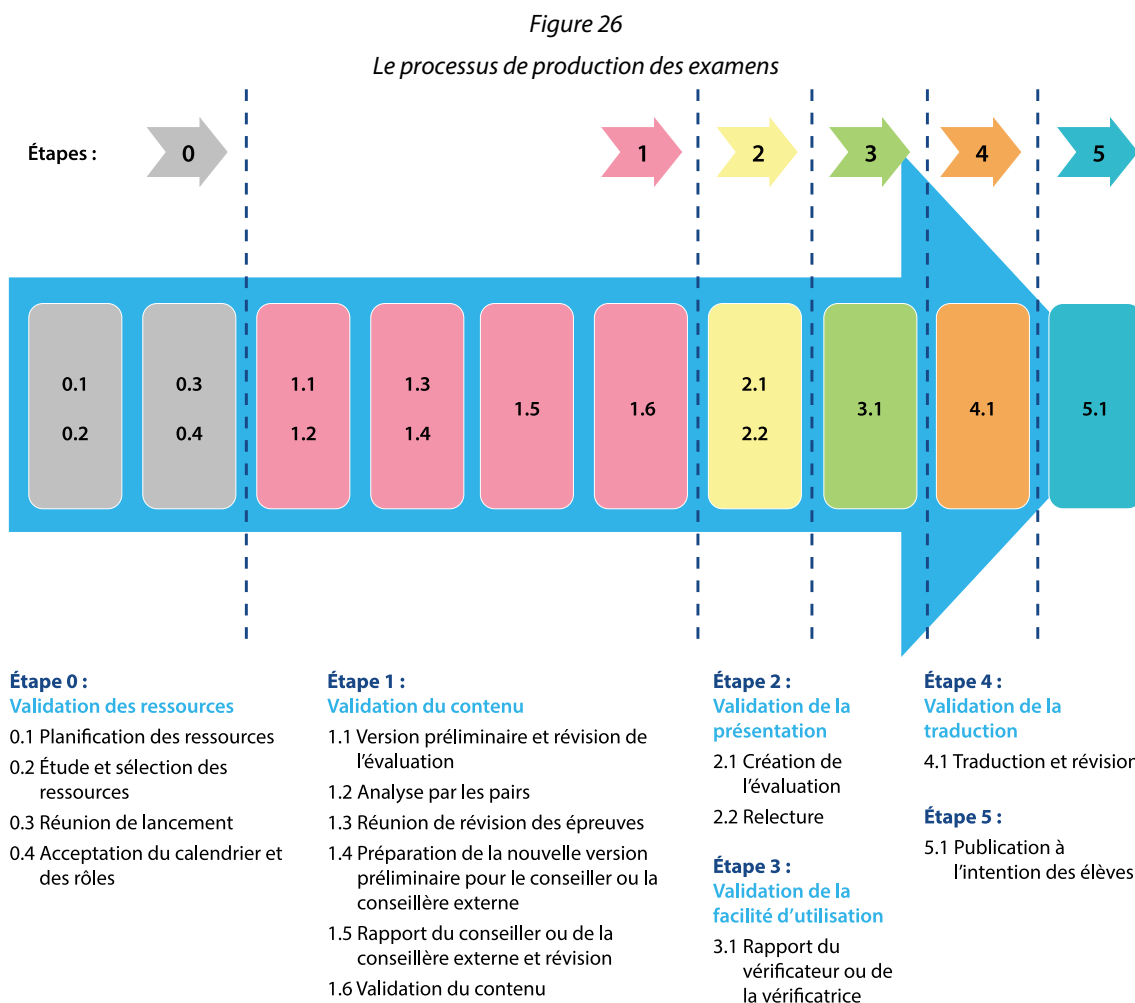
- L'évaluation doit donner aux élèves la possibilité de montrer leurs capacités et non indiquer ce qui leur est impossible de faire.
- L'évaluation doit être à la portée du plus grand nombre d'élèves possible, tout en permettant une différenciation raisonnable.
- Les meilleures évaluations sont accessibles à l'intégralité des élèves. Pour que les aménagements à des fins d'accès et d'inclusion nécessaires soient les moins nombreux possibles, l'IB se sert de la conception universelle dans tout le processus de production de l'évaluation.
- L'évaluation doit porter uniquement sur le programme d'études, tel qu'il apparaît dans le guide pédagogique.
- Les modalités de la notation sont décidées lors de l'élaboration de l'évaluation, afin que des points soient bien attribués à ce qui est censé être évalué. En d'autres termes, le barème de notation est élaboré en même temps que les examens et des orientations sont données sur les marges de tolérance prévues pour les examinateurs et examinatrices.

### Les mots-consignes

Les mots-consignes sont les termes et expressions clés utilisés dans le programme et les questions d'examen pour indiquer ce qui est demandé en réponse à une instruction particulière. Ils suggèrent également le type et le degré de réponse attendus des élèves.

Les mots-consignes constituent un défi particulier lorsqu'ils doivent être traduits dans d'autres langues, en raison des subtilités qui peuvent facilement être perdues. Pour que cela ne soit pas le cas, l'IB publie dans les guides pédagogiques ses mots-consignes dans chaque langue d'usage et explique leur sens. Lors de la vérification des examens traduits, il est donc fondamental de veiller à ce que le bon mot-consigne ait été utilisé.

# Présentation de la production des examens



La figure 26 illustre les différentes étapes qui constituent le processus de production des examens. La création d'une évaluation est un processus lent qui nécessite environ 18 mois au cours desquels interviennent de nombreux spécialistes externes et des membres du personnel de l'IB : auteurs et autrices, responsables de matière, spécialistes de matière, membres du service d'édition, concepteurs et conceptrices, vérificateurs et vérificatrices externes et linguistes, entre autres.

Il est crucial que le barème de notation soit élaboré en même temps que l'examen. Cette simultanéité garantit que les questions et la notation seront alignées dans ce qu'elles cherchent à évaluer.

## Le processus conduisant à la validation du contenu

La première étape de ce processus consiste à déterminer qui rédigera les examens. Il sera généralement demandé aux examinateurs principaux et examinatrices principales de s'en charger, mais l'IB vérifie certains éléments, comme les éventuels conflits d'intérêts en raison de liens avec certains établissements scolaires ou de l'animation d'ateliers de formation.

Comme indiqué ci-après, différents rôles sont associés à la production des examens.

- Auteur ou autrice (membre externe)
- Conseiller, conseillère ou responsable de la normalisation (membre externe)
- Vérificateur, vérificatrice ou personne parlant couramment la langue (membre externe)
- Responsable de matière (membre du personnel de l'IB)
- Éditeur ou éditrice responsable de la production (membre du personnel de l'IB)
- Correcteur ou correctrice (membre du personnel de l'IB)

- Concepteur ou conceptrice (membre du personnel de l'IB)
- Traducteur ou traductrice (membre externe)
- Réviseur ou réviseuse linguistique (membre externe)

Lorsque l'IB produit des examens pour plusieurs zones horaires, des processus de production des examens totalement indépendants sont suivis.

### La rédaction et la validation du contenu

Les auteurs et autrices doivent produire une version préliminaire de l'examen et une version préliminaire du barème de notation. Ces versions sont ensuite confiées à d'autres auteurs, autrices et responsables de matière, pour assurer plusieurs niveaux de révision. Pour certaines matières, il s'ensuit une réunion officielle de révision des épreuves, au cours de laquelle les éléments suivants sont passés en revue :

- niveau d'exigence cognitive des questions et difficulté globale de l'examen ;
- éventuel biais ;
- accessibilité ;
- programme d'études couvert ;
- correspondance entre l'examen et le modèle d'évaluation publié ;
- lien entre le nouvel examen et les examens et échantillons antérieurs ;
- utilisation de questions différentes des exemples donnés dans les ressources publiées.

Cette réunion aboutit à la production d'une autre série de versions préliminaires, qui sont passées en revue soit par le conseiller ou la conseillère externe, soit par la ou le responsable de la normalisation.

Il s'agit d'une autre personne spécialiste de la matière qui n'est pas encore intervenue dans la rédaction de ces examens et barèmes de notation. Elle offre une nouvelle perspective sur ce qui a été produit et doit commenter le niveau de difficulté des examens pris ensemble, leur correspondance avec les examens antérieurs et leur adéquation avec le programme d'études. Elle propose des changements auxquels réfléchissent ensuite l'auteur ou l'autrice et la ou le responsable de matière pour savoir quelle orientation donner aux versions définitives.

Les versions définitives font l'objet d'une révision du contenu en interne : des vérifications spécialisées sont effectuées par un correcteur ou une correctrice et un concepteur ou une conceptrice, avec l'aide éventuelle de la ou du responsable de matière et de l'auteur ou autrice.

La validation du contenu définitif de l'examen revient à l'auteur ou autrice, et la ou le responsable de matière de l'IB doit confirmer que l'examen a été passé en revue dans le respect des normes de qualité de l'IB établies.

Le barème de notation n'est pas encore validé à ce stade, mais il est passé en revue lors de la normalisation, à la lumière des réponses des élèves.

### L'élaboration et la relecture

Une fois le contenu validé, le texte des questions ne peut plus être modifié. Le concepteur ou la conceptrice met alors les questions au format requis pour les examens (papier ou numérique).

Pour les examens traditionnels (sur papier), cela signifie ajouter la page de couverture avec les instructions (ou les consignes), appliquer les règles stylistiques appropriées, et ajouter les différents codes-barres et pages vierges pour constituer le livret de questions. Ce processus implique également de redessiner certains diagrammes pour obtenir le niveau de qualité souhaité ou de se procurer les images requises en ligne.

Pour l'évaluation numérique, le concepteur ou la conceptrice doit créer les examens directement dans l'environnement de développement qui convient. Cela vaut également pour les instructions, les consignes et les règles stylistiques.

Dans les deux cas, un éditeur ou une éditrice responsable de la production compare les évaluations, après édition, au contenu validé.

## La validation de la facilité d'utilisation

Le vérificateur ou la vérificatrice est une personne spécialiste de la matière qui n'est pas encore intervenue dans le processus de production de l'évaluation à ce stade. Elle passe l'examen comme le feraient les élèves, en tenant notamment compte du temps pris pour répondre aux questions. Cette vérification a pour but de déceler toutes les erreurs et de déterminer si les instructions ou les questions comportent des ambiguïtés.

Elle passe également en revue le barème de notation après avoir terminé l'évaluation.

En cas de cours de langue présenté par un petit nombre d'élèves, ce rôle de vérification est assuré par une seconde personne parlant couramment la langue plutôt que par un ou une spécialiste littéraire.

Après examen des commentaires formels du vérificateur ou de la vérificatrice, l'évaluation est finalisée. Elle est prête pour l'envoi sécurisé aux établissements scolaires.

## Le contrôle de la qualité

Il est essentiel que les examens produits soient exempts d'erreur et d'ambiguïté, car de tels problèmes pourraient avoir des répercussions considérables sur les élèves qui passent les examens. L'IB prend donc ces questions très au sérieux.

En plus de la validation officielle, un contrôle de la qualité est ainsi effectué à chaque étape du cycle de production de l'évaluation.

## Les traductions

Pour les examens autres que ceux de langue, l'IB offre la possibilité aux élèves de faire leur choix parmi plusieurs langues d'usage. À l'heure actuelle, la plupart des matières sont disponibles en anglais, en espagnol et en français. Des accords ont également été passés localement pour que certaines matières soient proposées en allemand, en coréen et en japonais.

La nécessité de traduire les examens est prise en compte dès les premières versions préliminaires, et les équipes de rédaction et de correction de l'IB sont parfaitement informées des problèmes que cela peut poser. La traduction officielle commence dès que les examens en anglais ont été mis au point.

Il est fondamental que le processus de traduction n'altère en rien le sens des questions, pour que le fait de passer l'évaluation dans une langue ou dans une autre ne procure aucun avantage ni aucun désavantage aux élèves. L'IB engage des traducteurs et traductrices externes qui sont des spécialistes de la matière afin de garantir que le langage technique utilisé dans l'évaluation reste correct. La traduction est ensuite comparée à l'examen original (et pas uniquement le texte) par un réviseur ou une réviseuse externe bilingue. Il s'agit du dernier contrôle de la qualité.

## Les épreuves modifiées

L'idéal est de concevoir des évaluations adaptées à l'ensemble des élèves. Certaines demandes et exigences ne peuvent cependant pas être satisfaites par la simple adoption de la conception universelle. Le cas échéant, il est nécessaire de fournir des versions modifiées. Pour ce faire, la *Politique d'accès et d'inclusion* est appliquée rigoureusement, et les épreuves sont produites et envoyées aux établissements scolaires d'une manière totalement similaire aux évaluations non modifiées.

## Les examens

- Les examens ont pour but de permettre aux élèves de montrer leurs connaissances, leur compréhension et l'application de leurs compétences dans un cadre bien délimité présentant une expérience cohérente à l'ensemble des élèves et réduisant les risques de mauvaise conduite.
- Lorsqu'il prépare une session d'examens, l'IB doit trouver un équilibre entre les besoins de l'ensemble des élèves et la facilité d'application pour les établissements, en procédant à la notation le plus rapidement possible pour combler les attentes des élèves.

- L'IB publie des règles claires concernant le comportement attendu dans les lieux d'examen, afin de réduire les possibilités de mauvaise conduite. Ces règles sont régulièrement mises à jour pour suivre les évolutions technologiques et les changements d'environnement subséquents.
- Le calendrier des examens est un compromis entre de nombreuses priorités contradictoires et il constitue l'option « la moins pire » au niveau international.
- Les *Procédures d'évaluation* spécifiques à chaque programme contiennent tous les renseignements relatifs aux événements imprévus, aux aménagements à des fins d'accès et d'inclusion, aux circonstances défavorables et à la modification du calendrier.

Pour les élèves, les examens sont généralement synonymes de période stressante et d'attentes multiples. Afin de réduire la pression exercée sur les élèves, l'IB invite à respecter les principes de base suivants :

- limiter la durée des sessions d'examens ;
- éviter, dans la mesure du possible, les chevauchements avec d'autres examens (par exemple, des tests nationaux).

Il faut cependant trouver un juste équilibre entre ces éléments et la nécessité d'avoir une évaluation suffisante pour tirer des conclusions valides sur l'accomplissement des élèves, tout en empêchant les recours à la mauvaise conduite.

## La préparation et la gestion du lieu d'examen

Les établissements scolaires doivent organiser les examens en respectant le règlement strict présenté dans le document *Déroulement des examens* spécifique à chaque programme.

La conservation des examens en lieu sûr et le bon déroulement de chaque examen sont des conditions impératives, car si les établissements scolaires ou les parties prenantes ont le sentiment qu'il y a eu mauvaise conduite, alors la valeur des résultats des élèves s'en trouvera considérablement diminuée. Afin de maintenir un climat de confiance global, l'IB suit un certain nombre de politiques et de procédures qui sont expliquées dans la section « *L'intégrité de l'évaluation* ».

L'IB a indéniablement à gérer certains incidents faisant qu'un travail ne peut lui être remis en vue d'une notation, pour des raisons indépendantes de la volonté de l'établissement scolaire et de l'élève. L'IB tente alors de limiter les dégâts en estimant une note pour l'élève, ce qui n'est envisageable que s'il existe d'autres preuves sur lesquelles s'appuyer. Et pour s'en assurer avec les examens sur papier, l'IB demande à ce que les différentes réponses (copies d'examen) ne soient pas toutes envoyées au centre de numérisation le même jour, afin de réduire la possibilité de perdre toutes les copies durant leur acheminement.

Dans le cadre du Programme du diplôme et du POP, ce risque sera encore réduit avec l'évaluation numérique, puisqu'une copie du travail envoyé pour l'évaluation sera mise à disposition à chaque étape du processus de chargement.

## L'établissement du calendrier des examens

Parce que l'IB souhaite que la session soit la plus courte possible, plusieurs matières sont prévues au même moment dans le calendrier des examens. Même si l'IB met tout en œuvre pour tenir compte des différentes combinaisons de matières possibles pour les élèves, force est de reconnaître qu'il ne peut empêcher qu'un petit nombre de chevauchements se produisent, et la procédure à suivre dans de tels cas est indiquée dans les *Procédures d'évaluation* spécifiques à chaque programme.

Le calendrier des examens est établi de manière à éviter que les élèves aient plus de six heures et demie d'examens dans la même journée. Généralement, les deux ou trois examens d'une matière aux composantes multiples ont lieu les uns à la suite des autres, avec le premier examen commençant dans l'après-midi et le dernier se terminant le lendemain matin. Bien qu'elle ne puisse pas toujours être mise en place, cette option est préférable à l'organisation de tous les examens d'une même matière le même jour.

Pour la plupart des élèves, cela permet d'échelonner les examens de manière plus équilibrée sur toute la durée du calendrier des examens, à savoir deux semaines pour le PEI et trois semaines pour le Programme

du diplôme ou le POP. Les élèves ayant le sentiment de ne pas avoir montré leur juste valeur lors d'un examen ont ainsi l'occasion de se reprendre ultérieurement.

## Les principes de conception du calendrier des examens

Les points suivants donnent un aperçu général des principes qui sous-tendent l'établissement du calendrier.

Il n'est pas toujours possible de suivre tous ces principes. Le cas échéant, un compromis doit être trouvé. L'IB publie le calendrier des examens au moins un an avant leur tenue.

- En raison du grand nombre de pays et de régions (ou territoires) dans lesquels les programmes de l'IB sont proposés, il n'est pas possible de tenir compte des jours fériés et des fêtes nationales ou locales (à l'exception du 1er mai).
- Il serait souhaitable de ne pas organiser d'examen le jeudi et le vendredi, par égard pour les établissements scolaires du Moyen-Orient où ces jours sont instaurés comme jours de repos, mais il est actuellement impossible d'agir autrement, et ce, pour des questions pratiques.
- L'IB met tout en œuvre pour tenir compte des liens particulièrement étroits qui peuvent exister entre certaines matières et une région ou une culture en particulier. Par exemple, les examens de littérature ou de langue arabe ne sont pas prévus un vendredi.
- L'IB se sert des données d'inscription relatives aux combinaisons de matières pour s'assurer que les chevauchements touchent un minimum d'élèves à travers le monde.
- L'interrogation ne peut pas porter sur deux langues étrangères différentes le même jour (cette mesure ne concerne pas les langues d'usage).
- Afin de réduire les risques qu'un événement inattendu désavantage des élèves dans toutes les composantes des matières qui comptent plusieurs examens, ces derniers sont prévus sur deux jours ou plus.
- Les élèves doivent pouvoir concentrer leurs révisions dans chaque matière dans un espace de temps limité. De ce fait, les examens qui s'étalent sur plus d'une journée sont prévus sur des jours consécutifs, dans la mesure du possible.
- Les examens des élèves doivent être répartis sur toute la période prévue et non être concentrés sur une période plus courte. Cela signifie que, dans la mesure du possible, les matières qui comptent le plus grand nombre d'élèves (par exemple, l'anglais, l'histoire et les mathématiques) ne sont pas prévues sur des jours consécutifs.
- Pour la même raison, l'IB s'efforce de prévoir les examens de langue et de sciences dans chacune des deux (pour le PEI) ou trois (pour le Programme du diplôme ou le POP) semaines du calendrier des examens.
- Pour pouvoir satisfaire à des exigences de traitement des examens internes à l'IB, certaines matières doivent avoir lieu au début du calendrier. C'est notamment le cas des matières qui comptent un grand nombre d'élèves.

## La gestion des événements imprévus

Les *Procédures d'évaluation* spécifiques à chaque programme expliquent comment gérer la plupart des événements imprévus. Si vous avez le moindre doute à ce sujet, n'hésitez pas à contacter le service [L'IB vous répond](#) qui saura vous apporter tous les renseignements nécessaires.

## Les aménagements à des fins d'accès et d'inclusion et les circonstances défavorables

Ces deux points sont largement traités dans la section « [Le caractère prioritaire de l'équité universelle](#) ».

L'IB estime que les élèves dans leur intégralité doivent pouvoir démontrer leurs capacités dans des conditions d'évaluation aussi équitables que possible. Les conditions ordinaires d'évaluation peuvent

désavantager les élèves ayant des besoins en matière de soutien à l'apprentissage en les empêchant de montrer le niveau atteint. Le cas échéant, des aménagements à des fins d'accès et d'inclusion peuvent être autorisés. Il peut s'agir de dispositions spéciales pour passer les examens ou de modifications apportées aux examens.

Les événements échappant au contrôle de l'élève, qui se produisent au cours de la session d'examens et qui risquent d'avoir des conséquences négatives sur son travail, portent le nom de « circonstances défavorables ». Ces circonstances incluent les maladies ou blessures temporaires, le stress grave, les circonstances familiales particulièrement éprouvantes, le deuil ou les événements pouvant menacer la santé ou la sécurité de l'élève. Les circonstances défavorables peuvent également inclure des événements qui affectent la communauté scolaire dans son ensemble, comme des troubles civils ou une catastrophe naturelle. Elles n'incluent pas les insuffisances dues au fait de l'établissement scolaire.

Pour obtenir davantage d'informations sur ces aménagements, veuillez consulter les *Procédures d'évaluation* spécifiques à chaque programme.

## La modification du calendrier des examens

Modifier la date d'un examen présente des risques importants pour l'intégrité du déroulement de l'examen en question. Cela implique en effet que des élèves passeront l'examen avant ou après d'autres, une situation qui peut facilement conduire à une mauvaise conduite. Il ne sera jamais accepté qu'un examen soit avancé à la veille du jour initialement prévu, car le risque d'infraction au règlement pour la majorité des élèves passant l'examen serait bien trop élevé.

L'IB peut autoriser qu'un ou plusieurs examens soient passés à une heure ou à une date différentes de celles indiquées dans le calendrier des examens de l'IB dans des circonstances bien particulières. Ces situations ainsi que la marche à suivre pour demander la modification du calendrier des examens sont détaillées dans les *Procédures d'évaluation* spécifiques à chaque programme et dans la *Politique de préparation aux examens*, qui sont disponibles sur le [Centre de ressources pédagogiques](#).

## La notation

- Pour être réussie, la notation doit permettre aux élèves d'obtenir un résultat cohérent et exact reflétant la qualité de leur travail.
- Les opinions exprimées par l'examineur principal ou examinatrice principale sont correctes et doivent être reproduites par le reste des examinateurs et examinatrices dans leur notation.
- Notes et notes finales sont deux choses différentes : les élèves peuvent obtenir une meilleure note à un examen plus facile, mais doivent se voir attribuer la même note finale.

## Les notes et les notes finales

Pour bien conduire des évaluations sommatives et en tirer profit, il est important de comprendre la différence entre la notation du travail des élèves et l'attribution d'une note finale.

- Dans la notation, le travail produit par l'élève est récompensé au regard d'un barème de notation ou d'un autre cadre similaire. Il s'agit d'une indication de la proportion de la tâche d'évaluation que l'élève a réussie. La note en tant que telle n'a pas d'autre portée.
- En attribuant une note finale, l'examineur ou examinatrice émet un jugement sur la qualité du travail de l'élève au regard d'une norme définie et en prenant en compte la difficulté de la tâche ainsi que la proportion de la tâche qui a été réalisée. Par conséquent, la note finale a un sens ou une pertinence et doit en principe pouvoir être comparée aux résultats obtenus dans les autres évaluations.

Il serait possible d'obtenir une note finale élevée en réussissant uniquement une proportion limitée d'une question très difficile, comme il pourrait être impossible de réussir à obtenir cette même note finale en répondant correctement à de nombreuses questions bien moins complexes.

Tel qu'expliqué par la suite, il n'est pas nécessaire de faire des références aux aspects réussis par l'élève pour expliquer le niveau correspondant à la note finale, même si l'IB suit une telle approche. En effet, il existe d'autres systèmes tout à fait cohérents et très réputés, dans lesquels le niveau de l'élève est déterminé par rapport au résultat de ses pairs.

Dans ses évaluations, l'IB se sert habituellement des notes pour avoir une idée des résultats globaux. Il s'intéresse ensuite à ce qui est exprimé par les élèves ayant obtenu ce nombre de points pour fixer une limite (seuil d'attribution des notes finales) au-delà de laquelle attribuer une certaine note finale aux élèves ayant obtenu plus de points. Ce processus est expliqué plus en détail dans la section « [L'attribution des notes finales et la totalisation](#) ».

## Les approches de la notation

L'IB comptant une multitude d'outils d'évaluation, il est crucial de choisir la bonne approche de la notation pour chaque évaluation. Cette approche est déterminée lors du processus de développement du programme d'études.

### Les barèmes de notation analytiques

Les barèmes de notation analytiques sont utilisés lorsque la gamme de réponses attendues des élèves est étroite.

Ces barèmes peuvent donner aux examinateurs et examinatrices des instructions spécifiques sur la manière de répartir le total des points prévus pour une question entre les différentes parties de la réponse. Même avec les questions structurées qui appellent des réponses hautement spécifiques, les barèmes de notation doivent fournir aux examinateurs et examinatrices des informations suffisantes pour leur permettre de

noter de manière cohérente, en tant compte des différents types d’approches que les élèves pourraient adopter et des erreurs courantes qui pourraient être commises. Les examinateurs et examinatrices devront inévitablement faire appel à leur jugement professionnel pour attribuer des points aux réponses inattendues ou aux autres réponses valides, mais les barèmes de notation devront leur apporter la meilleure orientation possible pour les aider à émettre un tel jugement.

Il est rarement possible de prévoir quelles parties d’une question les élèves réussiront ou non. Cette incertitude peut se révéler particulièrement problématique avec les questions structurées à réponse développée, dans lesquelles une erreur commise dans une des premières parties peut avoir des répercussions par la suite. L’IB conçoit ces questions de sorte que si l’élève commet une erreur dans la première partie de la question, cela ne l’empêche pas de poursuivre la question. Les barèmes de notation analytiques doivent malgré tout apporter des orientations explicites quant à la manière de noter certains types de réponses incorrectes et de gérer le travail de l’élève lorsqu’une erreur a été commise dans une partie d’une question.

## Les critères d’évaluation

Lorsqu’un barème de notation détaillé ne s’avère pas approprié aux évaluations, des critères d’évaluation peuvent être utilisés à la place.

Les critères d’évaluation portent davantage sur le type d’accomplissement attendu des élèves, indépendamment des renseignements apportés dans chaque réponse. Les différents niveaux de compétence possibles sont illustrés par les descripteurs de niveaux.

L’IB suit une approche dite « de meilleur ajustement » pour appliquer les descripteurs de niveaux de chaque critère. L’examinateur ou examinatrice qui utilise le critère d’évaluation doit choisir le niveau de réussite qui correspond le mieux au travail évalué. Il n’est pas nécessaire que chaque aspect du niveau de réussite soit satisfait dans le détail pour attribuer ce niveau. Il faut souligner que le plus haut niveau d’un critère, quel qu’il soit, ne décrit pas un travail parfait.

Dans la plupart des composantes notées selon des critères d’évaluation, il est plus courant d’utiliser de multiples critères qu’un seul. Dans de telles situations, il est important que les critères soient indépendants les uns des autres. En effet, il ne serait pas juste que l’élève reçoive des points à plusieurs reprises pour le même aspect de sa réponse.

En règle générale, les mêmes critères d’évaluation sont réutilisés chaque année, même si les questions posées aux élèves sont différentes. Cela s’explique par le fait que ce qui est évalué reste de même nature. Toute nouvelle question porte généralement sur des détails spécifiques qui ne figurent pas explicitement dans les critères. L’examinateur principal ou examinatrice principale préparera souvent des remarques sur la notation pour expliquer aux examinateurs et examinatrices comment appliquer les critères d’évaluation à chaque question, et inclura des exemples de détails pertinents.

## Les critères globaux : les bandes de notation

Il est parfois inapproprié de séparer les différents critères d’évaluation pour noter un travail. Cela se produit généralement lorsqu’il est impossible d’avoir des critères distincts, indépendants les uns des autres. Le cas échéant, des bandes de notation viennent remplacer les critères distincts. De fait, les bandes de notation constituent ensemble un critère global appliqué au travail, qui est jugé comme un tout. Étant donné la nécessité de disposer d’une gamme de notes raisonnable pour pouvoir différencier les accomplissements des élèves, plusieurs notes correspondent à chaque descripteur de niveau.

Tableau 7

Exemple de bandes de notation de l’IB

| Nombre de points | Descripteurs de niveaux  |
|------------------|--|
| 0                | Le travail n’atteint aucun des niveaux décrits ci-dessous.   |
| 1 – 3            | <ul style="list-style-type: none"> <li>La réponse montre une compréhension limitée de la question. Le vocabulaire spécifique à la matière n’est pas utilisé ou il est systématiquement utilisé de façon inappropriée.</li> </ul> |

| Nombre de points | Descripteurs de niveaux   |
|------------------|---|
|                  | <ul style="list-style-type: none"> <li>La réponse explique très peu la question. Les idées sont superficielles et manquent souvent de clarté.</li> <li>La réponse est descriptive. Toute analyse présente est superficielle ou incohérente. Les différents points de vue sont peu voire pas du tout discutés. Lorsqu'elle est présente, la conclusion est très superficielle ou n'est pas cohérente avec le reste de la réponse.</li> </ul>   |
| 4 – 6            | <ul style="list-style-type: none"> <li>La réponse montre une compréhension élémentaire de la question. Le vocabulaire spécifique à la matière est utilisé, mais souvent de façon inappropriée.</li> <li>L'explication de la question est sommaire et insuffisamment développée. Les idées sont souvent peu précises ou vagues, et il est souvent difficile de comprendre ce que la réponse essaie d'exprimer.</li> <li>L'analyse est limitée et la réponse est dans l'ensemble plus descriptive qu'analytique. Les différents points de vue sont discutés, mais de façon limitée. Une conclusion simpliste est fournie.</li> </ul>        |
| 7 – 9            | <ul style="list-style-type: none"> <li>La réponse montre une certaine compréhension de la question. Le vocabulaire spécifique à la matière est utilisé, parfois de façon appropriée.</li> <li>La réponse explique de façon satisfaisante la question, mais l'explication n'est parfois pas assez claire et développée. Des idées pertinentes sont présentées, mais elles ne sont pas assez précises et détaillées.</li> <li>La réponse présente une analyse, mais cette dernière n'est pas assez développée. Des points de vue différents sont quelque peu discutés. Une conclusion est fournie.</li> </ul>                               |
| 10 – 12          | <ul style="list-style-type: none"> <li>La réponse montre une bonne compréhension de la question. Le vocabulaire spécifique à la matière est utilisé, la plupart du temps de façon appropriée.</li> <li>La réponse explique clairement la question, mais l'explication n'est pas assez développée. Les idées présentées sont pertinentes et correctes, mais elles ne sont pas assez détaillées.</li> <li>La réponse présente une analyse critique, mais cette dernière n'est pas assez développée. Différents points de vue sont discutés. La réponse mène à une conclusion qui est en adéquation avec les arguments présentés.</li> </ul> |
| 13 – 15          | <ul style="list-style-type: none"> <li>La réponse montre une très bonne compréhension de la question. Le vocabulaire spécifique à la matière est utilisé avec exactitude et précision.</li> <li>L'explication de la question est claire et suffisamment développée. Les idées présentées sont pertinentes, correctes et détaillées.</li> <li>La réponse présente une analyse critique bien développée. Différents points de vue sont discutés de façon critique. La réponse mène à une conclusion raisonnée et clairement indiquée, qui est en adéquation avec les arguments présentés.</li> </ul>  |

Les descripteurs eux-mêmes ont tendance à être assez longs puisqu'ils couvrent tout un éventail de qualités pouvant être présentes dans les travaux des élèves, et ils ont un lien direct avec les objectifs du cours. Tout comme pour les critères d'évaluation, une approche de meilleur ajustement est utilisée. Les examinateurs et examinatrices doivent émettre un jugement en choisissant une note dans la gamme de notes disponibles dans chaque descripteur de niveau. Pour ce faire, il leur faut déterminer à quel point le travail de l'élève correspond à ce descripteur. Par exemple, un descripteur de niveau peut couvrir les notes allant de 7 à 9. L'examinateur ou examinatrice attribuera une note comprise dans cet intervalle en fonction de la mesure dans laquelle le travail de l'élève correspond à ce descripteur de la bande de notation.

## L'autonotation

Une autonotation s'applique à quelques composantes, à savoir celles qui comportent des questions à choix multiple. Ce processus consiste à utiliser la technologie pour évaluer le travail des élèves à l'aide d'un

barème de notation prédéfini. Pour ces questions à choix multiple, le barème de notation détermine objectivement si une réponse est correcte ou incorrecte. Les barèmes de notation de ces composantes sont extrêmement clairs et se passent de tout jugement.

## La notation nécessitant un jugement

Si l'autonotation est parfois possible, les évaluations de l'IB ont majoritairement besoin d'être notées par des examinateurs ou examinatrices, étant donné que des barèmes de notation analytiques ou des critères de notation doivent être appliqués.

Les examinateurs et examinatrices notent les travaux des élèves de manière électronique, en se servant d'un logiciel connu sous le nom d'« outil de notation ». Ce logiciel leur permet de cocher certains éléments, de consigner les points attribués et d'insérer des remarques sur la version électronique du travail de l'élève, qui peut inclure des éléments audio et visuels. Cela élargit considérablement l'éventail des travaux d'élèves qui peuvent être envoyés comme preuves (par exemple, des enregistrements des représentations ou des épreuves orales).

Pour obtenir de plus amples informations sur le système de notation de l'IB, veuillez vous référer à la vidéo intitulée « [Les deux méthodes de notation : comment la notation est-elle effectuée par les examinateurs, les examinatrices et le personnel enseignant ?](#) ».

## La normalisation

Différents aspects de la procédure d'évaluation se retrouvent dans le processus de normalisation. Son objectif est de préparer les examinateurs et examinatrices à noter en toute fiabilité.

Pour obtenir des résultats fiables, il est important que la note de chaque copie d'examen soit déterminée en fonction de la qualité du travail qu'elle contient et non en fonction de la personne qui corrige le travail. Pour ce faire, les examinateurs et examinatrices doivent montrer leur compréhension de la norme de notation établie par l'examineur principal ou examinatrice principale pour la composante concernée, et s'y conformer, avant de pouvoir procéder à la notation lors d'une session d'examens de l'IB en conditions réelles.

Le processus de normalisation commence par le recrutement d'examineurs et examinatrices possédant l'expérience et les compétences nécessaires. Une formation leur est ensuite fournie pour leur expliquer les principes et les pratiques à adopter en matière de notation.

Après la tenue d'un examen, l'équipe d'examineurs superviseurs et examinatrices superviseuses se réunit, sous la direction de l'examineur principal ou examinatrice principale, pour discuter du barème de notation et passer en revue des exemples de copies. Cette réunion de normalisation vise à vérifier que le barème de notation couvre convenablement la gamme de réponses que les examinateurs et examinatrices devraient avoir à corriger. Un barème de notation ne pourra sans doute jamais couvrir toutes les réponses d'élèves possibles. L'analyse d'exemples de copies effectuée lors de la réunion permet toutefois à l'examineur principal ou examinatrice principale et à l'équipe d'examineurs superviseurs et examinatrices superviseuses de formuler des instructions et des conseils clairs pour préparer les examinateurs et examinatrices au type de réponse attendue pour chaque note, sur la base du travail consulté.

La réunion de normalisation aboutit à une série de copies comportant la notation définitive de l'examineur principal ou examinatrice principale. Ces copies sont utilisées pour former les examinateurs et examinatrices à noter les travaux en respectant la norme appliquée par l'examineur principal ou examinatrice principale. Seules les personnes qui sont parvenues à se conformer à la norme de notation de l'examineur principal ou examinatrice principale sont autorisées à noter des copies pour une session d'examens en conditions réelles, leur norme de notation étant régulièrement vérifiée au cours de la période de notation à l'aide de procédures rigoureuses d'assurance de la qualité.

## Le modèle d'assurance de la qualité

Il est essentiel que, quelle que soit la personne qui corrige leur travail, les élèves obtiennent le même résultat que si leur travail avait été corrigé par l'examineur principal ou examinatrice principale.

Après la tenue de l'évaluation et la réception d'un certain nombre de réponses d'élèves, l'examineur principal ou examinatrice principale tient une réunion de normalisation avec les autres examinateurs superviseurs et examinatrices superviseuses. Il s'agit alors de s'accorder sur les notes à attribuer à un certain nombre de réponses qui seront utilisées pour créer le modèle d'assurance de la qualité. Les copies comportant les notes définitives sont classées selon trois catégories : réponses d'entraînement, de validation et de contrôle. Ensemble, elles constituent le modèle d'assurance de la qualité.

### L'entraînement

L'objectif des réponses d'entraînement est de permettre aux examinateurs et examinatrices de se familiariser avec la norme de notation établie par l'examineur principal ou examinatrice principale. Ces copies doivent montrer comment noter des exemples types et relever toute situation courante requérant une explication particulière du barème de notation.

Pour être de qualité, une série de réponses d'entraînement doit être très diversifiée et comporter des notes allant de faibles à élevées, pour aider les examinateurs et examinatrices à reconnaître où il est possible d'attribuer des points et comment distinguer les bonnes (et les mauvaises) réponses.

Après avoir passé en revue le barème de notation ou les critères de notation, et les éventuelles remarques sur la notation fournies par l'examineur principal ou examinatrice principale après la réunion de normalisation, les examinateurs et examinatrices peuvent se lancer dans la phase d'entraînement. Les réponses d'entraînement montrent la norme de notation appliquée par l'examineur principal ou examinatrice principale. Les examinateurs et examinatrices ont la possibilité de demander des conseils à leur chef d'équipe pendant la notation des réponses d'entraînement.

Une fois ce travail terminé, les examinateurs et examinatrices doivent avoir la certitude d'avoir compris la norme de notation de l'examineur principal ou examinatrice principale et de pouvoir l'appliquer, à la suite de quoi s'engagera le processus de validation.

### La validation

L'objectif des réponses de validation est de donner l'occasion aux examinateurs et examinatrices de montrer leur capacité à noter les travaux conformément à la norme de notation. Après avoir terminé avec succès la validation, les examinateurs et examinatrices seront en mesure de procéder à la notation lors d'une session d'examens en conditions réelles.

Chaque série comporte généralement cinq réponses de validation. Si l'examineur ou examinatrice n'attribue pas les mêmes notes que l'examineur principal ou examinatrice principale pendant la phase de validation, un retour d'information lui est fait pour lui expliquer quelles notes étaient attendues et pour quelles raisons. Une seconde possibilité de montrer sa compréhension de la norme de notation lui est alors donnée, à partir d'une autre série de réponses de validation.

Pour être de qualité, une série de réponses de validation doit être très diversifiée et comporter des notes allant de faibles à élevées. Elle doit également comporter des exemples exigeant que l'examineur ou examinatrice ait compris des indications précises dans le barème de notation pour l'aider à gérer des difficultés associées à la notation qui auront été anticipées par l'examineur principal ou examinatrice principale et l'équipe d'examineurs superviseurs et examinatrices superviseuses. Il s'agira des situations courantes révélées par l'analyse des copies réalisée lors de la réunion de normalisation. Les réponses de validation ne sont pas choisies dans l'objectif de piéger les examinateurs et examinatrices, mais de représenter les types de réponses qui ont des chances d'apparaître lors de la notation.

À cette étape, deux occasions sont offertes aux examinateurs et examinatrices de montrer leur compréhension de la norme de notation de l'examineur principal ou examinatrice principale et leur capacité à respecter cette norme pour les travaux envoyés pour l'évaluation. Les examinateurs et examinatrices doivent réussir à noter correctement une série de validation.

Chaque série de validation est constituée d'un certain nombre de copies comportant les notes définitives. Pendant la correction d'une série de validation, il est interdit de discuter des copies avec la ou le chef d'équipe. Si la marge de tolérance est respectée lors de la première notation des réponses de validation, alors les examinateurs et examinatrices peuvent passer à la notation en conditions réelles.

Si cette norme n'est pas respectée en revanche, la ou le chef d'équipe fournit un retour d'information aux examinateurs et examinatrices pour leur expliquer comment mettre leur notation en conformité avec la norme de l'examinateur principal ou examinatrice principale. Après avoir réfléchi sur ces éléments, les examinateurs et examinatrices se voient offrir une seconde possibilité de montrer leur compréhension de la norme de notation.

Si la marge de tolérance est respectée lors de la deuxième notation des réponses de validation, alors les examinateurs et examinatrices ont l'autorisation de procéder à la notation en conditions réelles.

### **La notation en conditions réelles et la notation des réponses de contrôle**

Les réponses de contrôle sont une constituante du processus d'assurance de la qualité. Elles servent à contrôler le travail des examinateurs et examinatrices, et à prouver que leur notation reste conforme à la norme attendue.

Au fil du temps, il est possible qu'un écart commence à se creuser avec la norme établie par l'examinateur principal ou examinatrice principale. Par conséquent, l'IB vérifie régulièrement la notation à l'aide de réponses de contrôle, qui comportent la notation définitive de l'examinateur principal ou examinatrice principale. Les réponses de contrôle ressemblent à toutes les autres copies. L'examinateur ou examinatrice n'a donc pas conscience de corriger une copie à des fins de contrôle de la qualité. En règle générale, une réponse de contrôle est glissée toutes les dix copies.

Si la note donnée à cette réponse de contrôle équivaut à celle de l'examinateur principal ou examinatrice principale, le travail de notation peut être poursuivi. En revanche, s'il s'avère que la norme de notation attendue n'est plus respectée car une note incorrecte est attribuée, l'IB se doit d'intervenir.

Cela commence par un retour d'information et des orientations provenant des examinateurs superviseurs et examinatrices superviseuses, afin de laisser la possibilité à l'examinateur ou examinatrice de se remettre en conformité avec la norme de notation. Si le réajustement n'est pas au rendez-vous, l'examinateur ou examinatrice ne pourra pas noter d'autres copies.

Des marges de tolérance sont définies pour chaque composante, pour autoriser un certain écart par rapport à la norme définitive. Pour en savoir plus, veuillez vous référer à la section intitulée « [Les marges de tolérance](#) ».

## Les indices d'une normalisation réussie

Lorsqu'un exercice de normalisation est réussi, les éléments suivants se retrouvent :

- des équipes d'examineurs superviseurs et examinatrices superviseuses faisant preuve d'implication et d'efficacité pour aider les examinateurs et examinatrices ayant des questions liées à la notation et pour garantir la cohérence de la notation ;
- des instructions claires et univoques, prenant la forme de barèmes de notation et de remarques sur la notation, expliquant clairement les motifs de l'attribution des notes ;
- des réponses d'entraînement, de validation et de contrôle notées de manière définitive pouvant servir à des fins d'entraînement et d'assurance de la qualité.

## Les marges de tolérance

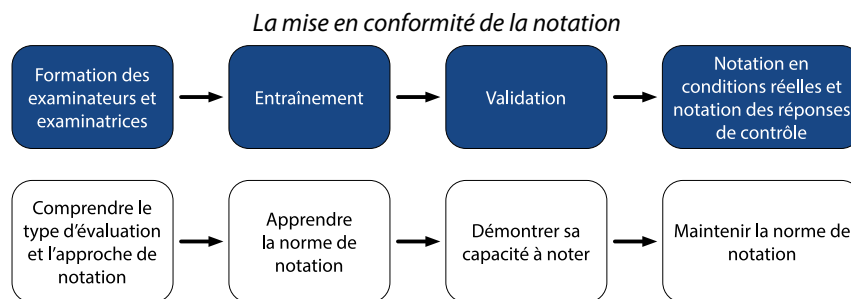
L'IB tient à poser des questions qui évaluent ce qui est important, et pas simplement ce qui est facile à corriger. Pour les questions appelant une réponse qui est soit bonne soit mauvaise, il est raisonnable de s'attendre à ce que l'examinateur ou examinatrice attribue exactement la même note que l'examinateur principal ou examinatrice principale. Pour les autres questions, notamment celles qui évaluent la

compréhension ou l'analyse, un certain degré de jugement intervient. Cela signifie concrètement que deux personnes ayant la même expertise et comprenant de la même façon comment bien répondre correctement à une question peuvent ne pas attribuer exactement la même note au travail corrigé, à un ou deux points près.

L'IB s'appuie sur l'idée de marge de tolérance pour rendre compte de ces différences d'opinion légitimes lors de l'examen du travail de notation. Par exemple, si l'examinateur principal ou examinatrice principale a attribué une note de 46 à une dissertation et qu'une marge de tolérance de 2 points était associée à la question, alors l'IB considère que les examinateurs et examinatrices attribuant une note comprise entre 44 et 48 respectent la norme de notation.

L'IB détermine la marge acceptable entre la note définitive (la note accordée par l'examinateur principal ou examinatrice principale) et la note accordée par les examinateurs et examinatrices à ces copies. Cette marge acceptable porte le nom de « marge de tolérance ». Pour les questions en plusieurs parties ou notées à l'aide de plusieurs critères, l'IB surveille les différences au niveau de la note globale, mais aussi au niveau de chaque partie ou critère. Pour montrer son respect de la norme de notation, un examinateur ou une examinatrice doit se trouver dans la marge de tolérance pour ces deux aspects.

Figure 27



## La formation des examinateurs et examinatrices

Avant le début de toute session d'évaluation, une formation est proposée aux examinateurs et examinatrices. L'objectif est de les aider à comprendre le type d'évaluation et l'approche générale à adopter pour la notation. La formation ne comporte aucun détail sur les examens qui seront prévus aux prochaines sessions.

## Les copies difficiles ou inhabituelles

En cas de copies difficiles à noter, les examinateurs et examinatrices doivent demander conseil à leur chef d'équipe. Si ces conseils ne sont pas suffisants pour garantir une notation équitable, il est possible de demander à une personne placée plus haut dans la hiérarchie de noter ces copies. Les copies particulièrement problématiques seront alors portées à l'attention de l'examinateur principal ou examinatrice principale, qui prendra une décision définitive quant à la manière de les noter.

De la même manière, toutes les copies inhabituelles, ou les copies des élèves ayant répondu à des épreuves d'examen modifiées, seront corrigées par l'examinateur principal ou examinatrice principale. Cette personne a en effet pour responsabilité de veiller de près à ce que la norme de notation pour ces copies soit équivalente à celle appliquée pour toutes les autres.

## Les liens avec un établissement scolaire

L'IB part du principe que les examinateurs et examinatrices ne doivent pas noter le travail de leurs élèves.

À cet effet, l'IB impose à l'ensemble des examinateurs et examinatrices de l'informer de tout lien avec des élèves et des établissements scolaires, afin de ne pas leur confier des travaux d'élèves qui induiraient des conflits d'intérêts. Pour de plus amples informations sur ce processus, veuillez consulter la section « [Les conflits d'intérêts](#) ».

## Les remarques et les annotations des examinateurs et examinatrices

L'évaluation sommative de l'IB a pour objectif de mesurer les accomplissements des élèves. Les examinateurs et examinatrices doivent se concentrer uniquement sur la notation des travaux des élèves conformément à la norme attendue. L'IB leur demande d'inclure des commentaires seulement si cela les aide dans leur travail de notation.

Pour rédiger des commentaires constructifs à l'intention des élèves ou des enseignantes et enseignants, il faut déterminer la note qui correspond au travail, puis tenter d'expliquer de quelle manière ce travail aurait pu être amélioré. L'IB est conscient que la seconde tâche n'est pas insignifiante, car elle constitue l'une des caractéristiques essentielles d'un bon enseignement. Elle requiert cependant du temps et une réflexion qui éloigneront l'examinateur ou examinatrice de sa mission principale, à savoir corriger le travail selon une norme de notation cohérente. En termes simples, l'IB s'attend à ce que les examinateurs et examinatrices effectuent une seule tâche (notation) en maintenant un niveau élevé de qualité, plutôt que deux tâches (notation et retour d'information) avec un niveau de qualité moindre.

L'examinateur ou examinatrice ne peut émettre un jugement qu'à partir du travail mis à sa disposition, alors que les enseignantes et enseignants avec de l'expérience pourront se baser sur tout un éventail d'informations pour décider du retour d'information à faire à l'élève. Le retour de l'examinateur ou examinatrice sera donc moins complet que celui de l'enseignant ou enseignante.

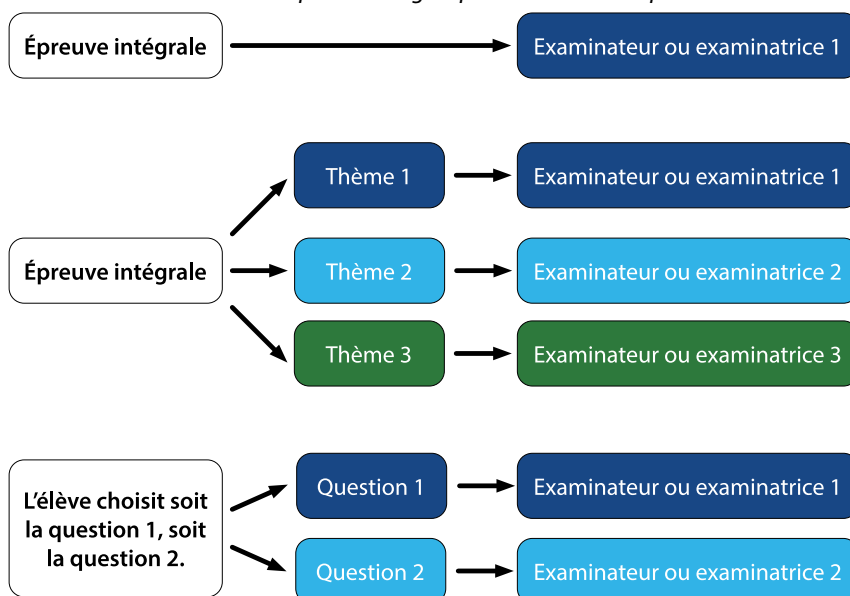
Pour toutes ces raisons, l'IB demande très clairement à ses examinateurs et examinatrices de noter uniquement le travail de l'élève en respectant la norme, et non d'ajouter des commentaires pour faire un retour d'information à l'élève ou à l'enseignant ou enseignante.

Pour faciliter le processus de notation, les examinateurs et examinatrices doivent indiquer clairement à quels endroits les points ont été attribués et, en cas d'ambiguïté, expliquer ce choix en ajoutant des commentaires appropriés. Cela permet à l'IB de vérifier les normes de notation et rend le processus d'attribution des points plus transparent pour les établissements.

## Les groupes d'éléments de question

Figure 28

*La division des épreuves en groupes d'éléments de question*



La création des groupes d'éléments de question repose sur l'idée qu'il est plus facile de corriger plusieurs fois la même question que de corriger toutes les questions d'une copie avant de recommencer avec la première question pour la copie suivante.

Pour chaque groupe d'éléments de question, il existera des réponses d'entraînement, de validation et de contrôle particulières, et les examinateurs et examinatrices devront prouver leur capacité à noter conformément à la norme attendue pour chacun. Même si cela peut paraître difficile, cette mesure permet en réalité à une personne n'ayant pas compris la norme pour une question précise d'un examen, de corriger les autres questions, au lieu d'être suspendue de la notation parce qu'elle ne peut pas noter l'examen dans son intégralité conformément à la norme.

## La totalisation

- Le terme « totalisation » désigne le processus d'addition des notes obtenues dans les composantes pour produire un résultat général.
- Pour que chaque composante contribue à la note définitive dans les bonnes proportions (pondération), il peut s'avérer nécessaire d'appliquer un coefficient aux notes des composantes.
- L'IB utilise un « système de compensation » grâce auquel les élèves peuvent compenser de mauvais résultats obtenus dans une question ou une composante par un niveau plus élevé atteint dans une autre.
- La note finale par matière de l'élève est calculée en totalisant les notes obtenues dans chaque composante, et non ses notes finales par composante.

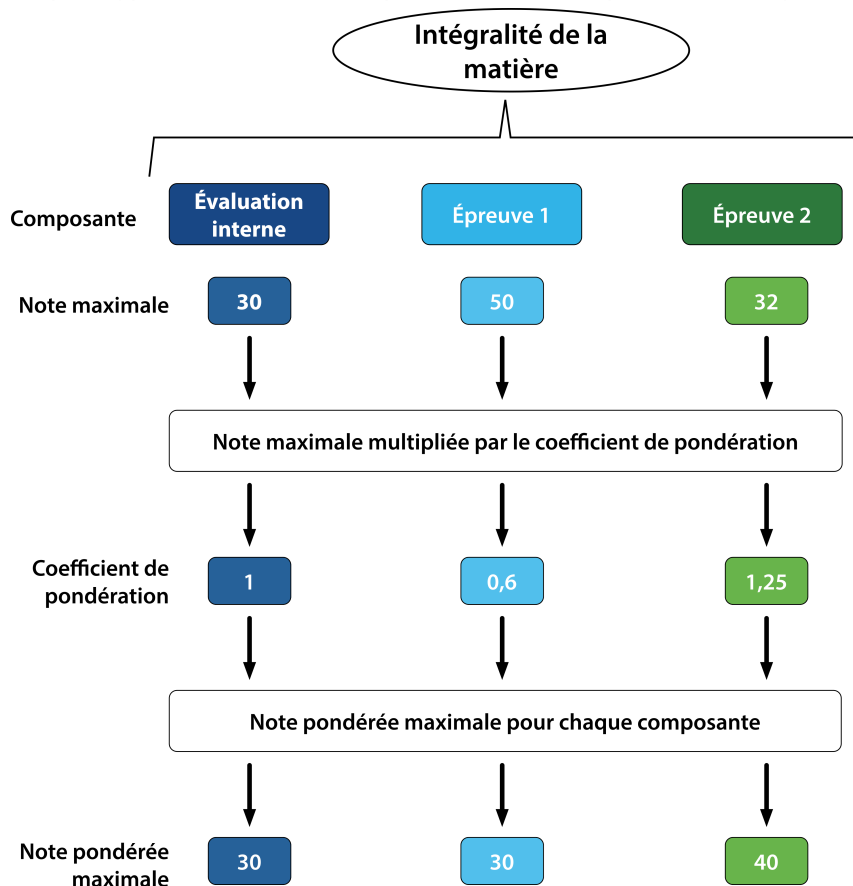
Le terme « totalisation » désigne le processus d'addition des notes (et des seuils) issues des différentes composantes pour produire une note définitive ou un seuil d'attribution des notes finales global. Pour y parvenir, il peut s'avérer nécessaire d'appliquer un coefficient aux notes globales (ou seuils) des composantes.

Un coefficient est appliqué afin de conserver la pondération voulue pour chaque composante pour ce qui est de sa contribution à l'évaluation globale du cours. Cela implique de multiplier ou de diviser les notes des composantes pour qu'elles contribuent comme il se doit au total global de la matière. Cela vaut également pour les seuils d'attribution des notes finales fixés pour la composante, qui auront initialement été déterminés à partir du nombre de points maximum pour les composantes.

La note pondérée maximale pour chaque composante correspond à la pondération indiquée pour la composante en question dans le guide pédagogique.

Figure 29

Exemple d'application d'un coefficient pour une matière comportant trois composantes



Le concept de pondération permet de refléter l'importance relative que l'IB accorde aux éléments évalués, en fonction de la contribution au résultat final qu'il souhaite pour chacun. Ainsi, si une composante sert avant tout à évaluer l'interprétation des données ou des sources, et que sa pondération est de 30 %, cela implique que, en comparaison des autres objectifs spécifiques du cours, l'interprétation représente 30 % de ce qui est important. Il est fréquent que plusieurs composantes évaluent les mêmes objectifs spécifiques, ce qui rend ce calcul moins pertinent.

Un aspect secondaire, mais cependant important, de la pondération, est qu'elle permet à l'IB de déterminer le nombre total de points d'une évaluation qui convient aux tâches et aux critères de notation, plutôt que d'essayer de créer artificiellement un total global. Les totaux peuvent ensuite être additionnés pour parvenir à une note finale.

Dans une évaluation, la totalisation implique de combiner des notes issues de plusieurs composantes pour déterminer la note finale de l'élève. Cette approche est suffisamment souple pour pouvoir fixer les notes en fonction de la difficulté des tâches, des critères et des intentions nourries lors de la conception de l'évaluation, au lieu de suivre un total prédéfini indiqué par les pondérations des composantes. Cela signifie que même si la pondération des composantes est prise en compte, les résultats obtenus peuvent être différents des résultats escomptés en raison, par exemple, de la facilité ou difficulté comparative imprévue d'une évaluation ou d'un élément. L'approche suivie par l'IB n'impose pas d'assurer une cohérence entre chaque note pour suivre rigoureusement les pondérations prédéfinies : ce sont les résultats généraux qui comptent le plus. Cette méthode, connue sous le nom de « système de compensation », permet aux élèves de compenser de mauvais résultats obtenus dans un domaine par de bons résultats dans un autre, en se concentrant sur la note totale plutôt que sur les notes obtenues individuellement aux différentes composantes. La note finale à une matière découle donc de la somme de toutes les notes obtenues aux différentes composantes, sans être la moyenne de ces notes. Cela signifie que des élèves ayant obtenu les mêmes notes finales aux composantes peuvent se voir attribuer des notes finales différentes pour les matières.

## La révision de notation

- La révision de notation désigne le processus qui consiste à vérifier les normes de notation des enseignantes et enseignants, et non à recorriger le travail des élèves.
- Une révision de notation réussie signifie que les élèves obtiendraient la même note d'évaluation interne en étudiant dans un autre établissement situé à l'autre bout du monde. C'est ce que l'IB qualifie de « norme mondiale ».
- Les preuves analysées pendant la révision de notation ne se limitent pas à la qualité du travail des élèves, mais incluent les explications fournies par les enseignantes et enseignants pour justifier leurs notes.

La révision de notation, réalisée à partir d'un échantillonnage dynamique, est utilisée pour l'évaluation interne dans le cadre du Programme du diplôme et du POP, et pour les portfolios électroniques dans le cadre du PEI.

### Définition de la révision de notation

Bien souvent, il est difficile de mesurer la caractéristique souhaitée par l'IB en évaluant l'élève à l'aide d'un examen officiel limité dans le temps. L'approche la plus valide consiste donc généralement à demander à l'enseignant ou enseignante de s'en charger au moyen d'une évaluation interne. La section intitulée « [Le rôle de l'évaluation se déroulant en classe et de l'évaluation interne](#) » explique en quoi une telle approche est adaptée.

Malgré les résultats importants qu'elle produit, cette approche s'accompagne du risque d'une différence d'interprétation de la norme de notation. Ainsi, deux enseignantes ou enseignants dans deux établissements différents pourraient attribuer deux notes différentes à un même travail. Pour que la totalité des examinateurs et examinatrices comprenne de la même façon la norme attendue, l'IB s'appuie sur un processus complet de formation et de test, mais il ne peut pas faire de même avec l'ensemble du personnel enseignant.

Au lieu de cela, il croit en la capacité des enseignantes et enseignants à noter tous les travaux avec cohérence ; c'est pourquoi il n'a plus qu'à appliquer un ajustement à leurs notes, si cela est nécessaire, pour les ramener au niveau de la norme mondiale. Pour ce faire, il demande aux enseignantes et enseignants de fournir un échantillonnage de leur notation, qui sera comparé à la norme de l'examineur principal ou examinatrice principale. À partir des données produites par cette comparaison, l'IB calcule une formule mathématique, si nécessaire, pour ajuster toutes les notes fournies par chaque enseignant et enseignante. Sur la base d'une comparaison statistique entre les deux séries de notes, un ajustement est apporté aux notes attribuées par l'enseignant ou enseignante (pour cette composante) à l'ensemble des élèves de l'établissement, si cela est nécessaire. Si les notes attribuées par l'enseignant ou enseignante sont systématiquement trop basses ou trop élevées, l'ajustement sera le même pour chacune des notes attribuées. Toutefois, si ses notes sont trop basses ou trop élevées uniquement dans la tranche de notes supérieure ou inférieure, l'ajustement pourra varier dans toute la gamme de notes.

### Points importants à retenir

- La révision de notation a pour objectif de vérifier si les enseignantes et enseignants ont appliqué les critères d'évaluation avec précision et cohérence lors de la notation des travaux d'élèves.
- À l'issue de la révision de notation, les notes de l'établissement peuvent être augmentées ou diminuées. Elles peuvent également être laissées inchangées.

- Si un facteur de révision de notation est appliqué, cela ne signifie pas que la notation est mauvaise, mais simplement qu'elle ne respecte pas la norme mondiale.

Pour des raisons pratiques, l'IB révisé la notation au niveau des établissements scolaires et non au niveau individuel des enseignantes et enseignants. Il est donc fondamental que tout le personnel enseignant d'un même établissement veille à noter les travaux en respectant la même norme. Le facteur de révision de notation étant appliqué aux notes de l'ensemble des enseignantes et enseignants, l'échantillonnage doit être une représentation fidèle de la notation au sein de l'établissement pour garantir que l'ajustement sera approprié.

Lors de l'analyse des notes d'évaluation interne destinées à la révision de notation, le coordonnateur ou la coordonnatrice doit prêter une attention toute particulière aux éléments qui suivent.

- Dans l'établissement, est-ce que tout le personnel enseignant (chargé d'attribuer les notes d'évaluation interne) suit la même norme de notation ?
- Est-ce que les enseignantes et enseignants notent constamment tous les travaux des élèves en suivant la même norme ?
- Est-ce que les enseignantes et enseignants expliquent clairement les raisons de l'attribution des notes ?

Les enseignantes et enseignants trouveront des conseils précis sur la manière d'appliquer correctement les critères d'évaluation dans la section « [Les commentaires destinés aux établissements](#) » ainsi que dans les *Procédures d'évaluation* en vigueur pour le PEI, le Programme du diplôme et le POP.

Pour de plus amples informations sur les aspects techniques du calcul des facteurs de révision de notation, veuillez vous reporter à l'annexe intitulée « [Gros plan sur la révision de notation de l'évaluation interne](#) ».

## La révision de notation et les attentes du personnel enseignant

Le travail envoyé pour l'évaluation est noté en interne par l'équipe enseignante, puis révisé en externe par l'IB.

Les enseignantes et enseignants doivent se familiariser avec le guide pédagogique qui les concerne pour connaître le détail des critères d'évaluation interne. Chaque critère d'évaluation se compose de descripteurs de niveaux définissant les niveaux atteints spécifiques auxquels correspond une gamme de points. Bien que les descripteurs de niveaux portent sur les aspects positifs du travail, la notion d'échec peut être incluse dans la description des niveaux les plus bas.

Le personnel enseignant doit évaluer les travaux remis pour l'évaluation interne à l'aide des critères, en utilisant les descripteurs de niveaux et en s'appuyant sur les clarifications. Les critères doivent être appliqués selon l'approche de meilleur ajustement. Lorsqu'un travail présente des aspects d'un critère à des niveaux différents, le nombre de points attribués doit refléter le mieux possible l'accomplissement dans son ensemble par rapport audit critère. Il n'est pas nécessaire que tous les aspects du descripteur de niveau soient présents pour que le ou les points correspondants soient attribués. Les descripteurs les plus élevés ne correspondent pas nécessairement à un travail parfait.

Lorsqu'un niveau contient deux notes ou plus, le personnel enseignant doit donner la note la plus élevée si le travail de l'élève correspond dans une large mesure aux qualités décrites (la qualité du travail est alors proche du niveau supérieur). Le personnel enseignant doit donner la note la plus basse si le travail de l'élève démontre les qualités décrites dans une moindre mesure (la qualité du travail est alors proche du niveau inférieur).

Seuls les nombres entiers seront retenus. Les notes partielles, c'est-à-dire les fractions et les décimales, ne sont pas acceptées.

Les critères doivent être considérés indépendamment les uns des autres. Un travail d'élève qui a atteint un niveau élevé pour un critère donné n'atteindra pas nécessairement un niveau élevé pour les autres critères. De même, l'obtention d'un niveau peu élevé pour un critère donné ne signifie pas nécessairement que le travail atteindra un niveau peu élevé pour les autres critères. Le personnel enseignant ne doit pas s'attendre à voir l'évaluation de l'ensemble des élèves suivre une distribution particulière de points.

Dans certaines matières, les descripteurs de niveaux peuvent faire référence à des mots-consignes spécifiques. Ces termes doivent être interprétés conformément à la définition fournie dans la section pertinente du guide pédagogique.

Le personnel enseignant doit garder à l'esprit que lorsqu'un facteur de révision de notation est appliqué au travail évalué en interne, cela n'est pas le signe d'une irrégularité dans la procédure de notation. L'ajustement ne vise qu'à ramener la notation au niveau de la norme mondiale adoptée. Le rapport pédagogique qui paraît pour chaque matière après une session d'examens contient généralement des informations utiles sur les résultats à l'évaluation interne. Du matériel de soutien pédagogique est aussi disponible sur le Centre de ressources pédagogiques.

Le même facteur de révision de notation est appliqué à tous les travaux envoyés pour l'évaluation par chaque établissement. Par conséquent, si plusieurs personnes partagent la responsabilité de l'évaluation interne pour une même matière au sein d'un établissement, il est crucial de mettre en place une normalisation interne pour que la norme de notation adoptée au sein de cet établissement soit la plus cohérente possible. Il pourra notamment s'agir de faire précéder toute notation de discussions sur les critères d'évaluation et d'une analyse des travaux envoyés pour l'évaluation.

Le personnel enseignant et les élèves doivent comprendre les concepts liés à l'intégrité intellectuelle, notamment l'authenticité et la propriété intellectuelle. Les travaux des élèves doivent être entièrement leurs et avoir été préparés dans le respect des exigences énumérées dans le guide pédagogique applicable. Lorsque la collaboration est autorisée, il est impératif que les élèves comprennent clairement la différence

entre collaboration et collusion. Pour de plus amples informations, veuillez consulter la section « [Le renforcement d'un état d'esprit éthique](#) ».

Les enseignantes et enseignants qui souhaitent avoir un résumé de ces informations peuvent imprimer l'affiche intitulée « Notation des travaux d'évaluation interne : le rôle du personnel enseignant dans le processus de révision de notation ». Cette ressource est accessible à partir de l'annexe « Ressources imprimables ».

## La sélection des travaux d'élèves

- Les élèves dans leur intégralité doivent recevoir une note juste. De ce fait, tout travail d'élève doit pouvoir illustrer la norme suivie par l'enseignant ou enseignante et servir d'exemple pour la révision de notation.
- La définition d'un très bon ou d'un très mauvais travail ne sera pas la même pour chaque membre du personnel enseignant ; il est donc essentiel que l'échantillonnage d'un établissement représente tout l'éventail des notes attribuées.
- Pour garantir la transparence et l'équité entre tous les établissements scolaires, l'IB doit pouvoir consulter des preuves montrant le travail à la base des notes définitives des élèves.

Pour garantir la transparence et éviter toute impression de mauvaise conduite, il est important que ce soit l'IB, et non l'établissement scolaire, qui sélectionne les travaux à inclure dans l'échantillonnage destiné à la révision de notation. L'IB suit des principes généraux pour s'assurer que le facteur de révision de notation est le plus fiable possible et, dans la limite de ces principes, que la sélection des travaux est aléatoire.

Le premier principe est que l'IB doit utiliser le plus petit échantillonnage possible pour obtenir un facteur de révision de notation fiable quel que soit l'établissement scolaire. Cette mesure réduit la charge qui repose sur l'établissement, mais aussi les coûts que l'IB pourrait avoir à répercuter sur les établissements en les incluant dans les droits d'examen. Il est toutefois légitime que le nombre de travaux inclus dans l'échantillonnage soit différent en fonction des établissements si l'IB a besoin de plus d'exemples de la notation du personnel enseignant pour déterminer un facteur de révision de notation solide.

Le second principe est que l'IB doit être certain que son facteur de révision de notation est équitable pour la totalité des élèves, quelle que soit leur note. L'expérience a prouvé à l'IB que les enseignantes et enseignants pouvaient avoir des attentes différentes à plusieurs niveaux de la gamme de notes et qu'une personne dont la notation est plus généreuse que la norme mondiale pour un travail de mauvaise qualité pouvait aussi être plus sévère que la norme mondiale pour un travail de bonne qualité. L'échantillonnage de travaux d'évaluation interne est donc soigneusement sélectionné de manière à assurer une représentation appropriée de toute la gamme de notes de l'établissement.

L'IB tend à ne pas sélectionner pour la révision de notation les élèves ayant obtenu la note maximale. De cette manière, les élèves de la tranche de notes supérieure ont la possibilité d'obtenir une meilleure note à la suite d'une révision de notation, si la notation initiale se révèle trop sévère. De même, l'IB ne sélectionne habituellement pas les élèves ayant obtenu un 0.

## Les situations irrégulières

La révision de notation a pour but de vérifier que les enseignantes et enseignants notent les travaux en respectant la norme mondiale, et ce, même si un travail est particulièrement difficile à noter. Le cas échéant, l'enseignant ou enseignante peut avoir à justifier avec plus de détails la note attribuée. Par exemple, si l'élève a bénéficié d'une aide supplémentaire de la part de l'enseignant ou enseignante pour réaliser son travail, cela doit être expliqué dans les commentaires sur la notation et l'examineur ou examinatrice prendra cet élément en compte lors de la révision de la notation.

## L'impossibilité de trouver un facteur de révision de notation

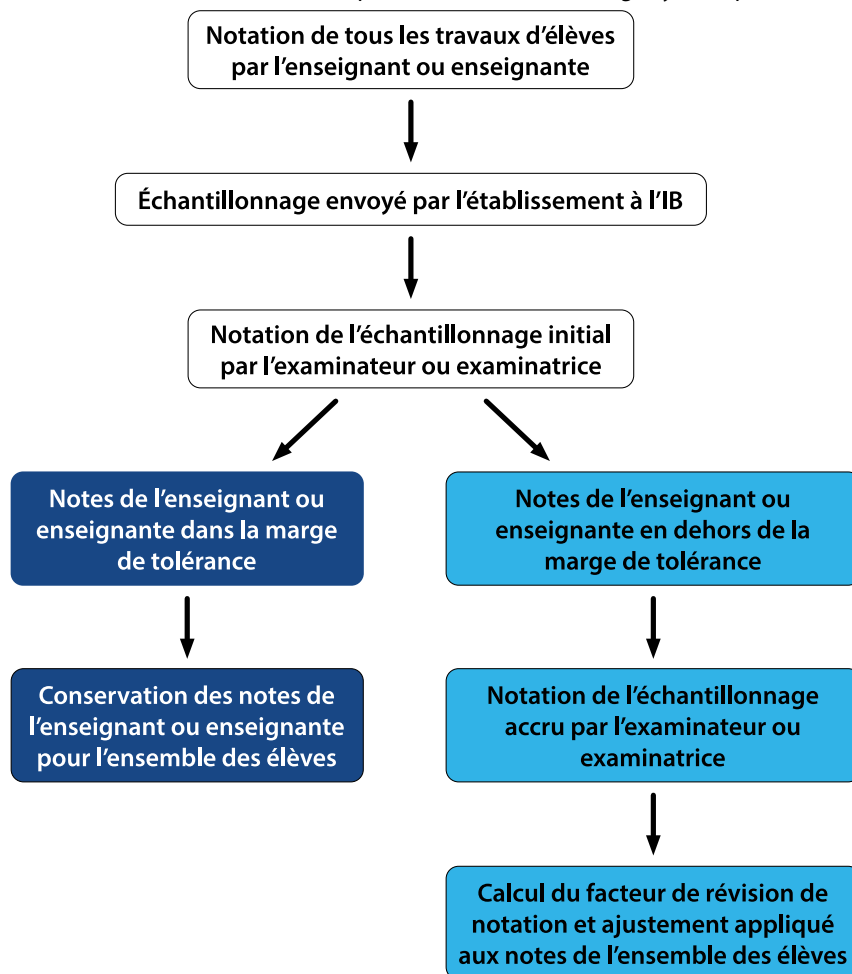
Dans certains cas, il est possible qu'un facteur de révision de notation ne puisse être calculé sur la base de l'échantillonnage de travaux envoyés. Cela se produit lorsque la différence entre la notation de l'examineur ou examinatrice et celle de l'enseignant ou enseignante n'est pas cohérente, ou que les notes de l'examineur ou examinatrice indiquent que la notation initiale a été trop généreuse ou au contraire trop sévère par rapport à la norme mondiale. Le cas échéant, l'IB demande à l'établissement d'envoyer des travaux supplémentaires afin de pouvoir garantir l'application d'un facteur de révision de notation équitable. C'est pourquoi tous les travaux d'élèves doivent être disponibles jusqu'à la clôture de la session d'examens. Dans les situations extrêmement rares où il est impossible d'appliquer un facteur de révision de notation équitable, les notes attribuées seront celles de l'examineur ou examinatrice.

## L'échantillonnage dynamique

- La révision de notation reposant sur l'échantillonnage dynamique signifie que si la notation d'un enseignant ou d'une enseignante se situe dans la marge de tolérance, l'IB accepte ses notes.
- Si certaines notes se situent en dehors, l'IB applique un facteur de révision de notation.
- Il existe un suivi à des fins d'assurance de la qualité pour chaque examinateur et examinatrice vérifiant les notes d'évaluation interne.

Le principe de la révision de notation reposant sur l'échantillonnage dynamique signifie que si un enseignant ou une enseignante montre sa compréhension de la norme mondiale en restant dans la marge de tolérance dans tout l'échantillonnage, il est reconnu que toutes ses notes sont conformes à la norme exigée. Si ses notes sont en dehors de la marge de tolérance en revanche, un facteur de révision de notation est calculé et appliqué. Ce système est expliqué dans la figure 30.

Figure 30  
La révision de notation reposant sur l'échantillonnage dynamique



Pour que la révision de notation soit équitable, les examinateurs et examinatrices doivent comprendre la norme mondiale définie par l'examineur principal ou l'examinatrice principale et la respecter pour procéder à la révision. Une formation sur la norme mondiale est dispensée aux examinateurs et examinatrices participant à la révision de notation, puis un suivi est assuré pour garantir le respect de cette norme dans leur révision de la notation. Tel que décrit dans la section « La normalisation », l'examineur principal ou l'examinatrice principale prépare trois types de travaux évalués en interne comportant des notes définitives :

- des **réponses d'entraînement**, pour expliquer la norme mondiale ;
- des **réponses de validation**, pour vérifier que les examinateurs et examinatrices comprennent la norme mondiale ;
- des **réponses de contrôle**, pour vérifier que les examinateurs et examinatrices respectent systématiquement la norme mondiale.

Une seule et même personne passera en revue tous les travaux évalués en interne d'un établissement donné. Les examinateurs et examinatrices reçoivent cependant des travaux issus d'établissements très divers sans qu'aucun ordre précis ne soit suivi. Cela réduit les risques de se forger une opinion sur la notation (jugée trop généreuse ou trop sévère) dès le premier travail d'élève et de rechercher ensuite cette tendance dans le reste de l'échantillonnage de l'établissement. Par ailleurs, cette méthode permet d'inclure des réponses de contrôle sans qu'elles apparaissent comme telles de façon évidente. Une fois qu'il aura été décidé s'il est nécessaire d'appliquer un facteur de révision de notation, il pourra être demandé aux examinateurs et examinatrices de fournir des commentaires sommaires sur la notation de l'établissement.

Le modèle d'assurance de la qualité est plus complexe avec l'échantillonnage dynamique qu'avec l'autre méthode de notation. L'IB doit également s'assurer que les examinateurs et examinatrices sont capables de déterminer lorsque les notes des enseignantes et enseignants sont trop sévères, trop généreuses ou correctes. La figure 31 illustre l'éventail de copies évaluées en interne requis par le modèle d'assurance de la qualité pour couvrir toutes les situations possibles.

Figure 31

*Éventail de copies évaluées en interne requis par le modèle d'assurance de la qualité*

|  | Cas de figure |  |
|--|---------------|--|
| Notation de l'examineur ou examinatrice en haut de la gamme de notes   | 1             | Notation de l'enseignant ou enseignante trop sévère                  |
|  | 2             | Notation de l'enseignant ou enseignante conforme à la norme mondiale |
|  | 3             | Notation de l'enseignant ou enseignante trop généreuse               |
| Notation de l'examineur ou examinatrice au milieu de la gamme de notes | 4             | Notation de l'enseignant ou enseignante trop sévère                  |
|  | 5             | Notation de l'enseignant ou enseignante conforme à la norme mondiale |
|  | 6             | Notation de l'enseignant ou enseignante trop généreuse               |
| Notation de l'examineur ou examinatrice en bas de la gamme de notes    | 7             | Notation de l'enseignant ou enseignante trop sévère                  |
|  | 8             | Notation de l'enseignant ou enseignante conforme à la norme mondiale |
|  | 9             | Notation de l'enseignant ou enseignante trop généreuse               |

Pour obtenir de plus amples informations sur le système de révision de notation de l'IB, veuillez vous référer à la vidéo intitulée « [Les deux méthodes de notation : comment la notation est-elle effectuée par le personnel enseignant, les examinateurs et les examinatrices ?](#) ».

## L'attribution des notes finales et la totalisation

- Les notes finales doivent avoir la même signification, quelle que soit la session pendant laquelle l'élève passe ses examens.
- Le processus d'attribution des notes finales détermine comment convertir les notes en notes finales pour garantir que ce soit le cas.
- Les seuils d'attribution des notes finales sont fixés à partir d'un éventail de preuves, notamment les résultats de la cohorte et l'avis de spécialistes sur les travaux envoyés pour l'évaluation.
- La note finale globale attribuée à l'élève est celle qui revêt la plus grande importance, et non la note accordée à chaque composante.

Il convient de garder à l'esprit que notes et notes finales sont deux choses différentes. Veuillez vous référer à la section « [La distinction entre les notes et les notes finales](#) » pour comprendre les raisons de cette différence. Le processus d'attribution des notes finales est un mécanisme permettant de décider de la correspondance à faire entre les notes et les notes finales pour chaque situation donnée.

Cette décision est prise à l'issue de plusieurs jours de discussions entre l'examineur principal ou examinatrice principale, l'examineur ou examinatrice en chef, les examinateurs superviseurs et examinatrices superviseuses, et l'IB. Diverses preuves (décrites plus en détail dans la section « [Les preuves utilisées dans le processus d'attribution des notes finales](#) ») sont étudiées pour convenir de ce qui est juste pour les élèves de l'année en cours, mais aussi pour les cohortes qui ont été évaluées dans cette matière les années précédentes. Enfin, l'examineur ou examinatrice en chef fait part à l'IB des résultats recommandés pour la session d'examens concernée.

L'objet principal de ces discussions est l'établissement des seuils d'attribution des notes finales, autrement dit la définition du nombre de points minimum requis pour pouvoir attribuer chaque note finale aux élèves. Parce qu'il n'est pas pratique d'émettre un jugement détaillé pour chaque note finale dans chaque matière, l'IB demande aux examinateurs et examinatrices de recommander plusieurs « seuils d'attribution des notes finales fondés sur une appréciation » et calcule lui-même les autres seuils de façon arithmétique. Pour en savoir plus, veuillez consulter la section « [Les seuils d'attribution des notes finales fondés sur une appréciation et calculés par interpolation](#) ».

Les seuils d'attribution des notes finales ne seront pas forcément les mêmes d'une session d'examens à l'autre. Cela s'explique par le fait que les questions posées aux élèves seront modifiées à chaque itération d'un examen et que les seuils d'attribution des notes finales devront être adaptés pour tenir compte des tâches plus faciles ou plus difficiles. L'IB fait son possible pour garantir que ses examens présentent le même niveau de difficulté chaque année. Néanmoins, en raison des enjeux élevés, il ne peut procéder à aucun examen d'entraînement avant la session, par crainte d'une divulgation des questions.

Pour déterminer où placer les seuils d'attribution des notes finales, les examinateurs et examinatrices et l'IB doivent prendre en considération les aspects présentés dans la figure 32.

Figure 32

*Les aspects pris en considération lors du processus d'attribution des notes finales*

|                        |   |
|------------------------|---|
| <b>Cohorte</b>         | <ul style="list-style-type: none"> <li>En quoi la cohorte qui passe les examens peut-elle être comparée aux cohortes des années précédentes ?</li> </ul>  |
| <b>Examen</b>          | <ul style="list-style-type: none"> <li>Dans quelle mesure l'examen a-t-il rempli les attentes de l'auteur ou autrice ?</li> <li>Qu'ont pensé les enseignantes et enseignants du niveau de difficulté de l'examen ?</li> </ul> |
| <b>Copies d'examen</b> | <ul style="list-style-type: none"> <li>Comment les travaux des élèves se situent-ils par rapport à ceux des années précédentes et aux descripteurs des notes finales ?</li> </ul>   |
| <b>Résultats</b>       | <ul style="list-style-type: none"> <li>En quoi les résultats des élèves sont-ils comparables à ceux des années précédentes ?</li> </ul>   |
| <b>Équilibre</b>       | <ul style="list-style-type: none"> <li>Toutes les preuves disponibles conduisent-elles à la même conclusion ou sont-elles contradictoires ?</li> </ul>  |

Bien souvent, l'IB doit combiner plusieurs composantes d'évaluation pour parvenir à une note finale globale. Le détail de cette manœuvre est expliqué dans la section « [La totalisation](#) ». Mais l'élément fondamental, que les examinateurs et examinatrices doivent garder à l'esprit pendant la réunion d'attribution des notes finales, est que le **résultat global** importe le plus. Si nécessaire, les seuils d'attribution des notes finales de certaines composantes peuvent être moins parfaits pour atteindre un résultat global juste.

Officiellement, le processus d'attribution des notes finales remplit plusieurs objectifs.

- Il permet de déterminer, à partir des travaux envoyés pour l'évaluation, ce qui indique un changement dans la qualité du travail justifiant de passer à un autre descripteur des notes finales.
- Il convient de déterminer pour chaque composante les notes qui correspondront aux différents seuils d'attribution des notes finales.
- Il faut garantir que l'ensemble des seuils établis au niveau de la matière représente une juste attribution des notes finales.

Le processus d'attribution des notes finales sera réussi si :

- les jugements portés sur les copies et les preuves tirées des résultats concordent largement, en tenant compte des informations sur la cohorte ;
- les écarts considérables au niveau des résultats des établissements peuvent être expliqués ;
- l'examineur ou examinatrice en chef et la ou le chef de la direction de l'évaluation de l'IB ont la certitude que les normes d'évaluation ont été respectées.

## Les seuils d'attribution des notes finales fondés sur une appréciation et calculés par interpolation

Les **seuils d'attribution des notes finales fondés sur une appréciation** sont ceux qui sont recommandés par l'examineur ou examinatrice en chef, à partir des discussions tenues lors du processus d'attribution des notes finales. Dans le PEI, le Programme du diplôme et le POP, ces seuils d'attribution des notes finales fondés sur une appréciation sont les suivants : 2 – 3, 3 – 4 et 6 – 7. Les autres seuils, à savoir 1 – 2, 4 – 5 et 5 – 6, sont calculés de façon arithmétique à partir des seuils d'attribution des notes finales fondés sur une appréciation ; c'est pourquoi ils sont qualifiés de « **seuils d'attribution des notes finales calculés par interpolation** ».

En effet, en cas de variation importante du nombre d'élèves ayant obtenu une note finale de 5 ou de 1, une discussion doit être menée en tenant compte de la cohorte. Cela peut conduire à une révision des seuils

d'attribution des notes finales fondés sur une appréciation ou, dans des circonstances exceptionnelles, à un réexamen des travaux envoyés pour l'évaluation se situant à ces seuils.

## Les preuves utilisées dans le processus d'attribution des notes finales

L'attribution des notes finales est un processus reposant sur des preuves, qui exploite diverses informations, et notamment :

- les commentaires des enseignantes et enseignants sur l'évaluation ;
- les jugements spécialisés des examinateurs et examinatrices sur les exemples de travaux d'élèves ;
- l'analyse des informations statistiques comparant les résultats des élèves de l'année en cours à ceux des cohortes des années précédentes.

Aucun type de preuve n'a plus d'importance que les autres : toutes doivent peser de façon égale pour tirer des conclusions.

### L'examen de la cohorte de l'année en cours

La première étape du processus d'attribution des notes finales consiste à se demander en quoi la cohorte qui passe l'évaluation se rapproche des cohortes des années précédentes. Si les établissements ayant inscrit des élèves pour l'année en cours sont majoritairement les mêmes que les années précédentes, il est probable que les différences de réussite s'expliqueront par la difficulté des examens. Si ce n'est pas le cas (par exemple, si un grand nombre de nouveaux établissements inscrivent des élèves à une matière pour la première fois ou si un pourcentage élevé d'élèves repassent l'évaluation), l'IB pourrait décider qu'il est nécessaire que les notes finales tiennent compte des différences de réussite.

Parmi les facteurs qui seront pris en considération lors de la comparaison des cohortes passant l'évaluation, il convient de citer :

- les variations du nombre d'élèves passant l'évaluation ;
- les variations du pourcentage d'élèves passant l'évaluation en anglais, en espagnol, en français et dans d'autres langues, le cas échéant ;
- le nombre de nouveaux établissements et le nombre d'élèves dans ces nouveaux établissements ;
- tout changement au niveau des options choisies par les élèves.

### Le retour d'information sur l'évaluation

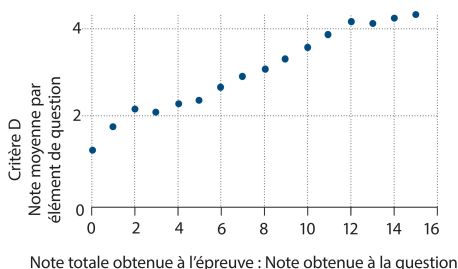
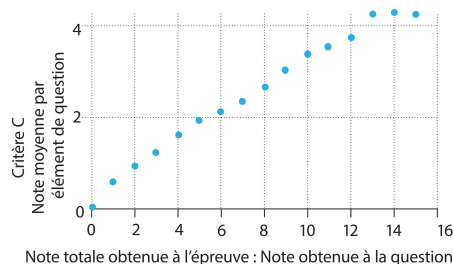
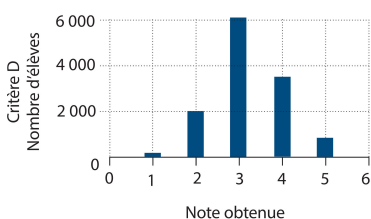
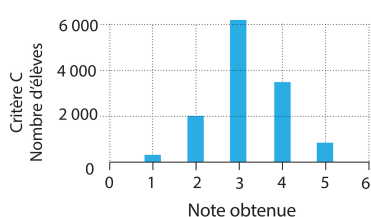
L'étape suivante consiste à se demander comment s'est déroulée l'évaluation par rapport aux attentes. Si une question en particulier se révèle bien plus difficile que prévu, ou si des questions censées faire une distinction entre les élèves méritant la note finale 6 et les élèves méritant la note finale 7 n'y parviennent pas, l'équipe chargée de l'attribution des notes finales doit en tenir compte lors de l'établissement des seuils d'attribution des notes finales.

Les membres de l'équipe utilisent les commentaires des enseignantes et enseignants sur les examens comme base de leur discussion ainsi que les statistiques sur les résultats obtenus à chaque question ou élément (voir la figure 33). Ces membres utilisent également les résumés des rapports provenant des examinateurs et examinatrices, et savent s'appuyer sur leur expérience de la notation des travaux des élèves lors de la session.

Figure 33

Exemple de statistiques à la disposition des examinateurs et examinatrices concernant les éléments

| Analyse statistique des éléments |                    |                                  |  |                                      |  |   | Corrélation entre l'élément et le reste de la composante |                            |
|----------------------------------|--------------------|----------------------------------|--|--------------------------------------|--|---|--|----------------------------|
| Question                         | Nombre de réponses | Nombre total d'élèves à l'examen | Pourcentage d'élèves ayant tenté de répondre à la question | Note moyenne par élément de question | Note la plus haute obtenue par élément de question | Écart type de la population des notes obtenues aux éléments de question | Question   | Coefficient de corrélation |
| 01 Critère A                     | 12 071             | 42 042                           | 28,71 %  | 2,7                                  | 5  | 0,95  | 01 Critère A   | 0,8044                     |
| 01 Critère B                     | 12 071             | 42 042                           | 28,71 %  | 2,6                                  | 5  | 0,94  | 01 Critère B   | 0,8018                     |
| 01 Critère C                     | 12 071             | 42 042                           | 28,71 %  | 2,8                                  | 5  | 0,84  | 01 Critère C   | 0,7983                     |
| 01 Critère D                     | 12 071             | 42 042                           | 28,71 %  | 3,2                                  | 5  | 0,79  | 01 Critère D   | 0,7253                     |
| 02 Critère A                     | 29 848             | 42 042                           | 71,00 %  | 2,8                                  | 5  | 0,98  | 02 Critère A   | 0,7954                     |
| 02 Critère B                     | 29 848             | 42 042                           | 71,00 %  | 2,7                                  | 5  | 0,93  | 02 Critère B   | 0,7944                     |
| 02 Critère C                     | 29 848             | 42 042                           | 71,00 %  | 2,9                                  | 5  | 0,83  | 02 Critère C   | 0,7865                     |
| 02 Critère D                     | 29 848             | 42 042                           | 71,00 %  | 3,2                                  | 5  | 0,79  | 02 Critère D   | 0,7320                     |



## La recherche de preuves dans les copies

Par la suite, les examinateurs superviseurs et examinatrices superviseuses comparent pour chaque note les descripteurs des notes finales à des copies présentant des notes proches des seuils d'attribution des notes finales prévus. Lors de cette étape du processus d'attribution des notes finales, il est essentiel de se concentrer non pas sur les notes attribuées, mais sur la nature des réponses des élèves et la mesure dans laquelle elles correspondent aux descripteurs des notes finales.

Avant de commencer à passer en revue les copies, il est souvent utile de revenir aux exemples de travaux des années précédentes, afin de se rappeler les attentes correspondant aux différentes notes finales. Les copies utilisées pour déterminer les seuils doivent être disponibles pendant tout le processus d'attribution des notes finales.

Avant la réunion de délibérations, les membres de l'équipe d'examinateurs superviseurs et examinatrices superviseuses, et surtout les examinateurs principaux et examinatrices principales, proposent des seuils d'attribution provisoires des notes finales pour chaque composante. Il s'agit d'indiquer au niveau de quelles notes il paraît nécessaire de fixer les seuils, sur la base des expériences passées des niveaux attendus. Après avoir pris connaissance de ces seuils provisoires, ainsi que de l'avis partagé sur le déroulement de chaque examen et de la distribution globale des notes, la ou le responsable de matière de l'IB dispose des renseignements et des preuves nécessaires pour suggérer la gamme de notes à représenter dans l'échantillonnage d'attribution des notes finales.

Chaque examinateur superviseur et examinatrice superviseuse doit passer en revue les copies d'élèves sélectionnées et déterminer quel descripteur des notes finales reflète le mieux la qualité du travail. Cette tâche s'avère compliquée, car même les élèves qui présentent des réponses d'un niveau relativement homogène montreront souvent des qualités associées à diverses notes finales au fil de leurs réponses. Il est acceptable, et même souvent utile, d'indiquer si une copie correspond tout juste à un descripteur des notes

finales ou si elle est proche de la note finale supérieure suivante (ce qui est souvent indiqué sous la forme « 7- » ou « 6+ » par exemple).

Il est essentiel d'éviter tout type de préjugé ou de biais pour mener cette tâche à bien. C'est pourquoi il ne faut pas donner son point de vue avant que chaque personne ait consigné son résultat de son côté. De même, il est utile de ne pas informer les examinateurs et examinatrices (à l'exception de l'examineur ou examinatrice en chef) des conclusions de l'analyse statistique avant la fin de cette étape, sauf renseignement important sur les résultats obtenus à un élément dans l'ensemble de la cohorte.

Déterminer la note finale qui correspond le mieux à une copie est un exercice très subjectif, d'où la grande possibilité d'écart entre les examinateurs et examinatrices. Par ailleurs, il n'est pas aberrant d'attribuer une note finale plus élevée à l'élève qui a montré une bonne compréhension mais fait quelques erreurs, qu'à l'élève qui a excellé à la tâche la plus facile (et obtenu de ce fait beaucoup de points) mais n'a pas montré le même niveau de compréhension, même si l'élève en question a obtenu un peu plus de points. Il ne faut pas non plus perdre de vue que cet examen des copies ne concerne qu'un petit nombre d'exemples de travaux d'élèves.

Des études ont révélé un don particulier chez les examinateurs et examinatrices pour repérer les copies affichant un niveau de qualité différent des autres copies déjà examinées (voir la section « [Les approches de la notation](#) »). L'IB leur demande donc de commencer avec la note la plus élevée de l'échantillonnage et de passer en revue les copies jusqu'à estimer ne plus trouver constamment des preuves des qualités associées à la note finale la plus élevée. Il leur faudra ensuite passer à la note la plus basse de l'échantillonnage et analyser les copies jusqu'à trouver constamment des preuves des qualités associées à la note finale la plus élevée. Si nécessaire, la gamme de notes de l'échantillonnage peut être augmentée.

Lorsque toutes les décisions des examinateurs superviseurs et examinatrices superviseuses en matière d'attribution d'une note finale sont rassemblées, elles doivent faire apparaître la gamme de notes au niveau de laquelle situer un seuil d'attribution des notes finales. Il s'agit d'une « zone d'incertitude », qui va de la note la plus basse à la note la plus élevée pour lesquelles suffisamment de preuves existent pour situer le seuil d'attribution des notes finales. Aucune définition précise n'indique comment décider de la zone d'incertitude. Les examinateurs superviseurs et examinatrices superviseuses doivent en discuter ensemble, en s'appuyant sur les décisions prises par chaque personne.

Figure 34

Exemple de résultats de l'analyse des copies d'examen

| Numéro de la copie d'examen | Note | Examineur ou examinatrice 1 (examineur principal ou examinatrice principale) | Examineur ou examinatrice 2 | Examineur ou examinatrice 3 | Examineur ou examinatrice 4 |
|-----------------------------|------|--|-----------------------------|-----------------------------|-----------------------------|
| 1                           | 51   | 4  | 4                           | 4                           | 4                           |
| 2                           | 51   | 4  | 4                           | 4                           | 4                           |
| 3                           | 50   | 3+   | 4                           | 4                           | 3+                          |
| 4                           | 50   | 4-   | 4                           | 4                           | 3+                          |
| 5                           | 49   | 4-   | 4-                          | 4-                          | 4-                          |
| 6                           | 49   | 3+   | 3                           | 4-                          | 3+                          |
| 7                           | 48   | 4-   | 3+                          | 4-                          | 3+                          |
| 8                           | 48   | 3+   | 4-                          | 4-                          | 3+                          |
| 9                           | 47   | 3  | 3                           | 3+                          | 3                           |
| 10                          | 47   | 3  | 3                           | 3                           | 4-                          |

Seuls les questionnaires à choix multiple ne sont pas visés par ce processus. L'expérience montre en effet qu'il est très difficile d'émettre des jugements sur les seuils d'attribution des notes finales à partir de la qualité du travail envoyé pour l'évaluation lorsque les examens sont composés uniquement de questions à choix multiple. Cela peut s'expliquer par le fait que les réponses contiennent peu de preuves à partir desquelles baser un jugement sur les réalisations réelles des élèves. Pour ce type d'examen, les seuils d'attribution des notes finales sont calculés de manière à ce que le pourcentage d'élèves à chaque note finale soit le plus proche possible de celui obtenu à partir d'une appréciation pour les examens qui s'en rapprochent le plus.

## L'analyse des statistiques sur les résultats

L'on pourrait croire que dans un système d'évaluation critériée dépendant du jugement professionnel des correcteurs et correctrices, les examinateurs superviseurs et examinatrices superviseuses devraient être capables d'établir des seuils d'attribution des notes finales en considérant simplement les questions de l'examen et les réponses attendues pour chacune. En réalité, il est très difficile de formuler ces jugements avant de connaître la façon dont les élèves ont répondu. Cresswell (2000) est parvenu à la conclusion que les correcteurs finaux et correctrices finales parvenaient le plus souvent à déterminer si des examens étaient plus faciles ou plus difficiles d'une session d'examens à une autre, mais pas à estimer dans quelle mesure ils étaient plus faciles ou plus difficiles.

L'objectif de la réunion d'attribution des notes finales étant de maintenir une signification ou un niveau cohérents des notes finales, il est intéressant de réfléchir à l'une des définitions de la comparabilité données par Cresswell.

Deux examens peuvent être considérés comme de niveau comparable si deux groupes d'élèves affichant le même éventail d'aptitudes et les mêmes résultats antérieurs, fréquentant des établissements semblables qui appliquent les mêmes politiques d'admission, étudiant avec des enseignants et enseignants aux compétences égales, et présentant le même niveau de motivation, reçoivent des notes finales qui suivent strictement la même répartition, après avoir étudié leurs programmes respectifs et passé les examens.

(Cresswell, 1996, p. 57 – 84, traduction libre)

Pour faciliter leur travail d'attribution des notes finales, des **seuils recommandés statistiquement** sont fournis aux examinateurs superviseurs et examinatrices superviseuses. Il s'agit des seuils qui produiraient exactement le même pourcentage cumulé (soit la proportion totale d'élèves ayant obtenu la note finale considérée ou une note finale plus élevée) que l'année précédente pour la note finale concernée.

Les examinateurs et examinatrices participant au processus d'attribution des notes finales reçoivent aussi des informations sur la note moyenne obtenue par les élèves et des histogrammes illustrant la répartition réelle des notes. Même si ces informations sont prises en compte dans les seuils recommandés statistiquement, il est souvent utile de les analyser de près.

Il est important de garder à l'esprit qu'un groupe d'élèves différent passera les examens à chaque session. De ce fait, il est probable que les postulats de la définition de la comparabilité donnée précédemment ne soient pas totalement vérifiés. Cela sera notamment le cas si les deux cohortes sont très différentes. De même, lorsqu'il existe un très grand nombre d'élèves avec des expériences scolaires similaires (par exemple, scolarisation dans des écoles du monde de l'IB), il est plus probable que les différences de résultats notables soient dues à la décision de répondre à des séries de questions différentes plutôt qu'aux écarts au niveau des accomplissements.

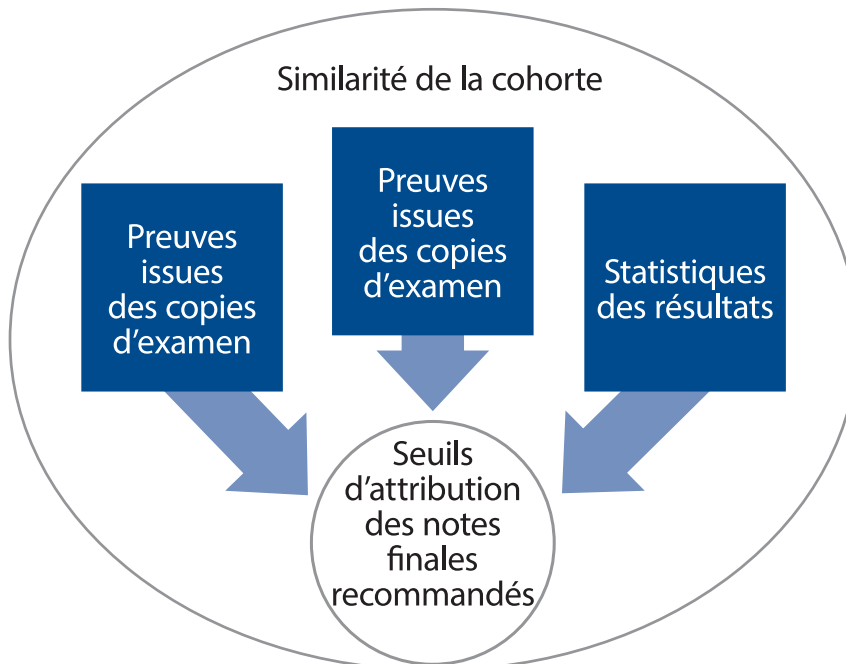
La taille de la cohorte doit être prise en considération lors de l'examen de l'importance des seuils recommandés statistiquement, mais l'IB ne suit aucune règle formelle sur la question, car d'autres facteurs sont aussi importants. En règle générale, si les élèves se comptent en dizaines seulement, des écarts considérables pourront être observés au niveau du résultat global, mais il est assez peu probable que cela se produise avec des milliers d'élèves.

## Trouver un juste équilibre entre toutes les preuves

Aucun type de preuve n'a plus d'importance que les autres. Le travail de l'équipe chargée de l'attribution des notes finales consiste à trouver un juste équilibre entre toutes les preuves pour faire une recommandation.

Figure 35

Preuves venant étayer le choix des seuils d'attribution des notes finales



Souvent, les différentes preuves suggèrent les mêmes résultats (par exemple, les seuils recommandés statistiquement se situent dans la zone d'incertitude), mais il peut arriver que certaines soient en contradiction. Dans de tels cas, il est essentiel d'expliquer cet écart dans la justification de la décision finale.

Lors de cette dernière étape du processus d'attribution des notes finales, il est possible de montrer les effets éventuels des différentes décisions sur les résultats globaux de la cohorte. Conformément au principe selon lequel c'est le résultat global qui importe, et non celui obtenu dans les différentes composantes, l'équipe chargée de l'attribution des notes finales pourrait comparer la note finale moyenne de la cohorte aux prévisions des enseignantes et enseignants, en se référant aux années précédentes pour comprendre la fiabilité d'une telle comparaison. Une autre méthode consiste à exclure les nouveaux établissements des résultats globaux pour vérifier s'ils affichent un niveau très différent des établissements plus expérimentés.

Enfin, l'équipe chargée de l'attribution des notes finales doit soumettre ses recommandations, accompagnées des justifications nécessaires, en vue de leur approbation. La totalisation peut alors avoir lieu. Cela désigne le processus d'addition des notes et des seuils issus des différentes composantes pour produire une note définitive ou un seuil d'attribution des notes finales global. Pour y parvenir, il peut s'avérer nécessaire d'appliquer un coefficient aux notes globales ou aux seuils des composantes. Pour plus d'informations, veuillez vous reporter à la section « [La totalisation](#) ».

## Les seuils d'attribution des notes finales fixes

Pour l'essentiel de l'évaluation interne, et une partie des composantes notées en externe, il est demandé aux élèves d'effectuer des tâches presque identiques à chaque session. C'est notamment le cas avec la préparation d'un portfolio d'arts et le projet personnel, pour ne citer que ces exemples. Dans de telles situations, il peut raisonnablement être supposé, puisque l'IB maintient la même norme de notation chaque année (grâce à la normalisation à l'aide des travaux des années précédentes), que les seuils d'attribution des notes finales resteront eux aussi les mêmes.

Lors du processus d'attribution des notes finales, l'examineur principal ou examinatrice principale et l'examineur ou examinatrice en chef devront se demander si des preuves indiquent que les seuils d'attribution des notes finales existants ne sont pas appropriés. Dans la plupart des cas, l'IB attend une conclusion confirmant que les seuils existants peuvent être conservés.

Malgré cela, il ne doit faire aucun doute que ces seuils sont vérifiés chaque année, et qu'ils peuvent être modifiés si des éléments prouvent qu'ils ne sont plus efficaces pour maintenir la comparabilité des notes finales entre les sessions. Cela pourrait se produire en cas de changement important du comportement des élèves (par exemple, l'influence des progrès technologiques sur la réalisation d'une tâche), de changements dans les travaux remis qui renforcent l'alignement sur le barème de notation sans améliorer la qualité globale du travail, ou de révision d'une approche de l'IB en matière de notation ou de révision de notation des travaux qui modifie la norme de notation.

Pour mieux comprendre pourquoi les seuils d'attribution des notes finales changent, et ce que cela signifie, veuillez vous référer à la vidéo intitulée « [Seuils d'attribution des notes finales : pourquoi les seuils d'attribution des notes finales changent-ils ? Qu'est-ce que cela signifie ?](#) ».

## L'étude et l'approbation des recommandations relatives aux seuils d'attribution des notes finales

L'équipe chargée de l'attribution des notes finales est parvenue à des recommandations quant aux seuils à retenir. Ces recommandations et justifications sont alors présentées dans le **rapport sur la répartition des notes dans la matière rédigé sous 24 heures**. Ce rapport comprend également les données sous-jacentes, comme des changements concernant la cohorte étudiant le cours et les notes finales attribuées avec les seuils proposés en comparaison des années précédentes. Ce rapport est ensuite étudié par des membres de la direction de la division de l'évaluation de l'IB qui décideront si les arguments avancés sont suffisamment solides.

Si l'équipe de direction de l'évaluation de l'IB n'est pas satisfaite des recommandations faites, elle expliquera ses inquiétudes à la ou au responsable de matière et lui demandera de donner la priorité à certains aspects des preuves ou d'approfondir son analyse pour mieux soutenir ses recommandations.

La ou le chef de la direction de l'évaluation de l'IB décide en dernier ressort des seuils d'attribution des notes finales, à la lumière des recommandations faites par l'équipe dédiée.

## L'octroi du diplôme ou du certificat du programme

Les certificats et diplôme de l'IB (certificat du PEI, diplôme de l'IB et certificat du POP de l'IB) ne sont pas décernés à l'issue du processus d'attribution des notes finales. Leur octroi est décidé pendant la préparation de la publication des résultats, selon le processus décrit dans la section « [La préparation de la publication des résultats](#) ».

## Les observateurs et observatrices

L'IB s'est engagé à accroître la transparence de ses processus d'évaluation et à faciliter la compréhension générale de son mode d'attribution des notes finales. Dans le cadre de cet engagement, des enseignantes et enseignants viennent observer les réunions de délibérations (les modalités dépendent de la façon dont se tiennent ces réunions, c'est-à-dire en présentiel ou en ligne). Ces observateurs et observatrices devront rendre compte de leur expérience à leurs collègues et préparer un rapport à l'intention de toute la communauté du réseau de collaboration de l'IB (IBEN). Pour obtenir de plus amples informations, veuillez vous adresser au service [L'IB vous répond](#).

## Les principes du processus d'attribution des notes finales

Le processus d'attribution des notes finales de l'IB repose un certain nombre de principes fondamentaux qui sont développés ci-après.

- Les seuils d'attribution des notes finales 3 – 4, 6 – 7 et 2 – 3 sont déterminés (dans cet ordre) à partir de toutes les preuves disponibles (appréciations et statistiques). Lorsqu'aucun travail envoyé pour

l'évaluation ne peut servir directement à l'établissement de ces seuils, les preuves disponibles sont utilisées pour établir le seuil le plus approprié.

- Les autres seuils d'attribution des notes finales sont ensuite déterminés de façon arithmétique, en suivant la procédure adaptée.
- Les décisions relatives aux seuils d'attribution des notes finales sont prises sur la base d'une analyse triangulaire des preuves issues des jugements des examinateurs et examinatrices, des preuves statistiques et des informations sur la cohorte. Toutes ces preuves doivent peser de façon égale et un compromis doit être trouvé.
- Si la cohorte passant l'évaluation présente de grandes similitudes avec celles des années précédentes, l'IB s'attend à constater des résultats globalement similaires aux précédents. Quelques particularités sont cependant à souligner.

Les cohortes changent souvent beaucoup d'une année à l'autre, notamment dans les matières présentées par un petit nombre d'élèves.

Si les résultats sont très différents des années précédentes, l'IB attend des preuves solides pour en comprendre les raisons.

Si les tâches d'une évaluation sont globalement similaires à celles des années précédentes, l'IB s'attend à retrouver des seuils d'attribution des notes finales similaires aux précédents.

Tous les seuils d'attribution des notes finales peuvent changer chaque année, y compris pour les tâches d'évaluation interne.

L'IB s'efforce de s'assurer que les évaluations présentent le même niveau de difficulté chaque année, mais ce niveau pourra varier dans certains examens.

La priorité est donnée aux seuils d'attribution des notes finales par matière. C'est ce résultat qui revêt une réelle importance pour l'élève et qui sert de base aux parties prenantes pour prendre les décisions.

Les seuils d'attribution des notes finales par composante sont une étape essentielle pour obtenir un résultat global solide par matière, mais l'accumulation d'effets limités au niveau des composantes peut avoir un effet considérable sur les résultats dans la matière prise dans son ensemble.

## Les vérifications de la qualité

- Le résultat le plus important de l'évaluation est la note finale que l'élève reçoit ; c'est pourquoi cette note est au cœur des vérifications finales.
- Ces vérifications finales visent à mettre en lumière tout élément inhabituel.

Après l'établissement des seuils d'attribution des notes finales, certaines vérifications supplémentaires de la qualité sont effectuées de manière à garantir que les notes finales attribuées aux élèves sont justes et correctes.

L'une de ces vérifications est la recorection de copies dites « à risque », qui permet de repérer les élèves ayant une note finale étonnamment basse. Certaines recorections de copies à risque portent sur le travail des examinateurs et examinatrices dont la notation s'est imposée comme nécessitant un examen complémentaire. Cette recorection est prise en charge par des examinateurs et examinatrices ayant prouvé leur capacité à appliquer de manière cohérente la norme de notation de l'examineur principal ou examinatrice principale. La recorection de copies à risque est effectuée non pas pour déterminer s'il est possible de trouver des points supplémentaires, mais pour vérifier que la note actuelle est appropriée et ainsi garantir des notes finales correctes pour les élèves.

Le personnel chargé de l'évaluation de l'IB procède également à une vérification pour s'assurer de l'adéquation des notes finales par matière destinées aux établissements.

## Le comité d'attribution des notes finales

- Le comité d'attribution des notes finales offre un cadre formel au sein duquel l'IB présente aux parties prenantes conviées les éléments de la session d'évaluation à examiner pour le PEI, le Programme du diplôme et le POP. Un tel comité n'existe pas pour le PP, car aucun résultat final n'est à vérifier ni à authentifier pour ce programme.
- Le comité d'attribution des notes finales est constitué de membres de plusieurs divisions de l'organisation, notamment de membres de la direction et d'examineurs et examinatrices en chef. En plus de donner son approbation finale pour la publication des résultats, il étudie également les cas contrevenant à la politique d'intégrité intellectuelle de l'IB et toute autre question pertinente.

Le comité d'attribution des notes finales est chargé de la dernière étape du processus d'approbation des résultats des élèves pour les programmes de l'IB assortis d'évaluations envoyées à l'IB. Il s'agit de l'organe qui passe en revue les recommandations de la ou du chef du bureau de l'éducation et du directeur ou de la directrice de l'évaluation relatives au respect des normes de l'IB pendant la session. Les résultats issus du comité doivent être attribués pour le compte du Conseil de fondation. Ce comité sert également de cadre pour établir la politique et les précédents qui s'appliquent à l'octroi des titres de l'IB.

La composition et les attributions précises du comité varient légèrement en fonction du programme, mais il est en principe constitué du même nombre de membres ayant droit de vote, à savoir :

- des membres de la direction du bureau de l'éducation et du bureau des services aux établissements de l'IB ;
- des examinateurs et examinatrices en chef des différentes disciplines.

Contrairement aux membres de la direction de la division de l'évaluation de l'IB qui sont responsables de l'approbation des seuils d'attribution des notes finales, le comité d'attribution des notes finales agit comme un organe de surveillance. Il examine les résultats au niveau global (par exemple, les taux de réussite globaux au programme) ainsi que tous les points portés à son attention par des membres de la division de l'évaluation de l'IB.

Le comité passe également en revue les cas de mauvaise conduite et de mauvaise administration, qui font l'objet de discussions approfondies au sein du sous-comité d'intégrité intellectuelle.

Le dernier rôle du comité d'attribution des notes finales consiste à mener une réflexion sur la réussite de la session d'examens et à faire des recommandations à l'IB pour les sessions ultérieures.

## Le conflit d'intérêts

Le comité d'attribution des notes finales est un organe de prise de décision important. Il est essentiel qu'il donne une image de transparence et d'indépendance. La politique adoptée est donc la suivante : dès qu'un conflit d'intérêts apparaît pour l'un des éléments à l'ordre du jour, la personne en question quittera la pièce pour la discussion concernée, une mesure de précaution qui est reconduite au début de chaque réunion.

## Les observateurs et observatrices

L'IB est très favorable à la présence d'observateurs et observatrices aux réunions du comité d'attribution des notes finales, et ce, pour les raisons suivantes :

- rendre les procédures plus transparentes ;
- donner la possibilité de faire des suggestions de modification et d'amélioration des processus ;
- reconnaître le partenariat entre les différentes parties constitutives de l'IB.

Pour que l'IB puisse atteindre ces objectifs, il demande aux observateurs et observatrices de remettre un rapport sur leur participation à la ou au chef du bureau de l'éducation dans les deux semaines qui suivent la réunion. Il convient d'y faire des observations d'ordre général sur les procédures et processus actuellement suivis, et de suggérer les modifications et améliorations jugées nécessaires. Aucun commentaire ne doit être fait sur un cas précis abordé par le comité.

En raison du caractère sensible des questions débattues par le comité, les observateurs et observatrices ont également des obligations à respecter en matière de confidentialité et de conflit d'intérêts. Les observateurs et observatrices n'ont pas de droit de vote pour les décisions du comité.

Pour savoir plus précisément comment devenir observateur ou observatrice d'une réunion du comité d'attribution des notes finales, veuillez envoyer un courriel à [support@ibo.org](mailto:support@ibo.org).

## La préparation de la publication des résultats

- Une fois toutes les évaluations notées et tous les seuils d'attribution des notes finales établis, plusieurs processus doivent être suivis pour garantir que les résultats des élèves sont prêts à être publiés.
- Dans certaines circonstances, lorsqu'il n'existe aucune preuve du travail envoyé pour l'évaluation, l'IB doit estimer la note en question en suivant la procédure pour note manquante.
- Pour les élèves étudiant un programme complet, les résultats par matière sont combinés de façon à déterminer s'il est possible de leur décerner le certificat du PEI, le diplôme de l'IB ou le certificat du POP de l'IB.
- La publication des résultats intègre un système de contrôle des changements efficace. Ainsi, toute modification des notes finales est détectée et les parties prenantes concernées, essentiellement les établissements scolaires et les universités, en sont informées.

L'établissement des seuils d'attribution des notes finales n'est pas directement suivi de la publication des résultats pour les élèves : plusieurs procédures doivent être entreprises entre ces deux étapes. Si quelques procédures, comme la conservation des notes pour les matières anticipées, concernent seulement des élèves en particulier, d'autres s'appliquent à leur intégralité.

### Les notes manquantes

- Lorsque l'IB n'est pas en mesure d'accéder à un travail pour une raison indépendante de la volonté de l'établissement scolaire ou de l'élève, une note sera estimée de façon à minimiser les éventuels effets négatifs sur l'élève.
- Une telle estimation doit se fonder sur des preuves. Si les travaux envoyés pour l'évaluation sont trop peu nombreux pour être exploitables, une autre solution doit être trouvée pour remplacer la note manquante.
- La procédure pour note manquante se base sur les résultats moyens de l'élève. L'estimation procurera donc un avantage ou un désavantage pour un nombre relativement équivalent d'élèves. Le résultat le plus juste sera toujours celui obtenu en corrigeant le travail réel envoyé pour l'évaluation.

Il peut arriver qu'un travail ne puisse être remis à l'IB en vue d'une notation pour des raisons indépendantes de la volonté de l'établissement scolaire et de l'élève. C'est notamment le cas lorsque des examens au format papier sont perdus par les services postaux ou que des élèves contractent subitement une maladie le jour de l'examen.

Dans de telles situations, l'IB peut engager la procédure pour note manquante afin d'estimer une note pour l'élève en question.

Il est pertinent de lancer cette procédure dans les cas où l'absence du travail de l'élève est due aux actions de l'IB ou de tiers (à l'exclusion de l'établissement) et où il ne serait pas raisonnable de demander à l'élève de recommencer l'évaluation.

Quelle que soit la situation, la procédure pour note manquante doit se fonder sur des preuves des accomplissements de l'élève. De ce fait, si l'IB dispose d'informations peu nombreuses, il est très peu probable qu'une note juste puisse être estimée.

La procédure pour note manquante s'inspire des résultats moyens obtenus par l'ensemble des élèves dans la composante d'évaluation concernée, qui sont comparés aux résultats obtenus dans d'autres

composantes. En utilisant cette moyenne, l'IB s'attend à ce qu'autant d'élèves obtiennent de moins bons résultats avec cette estimation que de meilleurs résultats. Par conséquent, le résultat sera toujours plus juste si l'IB peut noter le travail réel de l'élève.

Si un établissement a inscrit cinq élèves ou plus dans la matière et le niveau concernés, l'IB se sert des données de l'établissement pour déterminer les notes manquantes. Si l'établissement a inscrit moins de cinq élèves, l'IB utilise les données mondiales pour calculer les notes manquantes.

Figure 36

### Comment déterminer une note manquante

Calcul pour déterminer la note manquante d'un ou d'une élève :

$$\text{Quotient} = \frac{\text{Note totale pondérée de l'élève pour les autres composantes}}{\text{Note totale pondérée mondiale (ou de l'établissement) pour les autres composantes}}$$

$$\text{Note manquante de l'élève} = \text{Note pondérée mondiale (ou de l'établissement) pour la composante manquante} \times \text{Quotient}$$

Le quotient compare les notes de l'élève à la moyenne mondiale ou de l'établissement pour les composantes achevées par l'élève.

- Si les résultats de l'élève sont plus élevés que la moyenne, alors ce chiffre sera supérieur à 1.
- Si les résultats de l'élève sont plus bas que la moyenne, alors ce chiffre sera inférieur à 1.

Ce mode de calcul, comme toutes les autres procédures pour note manquante, n'est que la « meilleure supposition » possible à partir des preuves disponibles. Il est toujours préférable de pouvoir noter le travail réel de l'élève.

## La procédure pour l'attribution d'une note finale manquante

Concernant le PEI, s'il n'existe qu'une composante dans le modèle d'évaluation d'une matière, il est impossible d'utiliser la procédure pour note manquante, car aucune preuve ne pourra être exploitée pour estimer la note. Le cas échéant, l'IB adopte une approche différente : la procédure pour l'attribution d'une note finale manquante spécifique au PEI.

L'un des inconvénients de vouloir réduire la charge qui pèse sur les élèves est que, avec l'utilisation d'une seule composante d'évaluation, les preuves disponibles pour le PEI sont très minces. Cela signifie que la procédure pour l'attribution d'une note finale manquante est moins fiable aux yeux de l'IB que la procédure pour note manquante existant pour les autres programmes et donc qu'elle ne doit servir que dans des cas exceptionnels.

Si l'élève a respecté les exigences de l'évaluation et a dûment remis son travail, mais ne peut pas recevoir de note de l'IB pour des raisons qui échappent à son contrôle ou à celui de l'établissement, l'IB utilise la procédure pour l'attribution d'une note finale manquante.

Pour que la procédure pour l'attribution d'une note finale manquante puisse être engagée, l'élève doit avoir obtenu une note finale dans quatre autres cours du PEI au moins. L'IB n'est pas en mesure d'estimer des résultats pour des élèves ayant moins de notes finales, car il ne dispose pas de preuves suffisantes pour pouvoir faire une estimation éclairée.

Pour déterminer la note finale manquante, il convient de calculer la note finale moyenne à partir de toutes les autres matières pour lesquelles il existe une note finale obtenue à partir des travaux envoyés pour l'évaluation au cours des 18 derniers mois.

Si la décimale est supérieure ou égale à 5, la note finale moyenne doit être arrondie à l'entier supérieur.

Si la décimale est inférieure à 5, la note finale moyenne doit être arrondie à l'entier inférieur.

Ce mode de calcul, comme toutes les autres procédures pour note manquante, n'est que la « meilleure supposition » possible à partir des preuves disponibles. Il est toujours préférable de pouvoir noter le travail réel de l'élève.

## Les résultats au programme

- L'IB octroie un diplôme ou un certificat dès lors que certains critères précis sont respectés et non sur la base d'un jugement au cas par cas.
- Chaque programme possède ses propres conditions requises pour l'octroi du diplôme ou du certificat.

Une fois que l'IB connaît les résultats dans toutes les matières, il peut effectuer un calcul pour déterminer si l'élève peut prétendre au certificat du PEI, au diplôme de l'IB ou au certificat du POP de l'IB, selon les cas.

Le mode de calcul des résultats du PEI, du Programme du diplôme et du POP, et les critères de réussite, sont expliqués en détail sur le [site Web de l'IB](#) et dans les *Procédures d'évaluation* spécifiques à chaque programme.

## La publication des résultats

Du point de vue de l'IB, la publication des résultats consiste essentiellement à mettre en œuvre un protocole très strict de contrôle des modifications.

Au cours de la session d'examens, les informations sont constamment mises à jour, ce qui permet à l'IB d'avoir en permanence une idée précise de ce qui se passe. Une fois les résultats portés à la connaissance des élèves et des établissements scolaires, l'IB doit leur indiquer très clairement qu'aucune modification ne se produira sans que toutes les parties prenantes en soient dûment informées. Concernant le Programme du diplôme et le POP, cela vaut tout particulièrement pour les universités, qui recevront un relevé de notes rassemblant tous les résultats de l'élève.

Des modifications peuvent être autorisées après la publication des résultats en raison de réclamations concernant les résultats, d'une confirmation des résultats en suspens par les établissements scolaires ou de la résolution de problèmes d'intégrité intellectuelle.

## Les réclamations concernant les résultats et les demandes d'appel de l'évaluation

- Les réclamations concernant les résultats visent à permettre aux établissements d'indiquer à l'IB où ils estiment qu'une erreur a été commise au cours du processus de notation.
- Les normes de notation appliquées pendant les recorections effectuées dans le cadre des réclamations concernant les résultats doivent être les mêmes que celles suivies lors de la session d'examens principale. Pour s'en assurer, l'IB s'appuie sur les contrôles prévus par le modèle d'assurance de la qualité.
- Les évaluations externes se veulent davantage sommatives que formatives ; c'est pourquoi l'IB demande aux examinateurs et examinatrices de ne rédiger des commentaires que s'ils expliquent leur notation.
- L'IB dispose également d'une procédure d'appel officielle que les établissements peuvent engager s'ils estiment que les processus n'ont pas été correctement suivis.

Pour consulter une description des aspects juridiques et des procédures des réclamations concernant les résultats que propose l'IB, veuillez vous référer aux *Procédures d'évaluation* spécifiques à chaque programme.

### Les catégories de réclamations concernant les résultats

Après la publication des résultats, le coordonnateur ou la coordonnatrice peut soumettre une demande pour l'un des éléments suivants.

- **Réclamation concernant les résultats de catégorie 1** : recorection de toutes les composantes de l'élève évaluées en externe pour une matière donnée
- **Rapport de recorection de catégorie 1** : rapport sur la notation réalisée dans le cadre d'une réclamation concernant les résultats de catégorie 1
- **Réclamation concernant les résultats de catégorie 2** : retour des copies des composantes évaluées en externe
- **Réclamation concernant les résultats de catégorie 3** : nouvelle révision de notation d'une composante évaluée en interne

Des frais sont facturés pour chacune des catégories ci-dessus (à l'exception des cas où une note finale est modifiée lors d'une recorection de catégorie 1).

### Pourquoi proposer des réclamations concernant les résultats

Le processus de réclamation concernant les résultats constitue le dernier rempart contre les erreurs dans le système de notation. Il permet aux établissements scolaires d'indiquer à l'IB s'ils estiment qu'une erreur a été commise, et à l'IB d'effectuer les vérifications nécessaires et d'apporter les corrections requises, au besoin.

## Le maintien des normes de la notation en conditions réelles à la recorection faisant suite aux réclamations concernant les résultats

Les réclamations concernant les résultats visent à attribuer aux élèves les notes qui reflètent la qualité de leur travail conformément aux normes de l'IB. Il est toutefois indispensable que les examinateurs et examinatrices ne se laissent pas influencer par l'insatisfaction de l'élève à l'égard de sa note initiale ou par le fait que cette note est proche d'un seuil d'attribution des notes finales.

Dans le processus de réclamation concernant les résultats, l'IB fait appel uniquement aux examinateurs superviseurs et examinatrices superviseuses montrant le plus de constance. Il est donc considéré que le résultat de cette recorection est la note correcte.

Pour aider les examinateurs et examinatrices à respecter la norme, l'IB glisse des réponses de contrôle dans les travaux visés par une réclamation concernant les résultats, pour pouvoir leur envoyer une alerte électronique en cas d'éloignement de la norme convenue.

Pour garantir que la nouvelle note est plus fiable que la note originale, le personnel chargé de l'évaluation de l'IB vérifie tous les changements de notes proposés à l'issue d'une réclamation concernant les résultats. Ainsi, dès qu'une différence de note considérable et inexplicquée est observée entre les deux corrections, et que cela entraîne une modification de la note finale, un troisième avis sur la copie est demandé.

## La réclamation concernant les résultats de catégorie 1 : une recorection

Comme cela est expliqué dans la section consacrée à la [notation](#), il est possible que deux personnes attribuent des notes différentes à une même question tant que cette différence reste acceptable (marge de tolérance). Par conséquent, si le processus de réclamation concernant les résultats aboutit à une modification mineure du nombre de points, il est probable qu'aucune erreur n'ait été commise lors de la notation. Et c'est le cas dans la plupart des demandes de réclamation concernant les résultats reçues par l'IB.

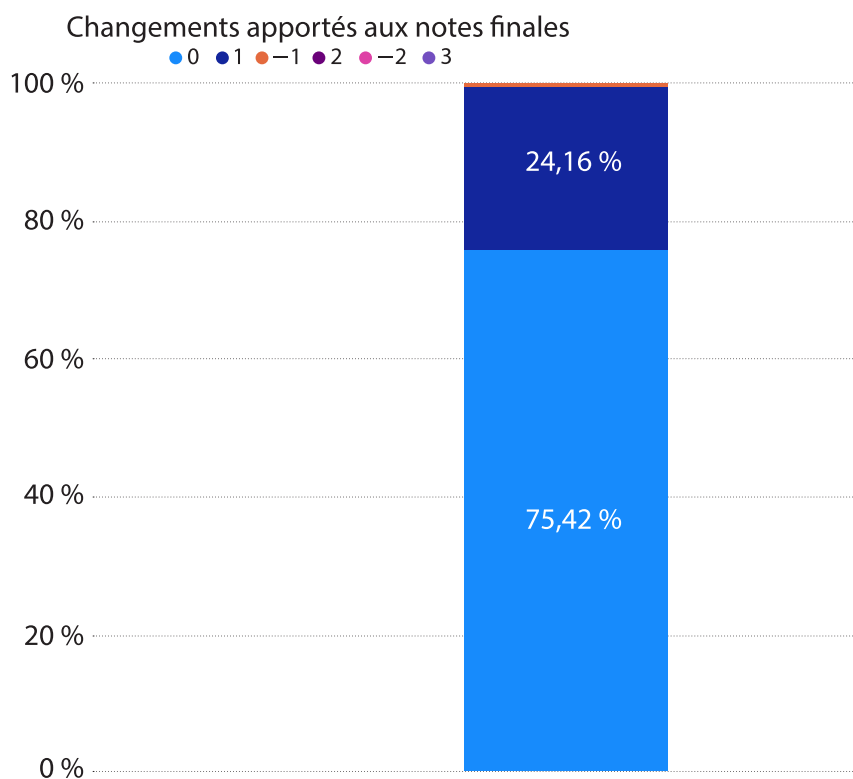
Pour réduire les effets de ces différences mineures, mais acceptables, de notation entre les examinateurs ou examinatrices, l'IB recorectre toutes les évaluations externes de l'élève lors d'une réclamation de catégorie 1. Cela limite en effet l'impact d'un changement mineur, car, en supposant que le changement ne résulte pas de problèmes récurrents, l'IB s'attend à ce qu'il s'annule sur plusieurs examens.

L'IB considère que la note décidée à l'issue d'une réclamation concernant les résultats est le reflet le plus fidèle du travail de l'élève ; c'est pourquoi elle constitue la note attribuée. Cela implique que le résultat de l'élève peut diminuer ou au contraire augmenter du fait d'une réclamation de catégorie 1. Les établissements doivent donc impérativement obtenir l'autorisation de l'élève avant d'envoyer une réclamation de ce type.

Étant donné que le système repose sur des seuils d'attribution des notes finales, il est toujours possible que des élèves se trouvent un point au-dessous d'un seuil ou au contraire juste au-dessus. Dans de tels cas, une réclamation concernant les résultats pourrait aboutir à un changement de note finale malgré l'absence d'erreur dans la notation. L'IB soutiendrait alors que l'élève se trouve à la limite de deux notes finales et que toutes deux sont un reflet fidèle de son accomplissement.

Figure 37

Modifications moyennes des notes finales décidées à l'issue d'une réclamation concernant les résultats de catégorie 1 sur cinq sessions d'examens du Programme du diplôme



## La réclamation concernant les résultats de catégorie 2 : le retour des copies des élèves

Les évaluations prises en considération pour une réclamation concernant les résultats de catégorie 2 sont les évaluations sommatives. Une copie renvoyée peut contenir des commentaires utiles de l'examinateur ou examinatrice à qui la copie a été affectée. Cependant, l'ajout de ces commentaires ne peut être garanti puisque leur rédaction n'est pas obligatoire dans le cadre du travail de notation des travaux d'élèves.

L'IB attend une notation appropriée de la part des examinateurs et examinatrices, pour garantir que les élèves obtiennent des résultats justes qui soient le reflet fidèle des travaux envoyés. Il leur est donc demandé uniquement d'inclure des commentaires pour expliquer leur notation, et non de faire un retour d'information aux élèves pour leur indiquer comment améliorer leur travail.

Pour savoir quoi faire si vous estimez que la notation d'un travail n'a pas été équitable, veuillez vous référer à la vidéo intitulée « [Contester des résultats : que faire si vous pensez que votre travail n'a pas été noté de manière équitable ?](#) ».

## La réclamation concernant les résultats de catégorie 3 : une nouvelle révision de notation

Pour ce qui est des réclamations de catégorie 3 (nouvelle révision de notation), tout changement apporté aux notes s'applique à la totalité de la cohorte. Si l'IB appliquait le principe de l'obtention de l'autorisation des élèves, cela signifierait qu'en ne donnant pas son autorisation ou en ne répondant pas à la demande, un seul individu pourrait empêcher toute la cohorte de voir son travail être de nouveau révisé. Cela ne serait évidemment pas acceptable. Par conséquent, pour que les notes finales ne soient pas diminuées sans que

les élèves aient officiellement accepté ce risque, les notes finales peuvent **uniquement être augmentées** (et jamais diminuées) à l'issue d'une réclamation de catégorie 3.

Une nouvelle révision de notation aura lieu uniquement dans les cas où l'écart entre la moyenne des notes révisées de l'évaluation interne des élèves et la moyenne de leurs notes brutes (les notes attribuées par le personnel enseignant) est **supérieur ou égal à 15 %** du nombre de points maximum pour la composante.

Ce seuil est en place parce que si l'IB n'impose aucune restriction, certains établissements demanderont systématiquement de nouvelles révisions de notation dans toutes les matières qu'ils proposent, dans l'espoir d'obtenir l'augmentation de certaines notes finales. Le nombre de réviseurs de notation superviseurs et de réviseuses de notation superviseuses de l'IB étant limité, il ne permettrait pas de satisfaire à une telle demande. Ce seuil permet de s'assurer que seuls les établissements qui ont clairement une source de préoccupation peuvent avoir recours à ce service. L'IB peut ainsi gérer la charge de travail de ses effectifs peu nombreux tout en proposant aux établissements une nouvelle révision de notation lorsque la sévérité des effets de la révision de notation sur les notes des élèves le justifie.

Figure 38

Une cohorte répondant aux critères pour demander une réclamation concernant les résultats de catégorie 3

| Numéro de l'élève                                   | Inclusion dans l'échantillon | Note brute | Note révisée | Ajustement de révision de notation |
|---|------------------------------|------------|--------------|------------------------------------|
| 2   | Oui                          | 19         | 14           | -5                                 |
| 10  |                              | 15         | 12           | -3                                 |
| 11  | Oui                          | 16         | 13           | -3                                 |
| 12  |                              | 20         | 15           | -5                                 |
| 13  |                              | 18         | 14           | -4                                 |
| 16  | Oui                          | 12         | 11           | -1                                 |
| 36  | Oui                          | 17         | 13           | -4                                 |
| 45  | Oui                          | 10         | 10           | 0                                  |
| Ajustement moyen en points                          |                              |            |              | -3,125                             |
| Ajustement moyen en pourcentage de la note maximale |                              |            |              | -15,63 %                           |

Dans l'exemple présenté dans la figure 38, le nombre de points maximum pouvant être obtenu pour la composante est de 20.

L'ajustement de révision de notation moyen est de -3,125 points, ce qui équivaut à 15,63 % du nombre de points maximum pouvant être obtenus.

L'échantillonnage répond donc aux critères pour une nouvelle révision de notation de catégorie 3.

## Les demandes d'appel de l'évaluation

L'appel donne la possibilité de demander à l'IB de revoir une décision prise dans le cadre du processus d'évaluation, c'est-à-dire l'octroi d'une note finale de l'IB ou le traitement d'un cas de mauvaise conduite.

Un appel ne peut pas servir à demander qu'une exception soit faite à une procédure de l'IB, mais peut remettre en question le caractère raisonnable de l'interprétation des règlements de l'IB. Un appel ne peut pas entraîner la modification des politiques de l'IB publiées et, lorsqu'il génère une clarification des règlements de l'IB, cette clarification doit être appliquée à l'ensemble des élèves de la session concernée. L'appel constitue le dernier recours pour traiter toute préoccupation concernant le processus d'évaluation de l'IB.

Pour en savoir plus, veuillez consulter les *Procédures d'évaluation* du PEI, du Programme du diplôme ou du POP, qui sont disponibles sur le Centre de ressources pédagogiques.

## Les élèves de reprise

Si sa note ne lui convient pas, l'élève peut repasser les examens dans la matière en question. Cela peut se faire à la session d'examens suivante qui a lieu six mois plus tard ou, en principe, au cours de n'importe quelle session ultérieure. Cependant, lorsqu'une matière est repassée, si le programme d'études ou les modalités d'évaluation ont été remaniés de manière importante, l'élève doit satisfaire à ces nouvelles exigences. Les élèves qui souhaitent repasser une ou plusieurs matières n'ont pas l'obligation de se réinscrire dans l'établissement où les matières en question ont été initialement passées. Si un établissement accepte des élèves qui souhaitent repasser une ou plusieurs matières, il doit accepter toutes les responsabilités pédagogiques et administratives les concernant.

Les élèves qui repassent une matière peuvent conserver les résultats obtenus pour une composante ne faisant pas l'objet d'un examen, à condition que le programme d'études ou les modalités d'évaluation n'aient pas changé de manière importante. La note obtenue à un examen écrit ne pouvant être conservée, tous les examens de la matière à repasser devront être repassés.

Pour soumettre à nouveau des travaux pour une composante ne faisant pas l'objet d'un examen, les élèves doivent suivre les cours donnés dans l'établissement où leur inscription a été effectuée pour la session de reprise. Cela résulte du fait que le personnel enseignant de la matière concernée doit les guider dans leur travail, noter le travail effectué pour l'évaluation interne et confirmer l'authenticité des travaux. Les élèves ne doivent pas se contenter d'apporter des changements mineurs au travail ne faisant pas l'objet d'un examen précédemment soumis. En principe, un travail complètement différent doit être remis. L'IB reconnaît toutefois que cela peut ne pas être possible pour certaines tâches requérant un travail considérable, en raison des contraintes de temps. Dans de tels cas, les élèves peuvent apporter des modifications importantes au travail original. Si un travail modifié est remis pour l'évaluation, il sera noté comme un nouveau travail et les élèves doivent accepter le fait que leur travail est susceptible d'obtenir une note plus basse. Dans le cas d'un travail modifié, les élèves doivent lui avoir consacré un temps supplémentaire équivalent à au moins 20 % du nombre total d'heures initialement passées dessus. En outre, le travail doit répondre à au moins un des critères énumérés ci-après. Le travail doit présenter :

- une modification de la question, de l'axe ou du thème de recherche ;
- un nouveau thème, une nouvelle preuve ou un nouveau point de vue ayant une incidence sur la conclusion ;
- une analyse supplémentaire importante ou des calculs supplémentaires importants ;
- une nouvelle exploration d'un aspect du travail de l'auteur, autrice ou artiste.

Si l'élève a réalisé un travail qui n'a pas été envoyé dans le cadre de son évaluation (par exemple, un portfolio d'Arts du Programme du diplôme), le temps déjà consacré à ce travail jamais vu auparavant peut être pris en compte pour atteindre l'objectif de 20 %.

Cette approche a été retenue pour assurer un équilibre entre l'obligation de ne pas faire profiter les élèves de reprise d'un nouveau cycle de retour d'information sur leur travail (pour ne pas désavantager les élèves qui présentent un nouveau travail) et la nécessité de reconnaître qu'il ne serait pas équitable de faire peser une charge excessive sur les élèves ou l'établissement scolaire pour obtenir, par exemple, des données entièrement nouvelles ou une recherche inédite qui viendraient se dresser en obstacle à la reprise.

L'approche de la comparabilité adoptée par l'IB est précisée dans la section intitulée « [La comparabilité](#) ».

L'IB a conscience de la nécessité de modifier le contenu pédagogique au fil du temps. Avant de lancer un nouveau cours, l'IB en informe les établissements scolaires par divers moyens, notamment à travers les actualités publiées pour le programme concerné et la production d'un nouveau guide pédagogique. Après la dernière session d'examens prévue pour un cours donné, l'IB n'est pas en mesure de proposer aux élèves de reprise de repasser les examens portant sur ce cours. Cela s'explique essentiellement par le fait qu'il serait ingérable à la fois pour les établissements scolaires et pour l'IB de faire porter l'évaluation d'une même matière sur deux cours différents lors d'une même session d'examens. Il peut toutefois être accepté

que les élèves de reprise conservent les notes obtenues à l'évaluation interne ou aux travaux évalués en externe. Mais si le nouveau cours est radicalement différent de l'ancien, il est probable que cette possibilité ne leur soit pas offerte.

Pour en savoir plus sur les élèves de reprise, veuillez consulter les *Procédures d'évaluation* propres à chaque programme.

## Les commentaires destinés aux établissements

- L'évaluation sommative de l'IB a pour objectif de mesurer les accomplissements des élèves, et tous les processus s'y rapportant sont pensés de manière à rendre les résultats les plus valides possibles.
- L'IB tient à rendre totalement transparents les résultats de l'évaluation communiqués aux établissements scolaires et aux élèves, pour que tant le personnel enseignant que les élèves puissent constater que les notes et les notes finales ont été attribuées avec la même constance et sans biais.
- Cette transparence est particulièrement importante pour l'évaluation interne, car son résultat est issu d'une révision de la notation de l'enseignant ou enseignante.
- L'évaluation sommative de l'IB, et notamment le processus de réclamation concernant les résultats, ne vise pas à indiquer aux établissements comment améliorer les résultats de leurs élèves, à moins que cela ne soit une conséquence indirecte des données requises pour une évaluation valide.
- En dehors du processus d'évaluation, les établissements peuvent bénéficier d'autres services de l'IB en matière de méthode d'enseignement et de perfectionnement professionnel.

L'évaluation sommative de l'IB a pour objectif de mesurer les accomplissements des élèves conformément aux programmes d'études suivis et de fournir un résultat de mesure fiable, sous forme de note finale. Elle se distingue de l'évaluation formative qui vise, elle, à fournir un retour d'information pour apporter un soutien à l'élève dans son apprentissage. Bien que le retour d'information constructif soit le pivot de l'évaluation formative et de l'enseignement efficace, il n'est pas adapté de se servir des résultats des évaluations sommatives à des fins formatives. Cela s'explique notamment par le fait que l'examineur ou examinatrice de l'IB ne peut émettre un jugement qu'à partir du travail mis à sa disposition (la copie d'examen), alors qu'en classe, les enseignantes et enseignants de l'IB pourront se baser sur tout un éventail d'informations et un point de vue personnel pour décider du retour d'information à faire à l'élève. De la même manière, les examinateurs et examinatrices ne fournissent pas de retour d'information écrit aux élèves ni aux enseignantes et enseignants lors de la notation du travail.

Les examinateurs et examinatrices doivent indiquer clairement à quels endroits les points ont été attribués et expliquer ce choix en ajoutant des commentaires appropriés, en cas d'ambiguïté. Cela permet à l'IB de vérifier les normes de notation et rend le processus d'attribution des points plus transparent pour les établissements.

## Les rapports pédagogiques

Après chaque session d'examens, l'IB produit un rapport pédagogique, qui est généralement le fruit du travail des examinateurs et examinatrices en chef. Ce rapport fournit des informations sur chaque composante d'évaluation et sur la réussite des élèves aux examens, en précisant les questions ou les thèmes particulièrement bien ou mal abordés. Il contient également des recommandations d'ordre général faites par les équipes d'examineurs et examinatrices pour expliquer au personnel enseignant comment préparer les élèves à des questions ou thèmes similaires pour les prochaines sessions. Il comprend aussi les seuils d'attribution des notes finales de la matière et les seuils d'attribution des notes finales de chaque composante.

Les enseignantes et enseignants devront évidemment replacer les informations ainsi rassemblées dans le contexte spécifique de leur classe. Par exemple, même si une question donnée a globalement produit des réponses médiocres, il est tout fait possible que l'ensemble des élèves de leur classe ait obtenu des notes élevées.

## Les commentaires sur l'évaluation interne

Les notes attribuées par les enseignantes et enseignants à l'évaluation interne sont soumises à un processus de révision de notation. Il est possible que le personnel enseignant interprète les critères d'évaluation interne différemment de l'examineur ou examinatrice de l'IB, ce qui peut s'illustrer par une correction plus sévère du travail et par l'attribution d'un nombre de points moins élevé. La révision de notation de l'évaluation interne vise à ajuster la norme de notation des enseignantes et enseignants pour la mettre en conformité avec la norme de l'examineur principal ou examinatrice principale ou avec la norme mondiale. Pour que les enseignantes et enseignants puissent comprendre leur position par rapport à la norme mondiale, l'IB fournit un retour d'information sur la notation des tâches d'évaluation interne, en cas de notation en dehors de la marge de tolérance. Cela permet au personnel enseignant de comprendre les raisons de l'application d'un facteur de révision de notation.

Le retour d'information qui est fait aux établissements ne vise pas à expliquer comment les élèves de l'échantillonnage auraient pu obtenir de meilleurs résultats. Il apporte plutôt des renseignements sur la qualité de la notation pour chaque critère ainsi que sur l'application générale des critères d'évaluation et l'adéquation des travaux remis.

## Remarques sur les manuels, les ateliers et les examens

Le programme d'études de chaque cours de l'IB est détaillé dans le guide pédagogique correspondant et sert de base pour la conception des évaluations. Tous les manuels, y compris ceux approuvés par l'IB, visent à aider les élèves et le personnel enseignant à étudier le cours conformément à la description qui en est faite dans le guide pédagogique, et non à définir le contenu du programme d'études.

Par conséquent, les questions d'examen pourront porter sur un thème ou une partie d'un thème qui figure dans le guide pédagogique, mais pas dans un manuel donné. L'ensemble du personnel enseignant est donc encouragé à se référer uniquement aux guides pédagogiques pour tout ce qui concerne le contenu du programme d'études. Dans le cadre du processus de rédaction des examens, l'IB passe fréquemment en revue les manuels, afin de vérifier que les questions d'examen envisagées ne figurent pas déjà dans ces ressources.

De même, les commentaires et les documents fournis dans les ateliers animés par l'IB ne se substituent pas aux énoncés figurant dans les guides pédagogiques, même s'ils peuvent aider les enseignantes et enseignants à mieux interpréter les formulations trouvées dans les guides. Toutes les modifications apportées aux guides pédagogiques seront officiellement publiées par l'IB et clairement indiquées comme telles.

## Définition des processus spécifiques aux programmes

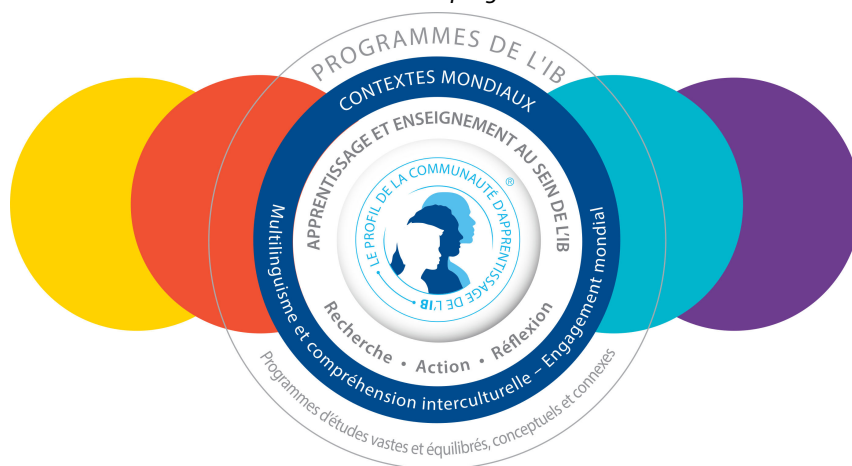
- Étant donné que tous les programmes du Baccalauréat International (IB) reposent sur la même philosophie éducationnelle, toutes les évaluations de l'organisation doivent suivre les mêmes principes et adopter les mêmes pratiques générales pour respecter cette philosophie. Les objectifs de l'évaluation évolueront naturellement avec la progression des élèves à travers les différents programmes et en fonction des stades de maturité scolaire correspondants.
- La grande idée derrière les évaluations sommatives externes des programmes de l'IB est de doter les élèves d'un récapitulatif valide de leurs résultats scolaires qui les aidera à poursuivre leurs études ou à intégrer le monde du travail.
- Même si l'évaluation est un élément important du Programme primaire (PP), ce programme ne se prête pas à une évaluation sommative externe.
- Chaque programme a des particularités qui imposent des processus spécifiques s'inscrivant dans les pratiques de l'évaluation.

## Le socle commun à tous les programmes

- Les programmes de l'IB proposent des programmes d'études ou des cadres pédagogiques qui se veulent vastes, équilibrés, conceptuels et connexes.
- Toutes les évaluations de l'IB doivent être conçues en tenant compte de ces caractéristiques sous-jacentes du système éducatif de l'IB, quand bien même elles ne les évaluent pas explicitement, afin qu'il y ait un effet de remous positif sur l'apprentissage et l'enseignement.

Figure 39

La continuité entre les programmes de l'IB



Le document intitulé *En quoi consiste le système éducatif de l'IB ?* (2019) met l'accent sur le fait que les programmes de l'IB favorisent l'apprentissage conceptuel, en se concentrant sur des idées structurantes et marquantes, qui sont pertinentes dans toutes les disciplines et qui aident à assimiler ce qui a été appris et à rendre le programme d'études plus cohérent. Les programmes de l'IB soulignent l'importance d'établir des liens, d'explorer les relations entre les matières scolaires, et d'apprendre des choses sur le monde, et ce, en allant au-delà du cadre de chaque matière. Ils permettent aux élèves d'accéder à un éventail de sujets d'étude et d'expériences d'apprentissage qui se veulent vastes, équilibrés, conceptuels et connexes.

- **Vaste et équilibré** : le système éducatif de l'IB offre une approche équilibrée qui permet aux élèves d'accéder à une large gamme de contenus dans de nombreuses matières.
- **Conceptuel** : l'apprentissage conceptuel se concentre sur de grandes idées structurantes et marquantes, qui sont pertinentes dans toutes les disciplines. Ces idées transcendent les frontières nationales et les barrières culturelles. Les concepts aident à assimiler l'apprentissage, à donner une plus grande cohérence au programme d'études, à approfondir la compréhension propre à la discipline et à renforcer la capacité à examiner des idées complexes. De plus, ils permettent d'utiliser ce qui a été appris dans de nouveaux contextes.
- **Connexe** : les cadres pédagogiques de l'IB accordent de l'importance à la simultanéité des apprentissages. Les élèves étudient de nombreuses matières simultanément durant leur parcours scolaire, apprennent à établir des liens et acquièrent une compréhension solide des corrélations entre la connaissance et l'expérience dans nombre de domaines d'étude. Les objectifs globaux des cours et les exigences spécifiques aux programmes permettent véritablement d'acquérir des connaissances sur le monde en allant au-delà des limites traditionnelles de chaque matière.

Lors de la conception des évaluations, le personnel enseignant, les établissements et l'équipe de rédaction de l'IB doivent réfléchir à ces objectifs sous-jacents afin d'éviter de créer des tâches faisant obstacle à de bons apprentissage et enseignement. L'objectif doit systématiquement être de produire un effet de remous positif.

## La dimension internationale

La mission de l'IB est d'aider les élèves à développer « la sensibilité nécessaire pour contribuer à bâtir un monde meilleur et plus paisible, dans un esprit d'entente mutuelle et de respect interculturel » et à « comprendre que les autres, dans leurs différences, puissent aussi être dans le vrai » (extraits de la déclaration de mission de l'IB, 2024). Les programmes de l'IB sont également étudiés par des élèves vivant dans de nombreux pays et régions (ou territoires) représentant toute une palette de nationalités. L'apprentissage et l'enseignement prônés par l'IB se caractérisent donc à la fois par un contexte international et par un accent mis sur le développement de la sensibilité internationale, deux éléments qui doivent se retrouver dans l'évaluation. Pour en savoir plus sur l'approche adoptée par l'IB en matière de sensibilité internationale, veuillez consulter la section « [La sensibilité internationale et la compréhension interculturelle](#) ».



## Des besoins et des solutions spécifiques aux programmes

Dans les sections précédentes, les processus qui constituent le cycle d'évaluation ont été expliqués. Ces pratiques généralisées valent pour tous les programmes de l'IB. Chaque programme poursuit cependant un objectif particulier, ce qui explique que des différences puissent être observées au niveau de certains aspects de l'évaluation de chaque programme ainsi que des règles permettant de déterminer si un diplôme ou un certificat peut être octroyé pour un programme.

Ces approches de l'évaluation, qui sont spécifiques à chaque programme, sont importantes. Cette section permettra d'expliquer en quoi elles consistent et comment elles sont gérées.

## Le Programme du diplôme

Les caractéristiques de l'évaluation sommative externe du Programme du diplôme sont les suivantes.

- Les élèves doivent étudier un nombre prescrit de matières pour obtenir le diplôme de l'IB.
- Le résultat global au diplôme de l'IB est décrit à l'aide d'un total de points, le score maximal étant de 45.
- Les composantes du tronc commun peuvent contribuer à hauteur de 3 points au résultat global au diplôme de l'IB, selon une matrice des points.
- Presque toutes les matières comportent plusieurs composantes d'évaluation, qui combinent évaluation externe et évaluation interne.
- Presque toutes les matières sont proposées au niveau moyen (NM) et au niveau supérieur (NS), et contribuent à part égale au résultat global au diplôme.

## Les objectifs globaux du Programme du diplôme

Pour être valides, les résultats de l'évaluation doivent être clairement reliés aux objectifs globaux des cours et du programme.

L'objectif premier du Programme du diplôme est de proposer aux élèves de 16 à 19 ans une expérience éducative vaste, équilibrée, stimulante et axée sur la réalité internationale. Dans ce but, les élèves doivent étudier simultanément six matières et les composantes du tronc commun du programme, et ce, pendant deux ans. Le Programme du diplôme est également conçu de façon à leur apporter les compétences scolaires fondamentales qui leur seront nécessaires pour leurs études universitaires, la poursuite de leur éducation et la profession de leur choix, ainsi que de façon à leur permettre de développer les valeurs et les compétences importantes pour mener une vie pleine de sens et accomplie. En d'autres termes, la validité recherchée au niveau des résultats de l'évaluation du Programme du diplôme, des cours et des composantes du tronc commun découle de l'intérêt porté à l'expérience scolaire intégrale de chaque élève.

## Les utilisations valides des résultats des évaluations du Programme du diplôme

Lorsque l'IB élabore ses modèles d'évaluation et ses programmes d'études, il s'efforce de faire en sorte que les notes finales obtenues dans les cours du Programme du diplôme et le total des points obtenus au diplôme puissent être utilisés pour :

- faciliter la sélection en vue de l'entrée à l'université ou vérifier l'éligibilité à l'emploi ;
- déterminer si les élèves ont déjà satisfait à certaines exigences d'un programme universitaire, ce qui peut soit donner droit à des crédits en vue de l'obtention d'un certificat ou diplôme, soit justifier la dispense de certains cours ;
- fournir des preuves de l'éligibilité des élèves à poursuivre leurs études dans une langue donnée.

Pour comprendre les raisons pour lesquelles il est important de tenir compte de ces utilisations valides, veuillez vous reporter à la section « [La validité](#) ».

## La structure du Programme du diplôme et de ses cours

Pour obtenir le diplôme de l'IB, l'élève doit présenter six matières et satisfaire aux exigences des composantes du tronc commun, qui est composé de la théorie de la connaissance (TdC), du mémoire et du programme créativité, activité, service (CAS).

Certaines matières sont étudiées au niveau moyen (NM) et d'autres au niveau supérieur (NS). Trois matières au moins et quatre au plus doivent être présentées au NS, tandis que les matières restantes sont présentées au NM. Les cours du NM et du NS diffèrent par leur étendue, mais sont identiques pour ce qui est des objectifs d'évaluation et des descripteurs des notes finales spécifiques au groupe de matières. Les élèves doivent faire preuve d'un niveau de connaissance, de compétence et de compréhension plus élevé dans les matières étudiées au NS.

Un nombre limité de matières interdisciplinaires permet de satisfaire aux exigences de plusieurs groupes de matières. Par exemple, le cours de systèmes de l'environnement et sociétés permet de satisfaire l'obligation de suivre un cours du groupe de matières Individus et sociétés et un cours du groupe Sciences.

Figure 41

Le modèle du Programme du diplôme



## Le mode de calcul du résultat au Programme du diplôme

La validité du Programme du diplôme transparait dans le système de mesure des accomplissements des élèves. Le nombre total de points obtenus au Programme du diplôme est calculé en ajoutant les notes finales (allant de 1 à 7) attribuées dans chacune des six matières au nombre de points (entre 0 et 3) obtenus à chaque composante du tronc commun (TdC, mémoire et programme CAS). Pour calculer le nombre total

de points, une importance égale est accordée à toutes les matières, qu'elles soient étudiées au NM ou au NS. Le nombre de points maximum est 45.

## La matrice des points pour le tronc commun

À la différence des matières, l'échelle de notes va de A à E pour la TdC et le mémoire. En raison de sa nature particulière, la troisième composante, à savoir le programme CAS, ne peut être associée à aucune note finale.

Les accomplissements enregistrés dans le tronc commun peuvent rapporter entre 0 et 3 points qui seront ajoutés au nombre total de points obtenus au Programme du diplôme. Il est possible d'échouer au diplôme de l'IB soit en obtenant la note finale E en TdC ou au mémoire, soit en ne satisfaisant pas aux exigences relatives au programme CAS. Le nombre de points pouvant être attribués au tronc commun est calculé comme l'illustre la figure 42.

Figure 42

L'attribution des points pour la TdC et le mémoire

|            |                                   | Théorie de la connaissance (TdC) |   |   |   |                                   |
|------------|-----------------------------------|----------------------------------|---|---|---|-----------------------------------|
| Mémoire    | Note finale attribuée             | A                                | B | C | D | E ou « N »                        |
|            | A                                 | 3                                | 3 | 2 | 2 | Note éliminatoire pour le diplôme |
|            | B                                 | 3                                | 2 | 2 | 1 | Note éliminatoire pour le diplôme |
|            | C                                 | 2                                | 2 | 1 | 0 |                                   |
|            | D                                 | 2                                | 1 | 0 | 0 |                                   |
| E ou « N » | Note éliminatoire pour le diplôme |                                  |   |   |   |                                   |

## Les modalités d'octroi du diplôme de l'IB

Le diplôme de l'IB sera octroyé si les conditions suivantes sont réunies.

- Les exigences du programme CAS ont été satisfaites.
- Le total des points est supérieur ou égal à 24.
- La note « N » (indiquant l'absence d'une note finale) n'a pas été attribuée pour la TdC, le mémoire ni l'une des matières concourant au diplôme de l'IB.
- Une note finale minimale de D a été attribuée pour la TdC et pour le mémoire.
- Une note finale de 1 n'a été attribuée dans aucune matière ni à aucun niveau.
- Une note finale de 2 n'a pas été attribuée dans trois matières ou plus (au NS ou au NM).
- Une note finale égale ou inférieure à 3 n'a pas été attribuée dans quatre matières ou plus (au NS ou au NM).
- L'élève a obtenu 12 points ou plus pour les matières au NS (seules les trois notes finales les plus élevées comptent en cas d'inscription à quatre matières au NS).
- L'élève a obtenu 9 points ou plus pour les matières au NM (en cas d'inscription à deux matières au NM, il faut obtenir au moins 5 points au NM).

## Le diplôme bilingue de l'IB

À la place du diplôme de l'IB habituel, un diplôme bilingue peut être décerné à l'élève qui remplit **un ou plusieurs** critères suivants :

- présente deux langues sélectionnées dans le groupe de matières Études en langue et littérature, et obtient une note finale égale ou supérieure à 3 pour chacune de ces deux langues ;
- présente une matière du groupe Individus et sociétés ou du groupe Sciences dans une langue différente de celle choisie pour le groupe Études en langue et littérature. Une note finale égale ou supérieure à 3 doit alors être obtenue à la fois dans le cours de langue du groupe Études en langue et littérature et dans celui du groupe Individus et sociétés ou du groupe Sciences.

Les matières pilotes et les matières interdisciplinaires peuvent contribuer à l'obtention d'un diplôme bilingue, si les conditions ci-dessus sont remplies.

Les éléments suivants ne peuvent contribuer à l'obtention d'un diplôme bilingue :

- le mémoire ;
- un programme propre à l'établissement ;
- une matière choisie en complément des six matières du Programme du diplôme (option connue sous le nom de « matière supplémentaire »).

## Le maintien des normes d'une cohorte à l'autre

Les cohortes inscrites aux matières du Programme du diplôme peuvent aller de 1 à plus de 40 000 élèves. Il ne fait aucun doute que toutes les matières sont d'importance équivalente. Cependant, pour des raisons pratiques, quelques disparités existent au niveau des processus s'appliquant aux matières, en fonction du nombre d'inscriptions.

### Le maintien de la qualité du travail des examinateurs et examinatrices

Pour les matières qui comptent un très petit nombre d'examineurs et examinatrices, il peut ne pas être approprié de suivre le [modèle d'assurance de la qualité](#) qui passe par des réponses d'entraînement, de validation et de contrôle, tel que décrit dans la section « [La normalisation](#) ».

Si l'ensemble des examinateurs et examinatrices participe à la discussion menée au moment de la normalisation, il est superflu d'utiliser des réponses d'entraînement ou de validation, car chaque personne aura déjà participé à la définition de la norme. Le recours aux réponses de contrôle conserve son utilité pour s'assurer que la notation reste conforme à la norme requise. Même s'il ne s'avère finalement pas utile, un modèle d'assurance de la qualité est cependant généralement prévu.

Il importe avant tout de maintenir la qualité de la notation, même si le nombre d'inscriptions à une matière est faible.

### Les notes finales : le processus complet d'attribution des notes finales (en ligne ou en présentiel)

Pour les matières qui comptent plusieurs examinateurs et examinatrices en dessous de l'examineur en chef ou principal ou de l'examinatrice en chef ou principale dans la hiérarchie, une réunion officielle est organisée pour appliquer les pratiques expliquées dans la section « [L'attribution des notes finales et la totalisation](#) ». Un ou une responsable de matière de l'IB participe à la réunion pour aider les examinateurs et examinatrices et faire un contrôle de la qualité du processus. Cette réunion peut se dérouler en présentiel ou en ligne.

### Les notes finales : le processus dirigé d'attribution des notes finales pour un petit nombre d'élèves (en ligne)

Lorsqu'il existe un nombre très réduit d'examineurs et examinatrices pour chaque composante, les examinateurs superviseurs et examinatrices superviseuses discutent en ligne. La ou le responsable de matière de l'IB leur fournit tout l'éventail de preuves nécessaires pour décider convenablement des notes finales à attribuer et contrôle le processus pour veiller au maintien de la qualité.

### Les notes finales : le processus classique d'attribution des notes finales pour un petit nombre d'élèves

Lorsque le nombre d'élèves est trop faible pour que les preuves statistiques accumulées au fil des années soient exploitables, il est demandé aux examinateurs et examinatrices de passer en revue tous les travaux envoyés pour l'évaluation plutôt que de se concentrer sur un échantillonnage pour recommander les seuils d'attribution des notes finales fondés sur une appréciation.

Des responsables de matière de l'IB viennent apporter leur soutien à ces réunions si nécessaire. Leur rôle n'est pas d'observer ces réunions, mais d'effectuer des contrôles aléatoires de la qualité.

Pour faciliter le travail des examinateurs et examinatrices des matières présentées par un petit nombre d'élèves (processus dirigé ou classique d'attribution des notes finales), il est conseillé d'engager des discussions entre les examinateurs et examinatrices des différentes matières d'un même groupe. Cela permet de s'assurer d'une compréhension commune de la signification des notes dans tout le groupe de matières.

### **La validation finale**

Quel que soit le nombre d'élèves ayant passé les examens, la dernière étape du processus d'attribution des notes finales est la suivante : soit l'examineur ou examinatrice en chef soit l'examineur principal ou examinatrice principale fait des recommandations aux membres de la direction de la division de l'évaluation de l'IB, qui passent ensuite en revue les preuves à l'appui de ces recommandations et confirment qu'elles leur conviennent.

Les matières font toutes l'objet du même examen attentif, indépendamment du nombre d'élèves qui composent la cohorte.

## Lectures complémentaires

Pour obtenir de plus amples informations sur le Programme du diplôme, veuillez consulter les ressources qui suivent.

- [\*Le Programme du diplôme : des principes à la pratique\*](#)
- [\*Procédures d'évaluation du Programme du diplôme\*](#)
- [\*Règlement pour les écoles du monde de l'IB\*](#)
- Les guides pédagogiques (voir les pages consacrées aux matières sur le Centre de ressources pédagogiques)
- Le matériel de soutien pédagogique publié pour chaque cours (voir les pages consacrées aux matières sur le Centre de ressources pédagogiques)
- [\*L'enseignement et l'apprentissage guidés par l'évaluation dans le Programme du diplôme\*](#)

## Le Programme à orientation professionnelle

Les caractéristiques de l'évaluation du Programme à orientation professionnelle (POP) sont les suivantes.

- Les élèves doivent satisfaire à un certain nombre d'exigences pour obtenir le certificat du POP de l'IB.
- Aucun résultat général sous forme de points n'est associé au certificat du POP de l'IB.
- Lorsque des élèves du POP suivent des cours avec des élèves du Programme du diplôme, une évaluation commune a lieu.
- Le cadre pédagogique du POP impose aux élèves de suivre une formation à orientation professionnelle qui n'est pas proposée par l'IB et ne fait pas l'objet d'une sanction officielle des études par l'organisation.

### Les objectifs globaux du POP

Pour être valides, les résultats de l'évaluation doivent être clairement reliés aux objectifs globaux des cours et du programme.

La particularité du POP est qu'il aide les élèves à se préparer à la vie active. Ce programme leur offre la garantie de développer des compétences transférables qui leur seront utiles tout au long de la vie, dans la suite de leurs études ou dans l'emploi pour lequel il leur semble bon de se former.

Le POP aide les élèves à :

- développer une vaste gamme de compétences et approfondir leur compréhension dans des domaines spécifiques de la connaissance grâce à leurs cours du Programme du diplôme et leur formation à orientation professionnelle ;
- développer des stratégies flexibles d'acquisition ou d'enrichissement de connaissances et de compétences dans des contextes variés ;
- cultiver des attitudes et des habitudes qui leur permettent d'apprendre tout au long de leur vie et d'explorer différentes perspectives, possibilités et difficultés ;
- se préparer à intégrer efficacement un monde en constante évolution ;
- s'impliquer dans un apprentissage qui développe leur capacité et leur volonté d'innover de manière positive.

### Les utilisations valides des résultats des évaluations du POP

Lorsque l'IB élabore ses modèles d'évaluation et ses programmes d'études, il s'efforce de faire en sorte que les notes finales obtenues dans les cours du POP et le certificat du POP de l'IB puissent être utilisés pour :

- faciliter la sélection en vue de l'obtention d'un emploi et de l'entrée dans un programme professionnalisant comme un apprentissage ;
- faciliter la sélection en vue de la poursuite des études dans le domaine professionnel correspondant ;
- faciliter la sélection en vue de l'entrée à l'université ;
- déterminer si les élèves ont déjà satisfait aux exigences d'un programme universitaire (ce qui peut soit donner droit à des crédits, soit justifier la dispense de certains modules, cours, etc.).

Lorsqu'un ou une élève a passé des évaluations dans une langue (d'usage), cela fournit également des preuves de sa capacité à poursuivre ses études dans la matière ou une carrière dans cette langue.

Pour comprendre les raisons pour lesquelles il est important de tenir compte de ces utilisations valides, veuillez vous reporter à la section « La validité ».

## La structure du POP

Le POP est un cadre pédagogique composé de trois parties, à savoir :

- au moins deux cours du Programme du diplôme, étudiés au niveau moyen (NM) ou au niveau supérieur (NS) ;
- les composantes du tronc commun du POP ;
- une formation à orientation professionnelle.

Figure 43

Le modèle du Programme à orientation professionnelle



### Les composantes du tronc commun du POP

Les quatre composantes du tronc commun du POP sont reliées entre elles de par leur nature, leurs objectifs d'apprentissage et leurs liens intentionnels. Dans son ensemble, le tronc commun vise à :

- fonder le programme sur la mission de l'IB en renforçant les aptitudes du profil de la communauté d'apprentissage et la sensibilité internationale ;
- contextualiser et enrichir les cours du Programme du diplôme et la formation à orientation professionnelle, en réunissant tous les aspects du cadre pédagogique ;
- encourager le développement de connaissances, de compétences et d'attitudes personnelles, scolaires et professionnelles durables ;
- montrer que l'apprentissage est un processus itératif et corrélé, et constitue la pierre angulaire du bien-être continu des personnes et des communautés.

Il est obligatoire de mener à bien les quatre composantes du tronc commun du POP, à savoir :

- l'engagement auprès de la communauté ;
- l'étude des langues et des cultures ;

- le cours de compétences personnelles et professionnelles ;
- le projet de réflexion.

Le projet de réflexion est évalué en externe, ce qui signifie que le projet final et la réflexion sont notés et l'élève reçoit de l'IB une note finale pour cette composante. L'établissement doit confirmer à l'IB que les élèves ont mené à bien les trois autres composantes du tronc commun, mais l'IB ne vérifie aucune évaluation réalisée par l'établissement scolaire. Les procédures et pratiques utilisées par les établissements qui proposent le POP sont examinées durant les visites d'évaluation de la mise en œuvre du programme.

## Les cours du Programme du diplôme dans le cadre du POP

Chaque élève choisit au moins deux cours qui sont également proposés dans le Programme du diplôme. Le même processus d'évaluation s'applique aux élèves du POP et aux élèves du Programme du diplôme. Aucun examen ni aucun processus d'attribution des notes finales n'est spécifique aux élèves du POP.

Même si les cours sont communs aux deux programmes, les élèves ne peuvent pas s'inscrire à la fois au Programme du diplôme et au POP. La portée des autres exigences spécifiques à chacun des deux programmes exclut toute réalisation simultanée.

## La formation à orientation professionnelle

L'IB n'évalue pas les résultats de la formation à orientation professionnelle du POP, et il ne procède pas non plus à un quelconque contrôle de leur qualité. La seule exigence est que l'établissement confirme que l'élève a effectivement mené la formation à bien. Les cours du Programme du diplôme ne se prêtent pas à une intégration dans la formation à orientation professionnelle du POP.

## Le mode de calcul du résultat au POP

Aucun résultat général sous forme de points n'est associé au certificat du POP de l'IB.

L'élève se verra décerner le certificat du POP de l'IB, sous réserve du respect des modalités qui suivent.

- L'établissement a confirmé que l'élève a mené à bien la formation à orientation professionnelle spécifiée.
- L'élève a obtenu une note finale égale ou supérieure à 3 dans au moins deux cours du Programme du diplôme.
- L'élève a obtenu une note finale égale ou supérieure à D pour le projet de réflexion.
- L'établissement a confirmé que l'élève a satisfait à toutes les exigences relatives au cours d'étude des langues et des cultures, à l'engagement auprès de la communauté et au cours de compétences personnelles et professionnelles.
- L'élève n'a pas reçu de sanction pour mauvaise conduite de la part du comité d'attribution des notes finales.

Les notes finales et les résultats au POP sont confirmés par le même comité d'attribution des notes finales que pour le Programme du diplôme.

## Le certificat bilingue du POP

À la place du certificat du POP de l'IB, un certificat bilingue du POP peut être décerné à l'élève qui remplit **un ou plusieurs** critères suivants :

- présente deux langues sélectionnées dans le groupe de matières Études en langue et littérature du Programme du diplôme, et obtient une note finale égale ou supérieure à 3 pour chacune de ces deux langues ;
- présente une langue du groupe Études en langue et littérature du Programme du diplôme ainsi qu'un cours du groupe Individus et sociétés ou du groupe Sciences du Programme du diplôme dans une langue d'usage différente de celle choisie pour le groupe Études en langue et littérature. L'élève doit obtenir une note finale égale ou supérieure à 3 dans chacun de ces deux cours.

## La gestion des normes et les travaux manquants

Pour le projet de réflexion et tous les cours communs avec le Programme du diplôme, le POP suit les approches du Programme du diplôme, qu'il s'agisse de la procédure pour note manquante ou de la gestion des normes selon la taille de la cohorte.

Pour en savoir plus sur les notes ou travaux manquants, veuillez consulter la section « [Les notes manquantes](#) ».

## Lectures complémentaires

Pour obtenir de plus amples informations sur le POP, veuillez consulter les ressources qui suivent.

- *Le Programme à orientation professionnelle : des principes à la pratique*
- *Procédures d'évaluation du Programme à orientation professionnelle*
- *Aperçu du Programme à orientation professionnelle*
- *Guide du projet de réflexion*
- *Guide d'étude des langues et des cultures*
- *Guide du cours de compétences personnelles et professionnelles*
- *Guide de l'engagement auprès de la communauté*

## Le Programme d'éducation intermédiaire

Le certificat du PEI reconnaît que l'élève a réussi tous les aspects du Programme d'éducation intermédiaire (PEI). Pour se voir délivrer le certificat du PEI, l'élève doit :

- obtenir un total d'au moins 28 points et aucune note finale 1 ou 2 au cours de l'évaluation électronique du PEI ;
- suivre de préférence le programme pendant la période recommandée de deux ans (ou pendant un an minimum) et satisfaire aux exigences de la 5<sup>e</sup> année ;
- réaliser les évaluations internes et réussir les examens dans un minimum de cinq et un maximum de huit disciplines choisies dans les huit groupes de matières (parmi lesquels devront figurer les groupes imposés) ;
- présenter au moins un portfolio électronique dans le groupe de matières Arts, Éducation physique et à la santé ou Design ;
- passer l'examen interdisciplinaire sur ordinateur ;
- réaliser et remettre un projet personnel ;
- satisfaire aux exigences minimales de l'IB concernant l'engagement auprès de la communauté.

## Les objectifs globaux du PEI

Pour être valides, les résultats de l'évaluation doivent être clairement reliés aux objectifs globaux des cours et du programme.

Le PEI est conçu comme un cadre pédagogique cohérent et complet qui place les élèves de 11 à 16 ans face à des défis éducatifs et leur permet de développer des compétences importantes [à l'adolescence]. Ces années représentent une période critique dans leur développement, et leur réussite scolaire est liée de près à leur bien-être personnel, social et affectif. À une période où les élèves cherchent à se construire du point de vue de leur identité et de leur estime de soi, le PEI peut les motiver et les aider à réussir dans le contexte scolaire et au-delà. Ce programme leur permet de consolider leurs points forts et de relever des défis dans les matières pouvant leur poser des difficultés. Le PEI donne aux élèves des occasions de développer leur potentiel, de découvrir leurs préférences en matière d'apprentissage et de prendre des risques calculés. Il leur permet également de prendre pleinement conscience de leur identité personnelle et de mener une réflexion sur celle-ci.

*(Le Programme d'éducation intermédiaire : des principes à la pratique, 2014, p. 4)*

Dans le PEI, il existe un alignement explicite entre les **objectifs spécifiques des groupes de matières et les critères de notation**. Tous les groupes de matières du PEI ont quatre critères d'évaluation qui correspondent aux quatre objectifs spécifiques du groupe. Tous les critères contribuent à hauteur équivalente au résultat final.

## L'inclusion d'un contexte mondial dans les évaluations électroniques

Dans le cadre du PEI, les contextes d'apprentissage doivent être (ou prendre modèle sur) de véritables événements, situations et circonstances qui surviennent dans le monde. Les contextes d'apprentissage du PEI sont choisis parmi les contextes mondiaux en vue d'encourager la sensibilité internationale et l'engagement mondial au sein du programme. [...] Le PEI identifie six contextes mondiaux pour l'enseignement et l'apprentissage qui s'inspirent des thèmes transdisciplinaires du PP et les renforcent.

*(Le Programme d'éducation intermédiaire : des principes à la pratique, 2014, p. 21 – 22)*

Chaque session d'examens sera élaborée et orientée en fonction d'un contexte mondial particulier et d'une piste d'exploration sélectionnés dans la liste publiée dans le document intitulé *Le Programme d'éducation intermédiaire : des principes à la pratique*.

Près d'un tiers des tâches de chaque examen sur ordinateur se rapportant à une discipline seront liées au contexte mondial sélectionné, s'en inspireront ou en découleront. La totalité de l'examen interdisciplinaire sur ordinateur s'inspirera du contexte mondial sélectionné.

Des plans de travail des unités partiellement remplis, qui tiennent compte du contexte mondial sélectionné, seront élaborés pour les groupes de matières Acquisition de langues, Arts, Design, et Éducation physique et à la santé.

## Les utilisations valides des résultats des évaluations du PEI

Lorsque l'IB élabore ses modèles d'évaluation et ses programmes d'études, il s'efforce de faire en sorte que les notes finales obtenues dans les cours du PEI et le total des points obtenus au certificat du PEI puissent être utilisés pour :

- faciliter la sélection en vue de la poursuite des études ou de l'entrée dans le monde du travail ;
- faire un retour d'information positif et un résumé des points forts pour les élèves poursuivant leurs études.

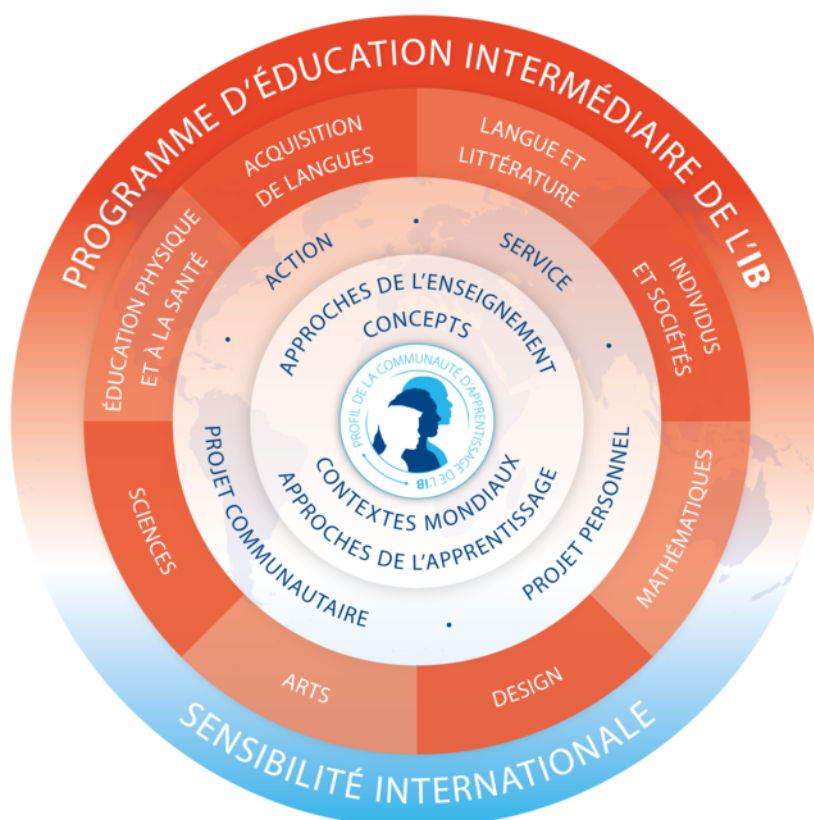
Lorsqu'un ou une élève a passé une évaluation dans une langue (d'usage), cela fournit des preuves de sa capacité à poursuivre ses études dans cette langue (par exemple, étudier dans un établissement d'enseignement anglais, espagnol ou français).

Pour comprendre les raisons pour lesquelles il est important de tenir compte de ces utilisations valides, veuillez vous reporter à la section « La validité ».

## La structure du PEI

Figure 44

Le modèle du Programme d'éducation intermédiaire



## Le mode de calcul du résultat au PEI

À la fin de la 5<sup>e</sup> année du PEI, il est possible d'inscrire les élèves aux évaluations externes de l'IB. Les résultats obtenus à ces évaluations seront consignés dans un document présentant les résultats de cours du PEI. Par ailleurs, ces élèves peuvent choisir de passer des évaluations qui conduiront à l'octroi du certificat du PEI.

Les établissements peuvent également délivrer une attestation de participation au PEI. Celle-ci s'adresse aux élèves qui ont suivi le programme pendant au moins deux ans et ont satisfait aux exigences de la 3<sup>e</sup> ou de la 4<sup>e</sup> année du PEI. Ces élèves n'ont pas d'inscription auprès de l'IB pour une quelconque forme d'évaluation. L'attestation de participation au PEI est un document propre à l'établissement qui n'est pas vérifié par l'IB.

Pour pouvoir obtenir le certificat du PEI, l'élève doit avoir suivi la dernière année du programme, après une participation recommandée de deux ans, et :

- passer une évaluation sur ordinateur ou réaliser un portfolio électronique dans six matières choisies dans les groupes suivants : Langue et littérature, Acquisition de langues (ou une deuxième matière du groupe Langue et littérature), Individus et sociétés, Mathématiques, Sciences et enfin, au choix, Arts, Éducation physique et à la santé ou Design ;
- obtenir une note finale égale ou supérieure à 3 dans chacune des six matières susmentionnées ;
- passer l'examen interdisciplinaire sur ordinateur et obtenir une note finale égale ou supérieure à 3 ;
- achever le projet personnel et obtenir une note finale égale ou supérieure à 3 ;
- obtenir un total général de 28 points ;
- satisfaire aux attentes de l'établissement scolaire concernant l'engagement auprès de la communauté.

Pour se voir délivrer un certificat bilingue du PEI, l'élève doit en outre obtenir des résultats satisfaisants aux examens sur ordinateur dans l'un des domaines suivants :

- un second cours de Langue et littérature (au lieu d'un cours d'Acquisition de langues) ;
- un cours (ou plus) du groupe de matières Sciences ou Individus et sociétés, ou l'examen interdisciplinaire, dans une langue autre que celle choisie pour le cours de Langue et littérature.

## Le déroulement de l'évaluation sommative externe (évaluation électronique du PEI)

L'évaluation électronique facultative du PEI prévoit deux façons d'évaluer les connaissances et les compétences des élèves :

- des portfolios électroniques contenant des travaux envoyés pour l'évaluation, réalisés dans le cadre des cours d'Acquisition de langues, d'Arts, de Design et d'Éducation physique et à la santé, qui font ensuite l'objet d'une révision de notation pour garantir une norme mondiale cohérente ;
- des examens sur ordinateur (d'une durée de deux heures chacun) portant sur les cours des groupes Langue et littérature, Acquisition de langues, Individus et sociétés, Sciences et Mathématiques, ainsi que sur l'apprentissage interdisciplinaire.

Par ailleurs, le projet personnel est remis à l'IB par voie électronique et fait l'objet d'une révision de notation. Contrairement aux autres éléments de l'évaluation électronique qui sont facultatifs, le projet personnel doit être réalisé par l'ensemble des élèves de 5<sup>e</sup> année du PEI.

### Les plans détaillés des examens

L'IB publie des plans détaillés des examens pour expliquer précisément aux établissements en quoi consisteront les évaluations électroniques. Ces plans détaillés permettent au corps enseignant et aux élèves de comprendre la nature et le but de l'évaluation électronique du PEI. Ils constituent une ressource importante pour aider les élèves à se préparer aux examens sur ordinateur, et à se concentrer sur les critères et les stratégies d'évaluation propres à chaque groupe de matières. Les plans détaillés comportent toujours quatre critères ayant la même pondération.

L'IB s'engage à veiller à ce qu'à chaque session, les examens ne s'éloignent pas de plus de 3 points de ces plans détaillés.

## **Les portfolios électroniques et les plans de travail des unités partiellement remplis**

Avec les portfolios électroniques, il devient possible d'évaluer les tâches (produits) des travaux réalisés sur une longue période dans le cadre des cours ou des représentations, chose qu'il est par nature difficile d'évaluer lors d'un examen. À la base des portfolios électroniques se trouvent des plans de travail des unités partiellement remplis qui sont suffisamment flexibles pour être adaptés aux contextes locaux et qui aident le personnel enseignant à s'assurer que le travail des élèves produit des preuves adéquates permettant d'émettre des jugements justes et pertinents. De nouveaux plans de travail des unités partiellement remplis sont fournis à chaque session.

Les plans de travail des unités devraient garantir que les tâches décidées par le personnel enseignant permettent aux élèves d'apporter des preuves dans l'ensemble de la gamme de notes finales du PEI. Les évaluations conçues par le personnel enseignant comportent le risque de désavantager les élèves avec un ensemble de tâches anormalement faciles ou difficiles. Lors de la révision de notation, l'IB ne peut attribuer des notes finales que sur la base des travaux envoyés pour l'évaluation disponibles. Si les tâches conçues par le personnel enseignant correspondent uniquement à une partie de la gamme des descripteurs des notes finales, aucune note finale ne pourra être attribuée en dehors de cette gamme.

## **L'évaluation unique et la gestion de la charge de travail pesant sur les élèves**

Pour les élèves, passer des examens et réaliser des travaux dans le cadre des cours est source de stress et de contraintes. Cela peut également empiéter sur le temps d'enseignement. Pour les élèves de 5<sup>e</sup> année qui suivent le PEI, l'IB estime que, tout bien considéré, il est plus adapté de minimiser la part d'évaluation sommative. Cela crée des difficultés, car les élèves n'ont qu'une seule occasion de montrer leurs capacités, mais l'IB les accepte afin de veiller au bien-être général des élèves.

## La méthode pour déterminer les notes finales manquantes

L'un des inconvénients de vouloir réduire la charge qui pèse sur les élèves est que, avec l'utilisation d'une seule composante d'évaluation, les preuves disponibles pour le PEI sont très minces. Cela signifie que la procédure pour l'attribution d'une note finale manquante est moins fiable aux yeux de l'IB que la procédure pour note manquante existant pour les autres programmes et donc qu'elle ne doit servir que dans des cas exceptionnels.

Pour en savoir plus sur le processus à suivre en cas de notes finales manquantes, veuillez consulter la section « [La procédure pour l'attribution d'une note finale manquante](#) ».

## Lectures complémentaires

Pour obtenir de plus amples informations sur le PEI, veuillez consulter les ressources qui suivent.

- [Le Programme d'éducation intermédiaire : des principes à la pratique](#)
- [Guide de la session d'examens du PEI](#)
- [Les guides pédagogiques du PEI](#)
- [Les guides des projets du PEI](#)
- [Procédures d'évaluation du Programme d'éducation intermédiaire](#)
- [Examens sur ordinateur du PEI – Exigences en matière de TI et responsabilités des établissements](#)
- [Familiarisation des élèves du PEI aux examens sur ordinateur](#)

## Le Programme primaire

Les caractéristiques de l'évaluation du Programme primaire (PP) sont les suivantes.

- L'évaluation nécessite une collaboration entre le personnel enseignant et les élèves en vue de contrôler l'apprentissage, de le consigner, de le mesurer, d'en rendre compte et de l'adapter.
- La promotion d'une culture de l'évaluation passe par le développement des capacités d'évaluation de chaque membre de la communauté d'apprentissage.
- L'IB n'impose aucune évaluation sommative externe.

### Les objectifs globaux du PP

Pour être valides, les résultats de l'évaluation doivent être clairement reliés aux objectifs globaux des cours et du programme.

Le cadre pédagogique du PP est centré sur l'apprentissage transdisciplinaire. Ce dernier organise le programme d'études et permet un apprentissage entre et à travers toutes les disciplines, et au-delà des limites traditionnelles des disciplines.

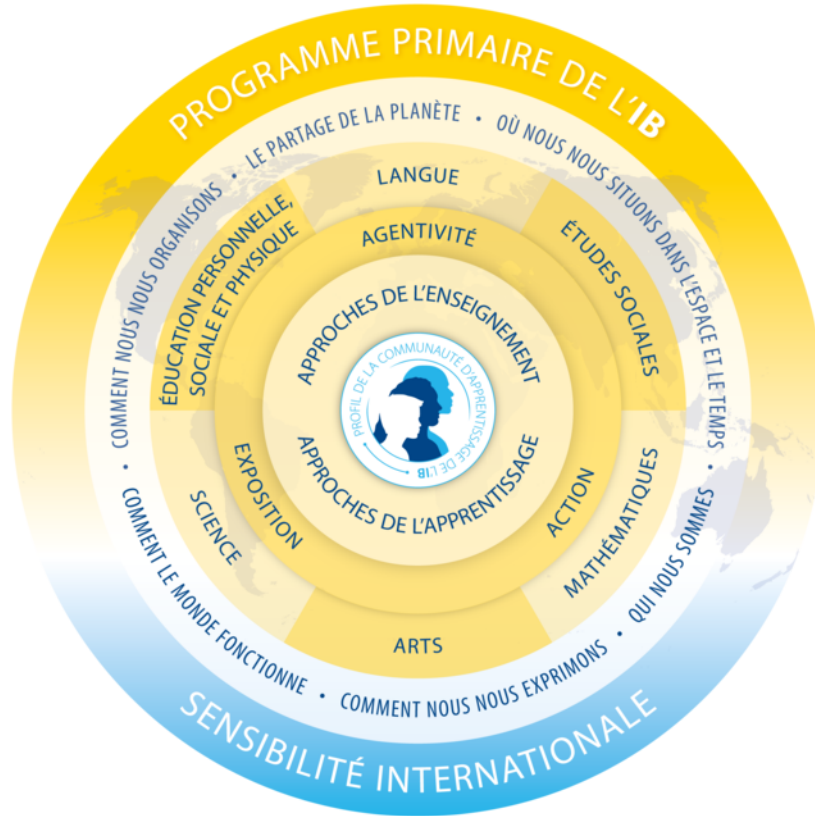
S'inspirant des théories de l'apprentissage constructivistes et socioconstructivistes, le cadre pédagogique du PP part du principe que les élèves sont au centre de leur apprentissage et des partenaires dans le processus d'apprentissage. Les élèves ont une capacité innée à faire des recherches, à questionner, à s'interroger et à formuler des théories au sujet de leur propre personne, des autres et du monde qui les entoure.

Cette compréhension de la façon dont les élèves apprennent est essentielle dans le modèle d'apprentissage et d'enseignement transdisciplinaire reposant sur la recherche et organisé autour de concepts moteurs. La participation au programme de recherche et la réflexion sur l'apprentissage permettent aux élèves du PP d'acquérir les connaissances, les compétences et les attitudes nécessaires pour s'impliquer dans le monde, en agissant en faveur du bien-être des personnes et de la planète.

## La structure du PP

Figure 45

Le modèle du Programme primaire



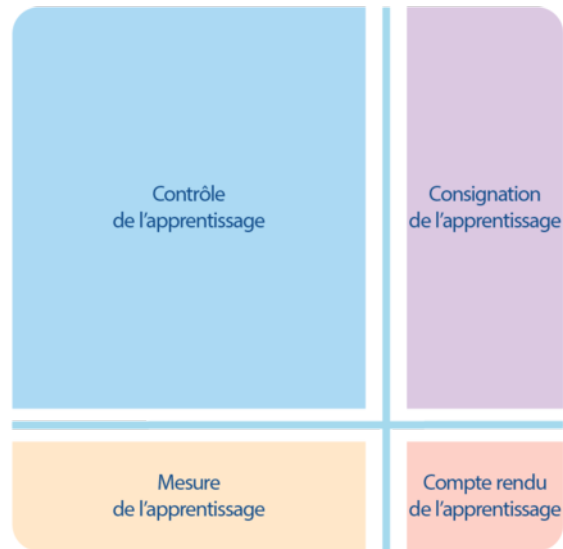
## L'évaluation au sein du PP

L'évaluation joue un rôle essentiel dans l'atteinte de l'objectif du PP, qui est de soutenir les élèves de façon réfléchie et efficace tout au long de l'acquisition des connaissances et des compétences spécifiques aux matières, de la compréhension conceptuelle ainsi que du développement des compétences spécifiques aux approches de l'apprentissage et des aptitudes du profil de la communauté d'apprentissage de l'IB.

Dans le PP, l'évaluation a quatre dimensions : contrôle, consignation, mesure et compte rendu de l'apprentissage. Chacun de ces aspects a une fonction distincte, mais tous visent à fournir des preuves afin de guider l'apprentissage et l'enseignement. Si ces quatre dimensions de l'évaluation n'ont pas le même poids, chacune a sa propre importance et sa propre valeur. Dans le PP, l'accent est mis sur le contrôle et sur la consignation de l'apprentissage, car ces dimensions sont essentielles pour fournir à l'élève une rétroaction exploitable.

Figure 46

Les dimensions de l'évaluation au sein du PP



L'évaluation au sein du PP repose sur les hypothèses qui suivent.

- L'évaluation est un processus collaboratif continu qui a cours entre le personnel enseignant et les élèves, mais aussi entre pairs, pour rassembler des preuves de l'apprentissage qui seront analysées afin de servir de base de réflexion et d'action pour orienter l'enseignement.
- Pour évaluer le travail des élèves, il convient d'employer diverses stratégies, en tenant compte des procédés variés, complexes et recherchés grâce auxquels les élèves comprennent leurs expériences.
- Les élèves participent activement à l'évaluation et à la réflexion sur leur apprentissage, et agissent en tenant compte des rétroactions fournies par leurs pairs et le personnel enseignant. Ces rétroactions leur servent de proactions pour les prochaines étapes de l'apprentissage.
- Les objectifs d'apprentissage et les critères de réussite sont élaborés collectivement et clairement communiqués.
- Les objectifs d'apprentissage et le processus d'apprentissage sont évalués.
- L'évaluation est conçue en adoptant une approche à la fois rétrospective et prospective.
- La promotion d'une culture de l'évaluation passe par le développement des capacités d'évaluation de chaque membre de la communauté d'apprentissage.

## Lectures complémentaires

Pour de plus amples informations sur le PP, veuillez consulter le document *Le Programme primaire : des principes à la pratique*, qui est disponible sur le Centre de ressources pédagogiques.

# Gros plan sur la révision de notation de l'évaluation interne

La révision de notation est employée pour les travaux évalués en interne dans le cadre du Programme d'éducation intermédiaire (PEI), du Programme du diplôme et du Programme à orientation professionnelle (POP) afin de veiller à l'application d'une norme commune dans tous les établissements. À l'issue de la révision de notation, les notes de l'établissement peuvent être augmentées, diminuées ou laissées inchangées. La révision de notation a pour objectif de vérifier si les enseignantes et enseignants ont appliqué les critères d'évaluation avec précision et cohérence lors de la notation des travaux des élèves.

## L'échantillonnage

Une fois que l'établissement scolaire a envoyé ses notes d'évaluation interne, le Baccalauréat International (IB) sélectionne les élèves à inclure dans l'échantillonnage envoyé pour la révision de notation. Cet échantillonnage est soigneusement constitué de manière à assurer une représentation appropriée de la gamme des notes de l'établissement. Le nombre d'échantillons envoyés pour la révision de notation peut être de dix, huit, cinq ou moins, en fonction du nombre d'inscriptions d'élèves dans la matière. Dans cet échantillonnage, les travaux de trois élèves intégreront l'échantillonnage initial pour illustrer la gamme de notes et déterminer si la notation de l'enseignant ou enseignante se situe dans la marge de tolérance de la norme mondiale. Dans le cas contraire, le reste de l'échantillonnage (que l'on appelle l'« échantillonnage accru ») sera également noté par l'examineur ou examinatrice. L'IB tend à ne pas sélectionner les élèves ayant obtenu les notes maximales pour la révision de notation. Cela permet d'offrir aux élèves des tranches de notes supérieures la possibilité d'obtenir une meilleure note à la suite d'une révision de notation.

Lorsque l'établissement inscrit suffisamment d'élèves dans une matière et un niveau donnés pour pouvoir les diviser en plusieurs classes, et que plusieurs enseignantes et enseignants prennent part à l'évaluation interne, l'IB exige de ces personnes qu'elles se partagent la tâche et travaillent ensemble afin d'uniformiser la façon dont elles se servent des critères d'évaluation. Un seul échantillonnage est alors requis qui, selon toute probabilité, contiendra des travaux envoyés pour l'évaluation notés par l'ensemble des enseignantes et enseignants. Toutefois, lorsque les classes utilisent des langues d'usage différentes pour la même matière, un échantillonnage différent doit être soumis pour chaque langue d'usage.

## Le calcul des facteurs de révision de notation

Toutes les composantes évaluées en interne sont notées à l'aide de critères d'évaluation ou de bandes de notation. Dans la majorité des cas, les enseignantes et enseignants disposent d'un nombre bien plus important d'informations sur le contexte de travail et la démarche suivie par les élèves pour effectuer ce travail que les examinateurs et examinatrices. De ce fait, les personnes qui révisent des composantes évaluées en interne doivent juger du caractère approprié de la notation de l'enseignant ou enseignante et non simplement recorriger le travail sans tenir compte des notes attribuées. Ces notes ne doivent être modifiées que lorsque les personnes chargées de réviser la notation ont la certitude qu'elles sont inappropriées et non conformes à la norme mondiale.

L'échantillonnage de travaux notés par l'enseignant ou enseignante est révisé par l'examineur ou examinatrice et, sur la base d'une comparaison statistique entre les deux séries de notes, un ajustement est apporté, au besoin, aux notes attribuées par l'enseignant ou enseignante (pour cette composante) à l'ensemble des élèves de l'établissement.

Si les notes attribuées par l'enseignant ou enseignante sont systématiquement trop basses ou trop élevées, l'ajustement sera le même pour chacune des notes attribuées par l'enseignant ou enseignante.

Dans l'exemple qui suit (tableau 8), l'enseignant ou enseignante a sous-noté ses élèves en les pénalisant de 5 points en moyenne ; c'est pourquoi toutes les notes de la cohorte seront remontées de 5 points.

Tableau 8

**Sous-notation constante**

| Notes de l'enseignant ou enseignante | Notes de l'examineur ou examinatrice | Différence entre la notation de l'enseignant ou enseignante et celle de l'examineur ou examinatrice | Notes définitives révisées |
|--------------------------------------|--------------------------------------|---|----------------------------|
| 15                                   | 20                                   | 5   | 20                         |
| 11                                   | 17                                   | 6   | 16                         |
| 10                                   | 14                                   | 4   | 15                         |
| 8                                    | 12                                   | 4   | 13                         |
| 4                                    | 10                                   | 6   | 9                          |

Si, dans la tranche de notes supérieure ou inférieure, les notes attribuées par l'enseignant ou enseignante sont trop basses ou trop élevées, l'ajustement peut varier en fonction du niveau initial des notes.

Dans l'exemple qui suit (tableau 9), la notation de l'enseignant ou enseignante n'est pas constamment conforme à la norme mondiale. Les notes sont jugées trop généreuses dans la partie supérieure de la gamme de notes, mais trop sévères dans la partie inférieure. L'ajustement appliqué à la notation ne sera donc pas le même pour toute la gamme de notes.

Tableau 9

**Notation inconstante**

| Notes de l'enseignant ou enseignante | Notes de l'examineur ou examinatrice | Différence entre la notation de l'enseignant ou enseignante et celle de l'examineur ou examinatrice | Notes définitives révisées |
|--------------------------------------|--------------------------------------|---|----------------------------|
| 37                                   | 35                                   | -2  | 35                         |
| 25                                   | 22                                   | -3  | 24                         |
| 17                                   | 18                                   | 1   | 17                         |
| 12                                   | 14                                   | 2   | 13                         |
| 6                                    | 8                                    | 2   | 8                          |

Si l'enseignant ou enseignante respecte la norme de notation, aucun ajustement ne sera apporté.

Les notes de l'enseignant ou enseignante qui sont présentées dans le tableau 10 se situent toutes dans la marge de tolérance de la norme mondiale. Les élèves dans leur intégralité se verront donc attribuer ces notes.

Tableau 10

**Notes respectant la norme de notation**

| Notes de l'enseignant ou enseignant | Notes de l'examineur ou examinatrice | Différence entre la notation de l'enseignant ou enseignante et celle de l'examineur ou examinatrice | Notes définitives révisées |
|-------------------------------------|--------------------------------------|---|----------------------------|
| 32                                  | 31                                   | -1  | 32                         |
| 29                                  | 30                                   | 1   | 29                         |
| 8                                   | 8                                    | 0   | 8                          |

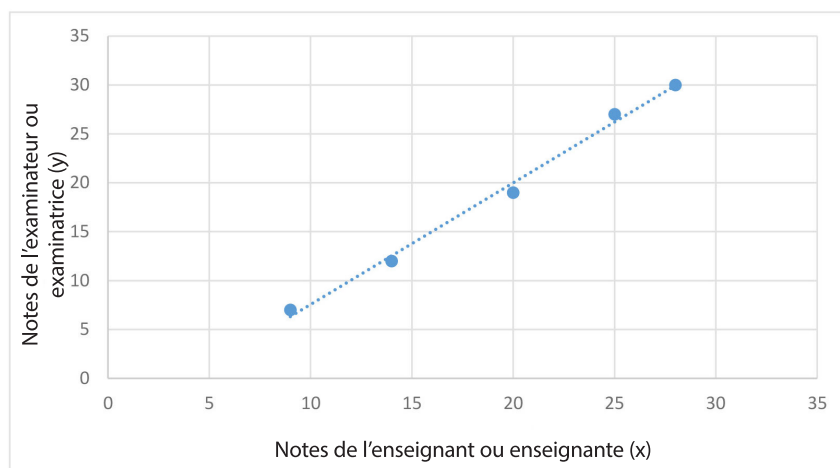
## La régression linéaire

Les données recueillies pour chaque échantillonnage envoyé pour la révision de notation font l'objet d'une analyse qui permet en premier lieu de déterminer si les notes attribuées par l'enseignant ou enseignante se situent dans la marge de tolérance. Dans le cas contraire, un ajustement approprié sera appliqué à toutes les notes attribuées par l'enseignant ou enseignante, en se basant sur la tendance générale que l'on trouve dans l'échantillonnage. La technique utilisée est appelée « régression linéaire » et consiste à calculer la droite se rapprochant le plus de tous les points du nuage de points de coordonnées qui représente les notes attribuées par l'enseignant ou enseignante et l'examineur ou examinatrice pour les travaux inclus dans l'échantillonnage. Un exemple de régression linéaire et l'ensemble de notes correspondant sont présentés dans la figure 47.

La droite de régression de la révision de notation qui est présentée dans la figure 47 correspond à un enseignant ou une enseignante tendant à avoir une notation un peu trop sévère tout en haut de la gamme de notes et trop généreuse en bas. Chaque point représente la paire de notes données à l'un des travaux de l'échantillonnage par l'enseignant ou enseignante et l'examineur ou examinatrice. La droite de régression continue est utilisée pour convertir les notes de l'enseignant ou enseignante en notes révisées.

Figure 47

Régression linéaire et ensemble de notes correspondant



| Notes de l'enseignant ou enseignante | Notes de l'examineur ou examinatrice | Notes définitives révisées |
|--------------------------------------|--------------------------------------|----------------------------|
| 28                                   | 30                                   | 30                         |
| 25                                   | 27                                   | 26                         |
| 20                                   | 19                                   | 20                         |
| 14                                   | 12                                   | 12                         |
| 9                                    | 7                                    | 6                          |

L'équation de la droite de régression, calculée à l'aide des données de l'échantillonnage, peut être utilisée pour convertir chaque note attribuée par l'enseignant ou enseignante ( $x$ ) en une note moyenne équivalente ( $y$ ) que l'examineur ou examinatrice aurait très probablement attribuée à l'élève en question. Cet ajustement apporté à la notation de l'enseignant ou enseignante suite à la révision de notation et basé sur l'extrapolation (on passe d'un échantillonnage à un ensemble de notes plus important) peut uniquement refléter les tendances générales observées dans la notation. Le but de la révision de notation est de s'assurer que les notes des élèves sont, dans l'ensemble, ramenées à des niveaux plus appropriés. S'agissant d'une approche de meilleur ajustement, les notes révisées ne correspondront pas aux notes de l'examineur ou examinatrice pour l'intégralité des élèves.

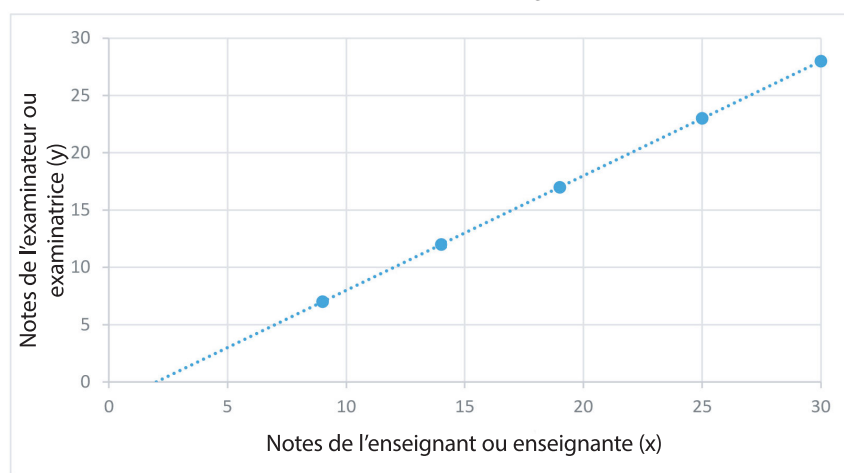
## L'échec à la révision de notation

Une vérification est automatiquement effectuée pour contrôler que la droite de régression linéaire (c'est-à-dire le facteur de révision de notation calculé) satisfait à certaines conditions avant d'être appliquée à toutes les notes de l'enseignant ou enseignante. Dans certains cas, il est impossible de calculer un facteur de révision de notation sur la base de l'échantillonnage de travaux reçu. L'une des mesures statistiques utilisées est le coefficient de corrélation (coefficient de corrélation de Pearson). Il mesure la cohérence de la relation entre la notation de l'enseignant ou enseignante et celle de l'examineur ou examinatrice.

Un coefficient de corrélation égal à 1 indique une parfaite cohérence dans la relation entre les notes et un accord dans le classement des élèves, du meilleur travail au plus mauvais (sans pour autant que les notes soient exactement identiques), comme l'illustre la figure 48.

Figure 48

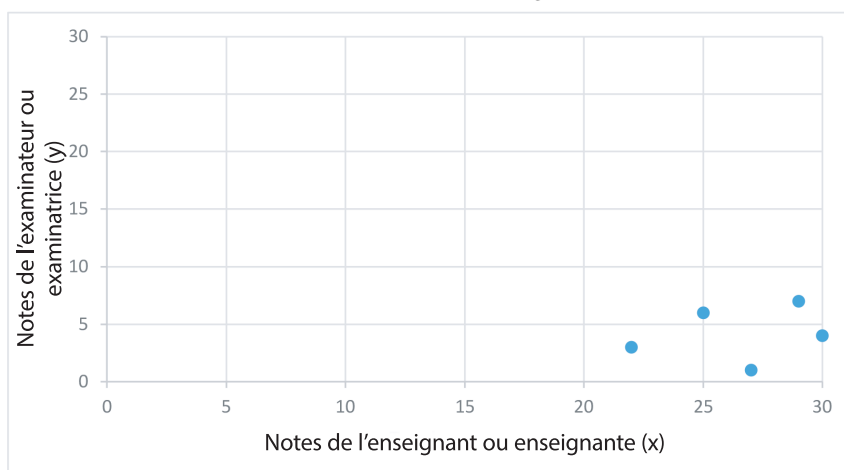
Coefficient de corrélation égal à 1



Un coefficient égal à 0 indique qu'il n'y a aucune relation, comme l'illustre la figure 49.

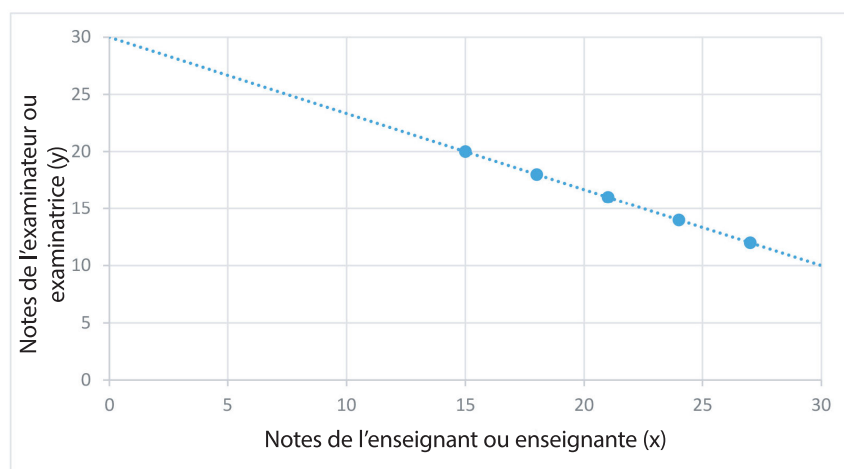
Figure 49

Coefficient de corrélation égal à 0



Un coefficient égal à  $-1$  indique un désaccord constant en ce qui concerne le mérite des travaux des élèves, l'enseignant ou enseignante et l'examineur ou examinatrice produisant des classements opposés, comme l'illustre la figure 50.

Figure 50  
Coefficient de corrélation égal à  $-1$



Pour que le facteur de révision de notation calculé soit acceptable, le coefficient de corrélation doit être égal ou supérieur à 0,85, ce qui indique un haut degré d'accord entre l'enseignant ou enseignante et l'examinateur ou examinatrice. Toutefois, un haut coefficient de corrélation seul ne suffit pas à garantir que le facteur de révision de notation est approprié. Il est par ailleurs vérifié que la pente de la droite de régression est comprise entre 0,5 et 1,5. Si la pente de la droite est trop faible, cela signifie que l'enseignant ou enseignante a utilisé une gamme de notes trop large, donnant relativement trop peu de points aux travaux faibles et trop de points aux bons travaux, même si cela est fait de façon cohérente. L'examinateur ou examinatrice a donc dû comprimer considérablement la gamme de notes attribuées par l'enseignant ou enseignante. Si la pente est supérieure à 1,5, la droite est trop abrupte et c'est alors le contraire qui se produit (en d'autres termes, l'enseignant ou enseignante n'a pas suffisamment différencié les mauvais et les bons travaux envoyés pour l'évaluation, et l'examinateur ou examinatrice a dû élargir la gamme de notes attribuées).

Un échantillonnage échouera aux vérifications automatiques de la révision de notation si le coefficient de corrélation est inférieur à 0,85 ou si la pente n'est pas comprise entre 0,5 et 1,5.

Les cas de tous les échantillonnages envoyés par des établissements qui échouent à la révision de notation sont étudiés (au cas par cas) par le personnel chargé de l'évaluation de l'IB, qui examine avec soin les données sous-jacentes et peut décider :

- de déclarer que la droite de régression calculée est appropriée pour la gamme de notes particulière de l'enseignant ou enseignante ;
- d'utiliser une autre sorte d'ajustement pour la révision de notation, qui sera appropriée pour la gamme de notes de l'enseignant ou enseignante ;
- de demander des données supplémentaires sur l'échantillonnage afin de clarifier la tendance ;
- de demander les travaux des autres élèves pour procéder à une recorection complète de la notation de l'enseignant ou enseignante.

Tous les travaux d'évaluation interne des élèves doivent être disponibles jusqu'à la clôture de la session d'examens pour que l'IB puisse résoudre tous les échecs à la révision de notation.

## L'adaptation du modèle linéaire

Le modèle de la ligne droite suivi pour la révision de notation est, dans une certaine mesure, modifié en utilisant la « déformation ». L'ajustement obtenu grâce à la ligne droite peut avoir des effets inappropriés sur les notes extrêmes de la gamme de notes possibles, empêchant l'obtention d'un 0. L'élève dont le travail vaut vraiment que l'on lui attribue un 0 peut obtenir quelques points alors que rien dans son travail

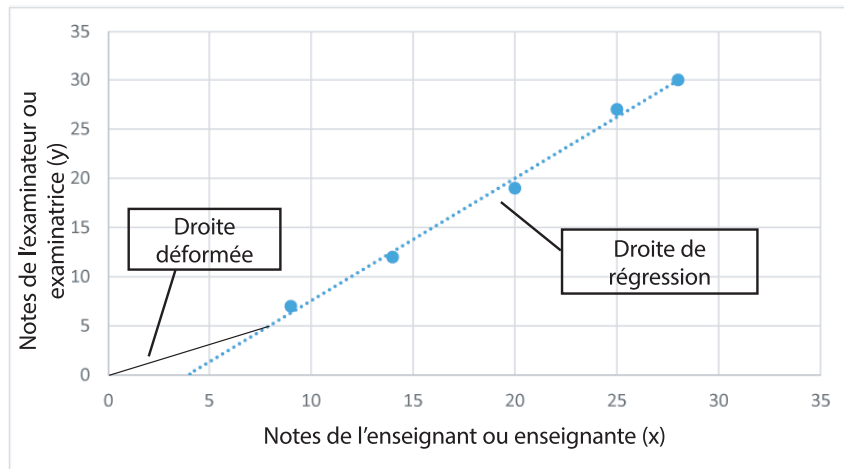
ne le justifie ou l'élève ayant produit un travail faible mais valant quelques points peut voir sa note transformée en 0 suite à la révision de notation.

Pour résoudre ce problème, il est possible de « déformer » la partie extrême de la droite, qui représente les notes situées dans les 20 % inférieurs de la gamme de notes disponibles. À cet extrême, la droite de régression est modifiée pour être remplacée par une nouvelle droite « déformée » qui relie la droite de régression aux points de coordonnées minimum. La déformation de la droite de régression n'est pas effectuée à l'extrémité supérieure de la gamme de notes, car il est souvent observé que le personnel enseignant attribue le nombre de points maximum à des élèves qui ne le méritent pas.

Le graphique présenté dans la figure 51 illustre la déformation de la droite de régression pour éviter que les notes des élèves ne soient ajustées en s'approchant trop des valeurs minimales.

Figure 51

Droite de régression déformée



La déformation garantit qu'une note révisée égale à 0 ne peut provenir que d'un 0 attribué par l'enseignant ou enseignante. Cela évite d'attribuer un 0 aux travaux qui valent quelques points et d'attribuer quelques points aux travaux qui n'en valent aucun. Cela suppose bien évidemment que la notation de l'enseignant ou enseignante passe avec succès le processus de révision de notation automatique. Si la notation de l'enseignant ou enseignante échoue à la révision de notation et ne peut pas être automatiquement révisée, la droite de régression n'est pas déformée.

## Bibliographie

- ANANIADOU, K. et CLARO, M. 21st century skills and competences for new millennium learners in OECD countries. *OECD Education Working Papers*. 2009, numéro 41. OECD [OCDE].
- ANDERSON, L. W. et KRATHWOHL, D. R. 2001. *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- BAIRD, J. A., CRESSWELL, M. et NEWTON, P. Would the real gold standard please step forward? *Research Papers in Education*. 2000, volume 15, numéro 2, p. 213 – 229. <https://doi.org/10.1080/026715200402506>
- BIGGS, J. Enhancing teaching through constructive alignment. *Higher Education*. 1996, volume 32, numéro 3, p. 347 – 364. <https://www.jstor.org/stable/3448076>
- BLACK, P. Assessment, learning theories and testing systems. In MURPHY, P. (sous la direction de). 1999. *Learners, learning & assessment*, p. 118 – 134. Paul Chapman Publishing in association with The Open University.
- BLOOM, B. S. (sous la direction de), ENGELHART, M. D., FURST, E. J., HILL, W. H. et KRATHWOHL, D. R. 1956. *Taxonomy of educational objectives: Handbook 1: Cognitive domain*. Longman.
- BROADFOOT, P. M. 1996. *Education, assessment and society: A sociological analysis*. Open University Press.
- BROWN, R. Cultural dimensions of national and international educational assessment. In HAYDEN, M., THOMPSON, J. et WALKER, G. (sous la direction de). 2002. *International education in practice: Dimensions for schools and international schools*. Routledge.
- CHAMBERLAIN, S. AQA: Public perceptions of reliability. In OFQUAL. 2010. *Reliability compendium*, chapitre 18. The Office of Qualifications and Examinations Regulation.
- COE, R., SEARLE, J., BARMBY, P., JONES, K. et HIGGINS, S. 2008. *Relative difficulty of examinations in different subjects*. Report for SCORE (Science Community Supporting Education). CEM Centre : Université de Durham.
- CRESSWELL, M. J. Defining, setting and maintaining standards in curriculum-embedded examinations: Judgemental and statistical approaches. In GOLDSTEIN, H. et LEWIS, T. (sous la direction de). 1996. *Assessment: Problems, developments and statistical issues*, p. 57 – 84. Wiley.
- CRESSWELL, M. J. Examination grades: How many should there be? *British Educational Research Journal*. 1986, volume 12, numéro 1, p. 37 – 54.
- CRESSWELL, M. J. The role of public examinations in defining and monitoring standards. *Proceedings of the British Academy*. 2000, volume 102, p. 69 – 120.
- CROOKS, T. J., KANE, M. T. et COHEN, A. S. Threats to the valid use of assessments. *Assessment in Education: Principles, Policy & Practice*. 1996, volume 3, numéro 3, p. 265 – 286. <https://doi.org/10.1080/0969594960030302>
- DOLAN, R. P., BURLING, K., HARMS, M., STRAIN-SEYMOUR, E., WAY, W. et ROSE, D. 2013. *A universal design for learning-based framework for designing accessible technology-enhanced assessments* (rapport de recherche). Pearson. <https://eric.ed.gov/?id=ED576691>
- FRITH, D. S. et MACINTOSH, H. G. 1984. *A teacher's guide to assessment*. Stanley Thornes.
- GIBBS, G. 1992. *Teaching more students: Assessing more students*. The Oxford Centre for Staff Learning and Development : Université d'Oxford Brookes.
- GIPPS, C. et MURPHY, P. 1994. *A fair test? Assessment, achievement and equity*. Open University Press.
- GLASER, R. Instructional technology and the measurement of learning outcomes: Some questions. *American Psychologist*. 1963, volume 18, numéro 8, p. 519 – 521. <https://doi.org/10.1037/h0049294>
- GOLDSTEIN, H. Group differences and bias in assessment. In GOLDSTEIN, H. et LEWIS, T. (sous la direction de). 1996. *Assessment: Problems, developments and statistical issues*, p. 85 – 93. John Wiley & Sons Ltd.

- GOOD, F. J. et CRESSWELL, M. J. 1988. *Grading the GCSE*. Secondary Examinations Council.
- HE, Q., OPPOSS, D. et BOYLE, A. A quantitative investigation into public perceptions of reliability in examination results in England. In OFQUAL. 2010. *Reliability compendium*, chapitre 19, p. 68. The Office of Qualifications and Examinations Regulation.
- HIERONYMUS, A. N. et HOOVER, H. D. 1986. *Iowa tests of basic skills: Manual for school administrators*. Riverside Publishing Company.
- HUGHES, D. C., KEELING, B. et TUCK, B. F. Effects of achievement expectations and handwriting quality on scoring essays. *Journal of Educational Measurement*. 1983, volume 20, numéro 1, p. 65 – 70. <https://www.jstor.org/stable/1434941>
- HUMPHREYS, L. G. An analysis and evaluation of test and item bias in the prediction context. *Journal of Applied Psychology*. 1986, volume 71, numéro 2, p. 327 – 333. <https://doi.org/10.1037/0021-9010.71.2.327>
- LAMBERT, D. et LINES, D. 2000. *Understanding assessment: Purposes, perceptions, practice*. Routledge Falmer.
- LINN, M. C. Gender differences in educational achievement. In PFLEIDERER, J. (sous la direction de). 1992. *Sex equity in educational opportunity, achievement, and testing*. Educational Testing Service.
- LLEWELLYN, D. 2014. *Inquire within: Implementing inquiry- and argument-based science standards in grades 3–8*, 3<sup>e</sup> édition. Corwin.
- MEYER, A., ROSE, D. H. et GORDON, D. 2014. *Universal design for learning: Theory and practice*. CAST Professional Publishing.
- MITRA, S. 2011. *Responsibility, leadership and education* [discours liminaire]. Conférence des chefs d'établissement tenue à Singapour du 13 au 16 octobre.
- MURPHY, P. 1999. *Learners, learning & assessment*. Paul Chapman Publishing en association avec The Open University.
- NATIONAL RESEARCH COUNCIL. 2011. *Assessing 21st century skills: Summary of a workshop*. The National Academies Press.
- NEWTON, P. A level pass rates and the enduring myth of norm-referencing. *Research Matters*. 2011. <https://doi.org/10.17863/CAM.100441>
- NEWTON, P. E. Clarifying the purposes of educational assessment. *Assessment in Education: Principles, Policy & Practice*. 2007, volume 14, numéro 2, p. 149 – 170. <https://doi.org/10.1080/09695940701478321>
- NEWTON, P. E. 2012. *We need to talk about validity* [exposé]. Réunion annuelle du National Council for Measurement in Education à Vancouver, Canada.
- NEWTON, P. E. et SHAW, S. D. 2014. *Validity in educational and psychological assessment*. Londres, Royaume-Uni : SAGE Publications.
- ORGANISATION DU BACCALAURÉAT INTERNATIONAL [IB]. 2019. *En quoi consiste le système éducatif de l'IB ?* Organisation du Baccalauréat International.
- ORGANISATION DU BACCALAURÉAT INTERNATIONAL [IB]. 2024. Déclaration de mission de l'IB. Organisation du Baccalauréat International.
- PETERSON, A. D. C. 1971. *New techniques for assessment of pupils' work*. Conseil de l'Europe. [Cet ouvrage est disponible en français sous le titre *Nouvelles techniques d'évaluation du travail scolaire*.]
- PETERSON, A. D. C. 2003. *Schools across frontiers: The story of the International Baccalaureate and the United World Colleges*, 2<sup>e</sup> édition. Open Court Publishing Company.
- POPHAM, W. J. 1978. *Criterion-referenced measurement*. Prentice Hall.
- RAND CORPORATION. 2012. *Teaching and learning 21st century skills: Lessons from the learning sciences*. Asia Society. Référence du 1<sup>er</sup> janvier 2017. [http://www.rand.org/pubs/external\\_publications/EP51105.html](http://www.rand.org/pubs/external_publications/EP51105.html)
- RAO, K., CURRIE-RUBIN, R. et LOGLI, C. 2016. *UDL and inclusive practices in IB schools worldwide*. CAST Professional Learning.
- SCHLEICHER, A. 2016. *The case for 21st-century learning*. OCDE. Référence du 4 avril 2025. <https://web.archive.oecd.org/2012-06-14/61660-the-case-for-21st-century-learning.htm>

- SHEPARD, L. A. Commentary: What policy makers who mandate tests should know about the new psychology of intellectual ability and learning. In GIFFORD, B. R. et O'CONNOR, M. C. (sous la direction de). 1992. *Changing assessments: Alternative views of aptitude, achievement and instruction*, p. 301 – 328. Kluwer Academic Publishers.
- SIRECI, S. G. On validity theory and test validation. *Educational Researcher*. 2007, volume 36, numéro 8, p. 477 – 481. <https://www.jstor.org/stable/4621103>
- SMITH, D. J. et TOMLINSON, S. 1989. *The school effect: A study of multi-racial comprehensives*. Policy Studies Institute.
- SNYDER, B. R. 1971. *The hidden curriculum*. MIT Press.
- SURGENOR, P. 2010. *Teaching toolkit: Effect of assessment on learning*. Université de Dublin (UCD).
- VYGOTSKY, L. S. 1962. *Thought and language*. MIT Press. [Cet ouvrage est disponible en français sous le titre *Pensée et langage*.]
- WILIAM, D. Validity, dependability and reliability in National Curriculum assessment. *The Curriculum Journal*. 1993, volume 4, numéro 3, p. 335 – 350. <https://doi.org/10.1080/0958517930040303>
- WINKLEY, J. et CRESSWELL, M. Introduction to the concept of reliability. In OFQUAL. 2011. *Reliability compendium*, chapitre 1. The Office of Qualifications and Examinations Regulation.
- WOOD, D. 1998. *How children think and learn: The social contexts of cognitive development*, 2<sup>e</sup> édition. Blackwell Publishing.
- WOOD, D., BRUNER, J. S. et ROSS, G. The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*. 1976, volume 17, numéro 2, p. 89 – 100. <https://doi.org/10.1111/j.1469-7610.1976.tb00381.x>
- ZANGA, G. et DE GIOANNIS, E. Discrimination in grading: A scoping review of studies on teachers' discrimination in school. *Studies in Educational Evaluation*. 2023, volume 78. <https://doi.org/10.1016/j.stueduc.2023.101284>

# Glossaire

L'IB s'engage à adopter un langage accessible et à employer le moins de termes techniques possible. Pour confirmer cet engagement, le glossaire proposé ci-après vient définir certains des mots-clés et concepts qui figurent dans la présente publication ainsi que dans d'autres ressources en rapport avec l'évaluation. Quelques-uns de ces termes sont également susceptibles d'apparaître dans des communications de l'IB.

| Terme  | Définition  |
|--|---|
| <b>Accès et inclusion</b>                                  | Finalité d'une évaluation tenant compte des besoins de l'ensemble des élèves, afin que chaque élève puisse montrer de manière équitable ses compétences dans la matière.  |
| <b>Alignement</b>  | Accord de principe qui s'opère également dans la pratique entre des valeurs et idéaux en matière d'apprentissage (le programme établi), les pratiques d'enseignement (le programme enseigné) et l'apprentissage réalisé par les élèves (le programme évalué).                 |
| <b>Aménagements à des fins d'accès et d'inclusion</b>      | Conditions modifiées ou supplémentaires lors du processus d'évaluation pour les élèves ayant des besoins en matière d'aménagement de la procédure d'évaluation. Ces aménagements permettent aux élèves de montrer le niveau atteint de manière plus équitable.                |
| <b>Appareil</b>  | Ordinateur portable ou ordinateur de bureau utilisé pour les évaluations numériques de l'IB.<br>Le terme « appareil » est préféré au terme « ordinateur » pour rendre compte de l'exploration en cours d'une compatibilité future avec des tablettes et d'autres plateformes. |
| <b>Approche comparée (voir Approche critériée limitée)</b> | Comparaison de l'accomplissement de l'élève à des descriptions prédéfinies de l'accomplissement (critères) et à l'accomplissement des élèves des cohortes précédentes. Il s'agit de l'approche adoptée par l'IB pour garantir le maintien des normes.                         |
| <b>Approche critériée (voir Approche comparée)</b>         | Comparaison de l'accomplissement de l'élève à des descriptions prédéfinies de l'accomplissement (critères) pour la notation.  |
| <b>Approche critériée limitée (voir Approche comparée)</b> | Comparaison de l'accomplissement de l'élève à des descriptions prédéfinies de l'accomplissement (critères) et à l'accomplissement des élèves des cohortes précédentes. Il s'agit de l'approche adoptée par l'IB pour garantir le maintien des normes.                         |
| <b>Approche normative</b>                                  | Pratique selon laquelle le résultat est déterminé en comparant l'accomplissement de l'élève à celui de l'ensemble de la cohorte pour laquelle l'évaluation est conçue.  |
| <b>Argument de validité</b>                                | Preuve et explication des décisions prises lors de la création d'une évaluation, qui justifient son adéquation avec l'objectif poursuivi.   |
| <b>Attribution des notes finales</b>                       | Processus par lequel la question de la conversion des notes en notes finales est réglée. Il garantit que les notes finales ont la même signification quelle que soit la session pendant laquelle l'élève passe ses examens.   |

| Terme  | Définition  |
|--|---|
| <b>Authentification</b>                          | Preuves et processus de vérification assurant que le travail a bien été réalisé par l'élève. Cela englobe notamment les signatures de l'enseignant ou enseignante et de l'élève confirmant que le travail est bien celui de l'élève en question.  |
| <b>Autonotation</b>                              | Utilisation de la technologie pour évaluer le travail des élèves à l'aide d'un barème de notation prédéfini qui considère objectivement une réponse comme correcte ou incorrecte.   |
| <b>Bandes de notation</b>                        | Gamme particulière de notes qui doivent être attribuées à une réponse laissant apparaître certaines qualités.   |
| <b>Barème de notation</b>                        | Indication de la façon d'attribuer les niveaux des critères pour une tâche particulière.  |
| <b>Barème de notation analytique</b>             | Barème de notation indiquant les bonnes réponses et les éléments pour lesquels il faut attribuer des points.  |
| <b>Biais</b>                                     | Effet conduisant un groupe défini à obtenir des résultats différents de ceux de la moyenne à une question ou une tâche particulière pour une raison autre que ses aptitudes à développer la compétence ou la connaissance évaluée.  |
| <b>Chef d'équipe</b>                             | Examineur ou examinatrice qui dirige une équipe composée d'examineurs et examinatrices.   |
| <b>Circonstances défavorables</b>                | Circonstances échappant au contrôle de l'élève et susceptibles d'être préjudiciables à ses résultats à l'évaluation. Ce sont, par exemple, un stress grave, des circonstances familiales particulièrement éprouvantes, un deuil ou des événements pouvant menacer la santé ou la sécurité des élèves. Il arrive que des circonstances défavorables concernent un groupe ou l'ensemble des élèves dans un établissement scolaire donné. Les circonstances défavorables n'incluent pas : <ul style="list-style-type: none"> <li>• les insuffisances imputables à l'établissement scolaire des élèves ;</li> <li>• l'incapacité des élèves à améliorer leurs résultats malgré les aménagements à des fins d'accès et d'inclusion autorisés et mis en place.</li> </ul> |
| <b>Circonstances exceptionnelles</b>             | Circonstances qui ne font pas partie de l'expérience courante des autres élèves ayant des besoins en matière d'aménagement de la procédure d'évaluation. L'IB se réserve le droit de déterminer les circonstances à qualifier d'« exceptionnelles » qui justifieraient par conséquent un aménagement particulier à des fins d'accès et d'inclusion.   |
| <b>Comparabilité</b>                             | Mesure dans laquelle un résultat particulier peut être considéré comme identique à un autre résultat. Elle permet généralement de comparer deux années ou deux matières.  |
| <b>Composante d'évaluation</b>                   | Tâche ou ensemble de tâches regroupées, représentant une partie de l'évaluation globale, laquelle peut prendre la forme d'un examen, d'un dossier de travail, d'un projet ou d'un travail de recherche, entre autres.   |
| <b>Conception universelle de l'apprentissage</b> | Orientation et cadre définis pour l'apprentissage afin de cibler l'intégralité des élèves. L'objectif est de créer des environnements inclusifs en proposant une multitude de moyens de s'investir, de représenter et de s'exprimer.  |
| <b>Conception universelle de l'évaluation</b>    | Concept selon lequel toutes les évaluations devraient être conçues sur la base d'une compréhension des différents besoins des élèves au lieu de traiter quelques élèves différemment. Cette idée fait partie intégrante de l'engagement de l'IB envers la conception universelle de l'apprentissage.  |

| Terme  | Définition  |
|--|---|
| <b>Configuration requise</b>                           | Spécifications techniques minimales à respecter par un appareil pour un déroulement des examens numériques de l'IB en toute fluidité et sécurité. Ces exigences ont été fixées pour garantir que le matériel informatique ne causera aucun préjudice aux élèves pendant les évaluations.<br><br>Ces exigences sont publiées à la fois pour le PEI et pour le Programme du diplôme et le POP. Les établissements scolaires doivent s'assurer que tous les appareils des élèves respectent les spécifications requises.                       |
| <b>Console d'administration</b>                        | Site Web sécurisé utilisé pour les examens sur ordinateur du PEI. Les coordonnateurs et coordonnatrices du PEI, le surveillant principal désigné ou la surveillante principale désignée peuvent y effectuer des activités administratives essentielles, comme enregistrer des appareils et accéder aux fichiers d'examen. L'accès à la console d'administration s'effectue à partir d'un lien affiché sur IBIS.<br><br>Remarque : pour les examens numériques du Programme du diplôme et du POP, voir <b>Système d'examens numériques</b> . |
| <b>Copie</b>   | Réponses de l'élève à un examen sur papier. Il peut également s'agir de tout travail envoyé pour l'évaluation.  |
| <b>Cours</b>   | Nombre prescrit de leçons ou d'heures d'enseignement au sein d'une période d'enseignement déterminée. Les établissements organisent l'apprentissage et l'enseignement dans les matières au moyen de cours disciplinaires et interdisciplinaires.  |
| <b>Créateur ou créatrice de contenu d'évaluations</b>  | Personne qui crée, examine minutieusement et révisé les questions d'examen proposées ainsi que les barèmes de notation associés qui seront utilisés pour les évaluations. Il s'agit presque toujours des examinateurs superviseurs et examinatrices superviseuses, bien que ce travail requière des compétences différentes.  |
| <b>Critères d'évaluation</b>                           | Critères permettant de mesurer l'accomplissement de l'élève.  |
| <b>Critères globaux</b>                                | Approche adoptée pour évaluer le travail de l'élève, qui consiste à examiner le travail dans son ensemble au lieu d'examiner séparément des éléments distincts (par exemple, communication, connaissances spécifiques à la discipline et qualité de l'argumentation).   |
| <b>Cycle d'évaluation</b>                              | Mesures prises pour créer, effectuer et noter les évaluations, y compris les examens. Il s'agit d'un cycle car l'IB tire des enseignements de chaque session d'examens afin d'améliorer les sessions suivantes.   |
| <b>Descripteurs de niveaux</b>                         | Descriptions des caractéristiques devant être présentes dans le travail envoyé pour l'évaluation pour chaque niveau.  |
| <b>Descripteurs des notes finales</b>                  | Ensemble des qualités attendues des élèves pour chaque note finale. Les descripteurs des notes finales peuvent être spécifiques à une matière, spécifiques à un groupe de matières ou généraux pour s'appliquer à l'ensemble du programme. Dans tous les cas, ils doivent décrire les mêmes caractéristiques. Les exemples plus précis ne font qu'expliquer ce que ces descriptions signifient dans le contexte spécifique à la matière.  |
| <b>Différenciation (apprentissage et enseignement)</b> | Modification des stratégies d'enseignement pour répondre aux divers besoins d'apprentissage des élèves, en utilisant des contenus, processus et outils variés.  |

| Terme  | Définition  |
|--|---|
| <b>Différenciation (évaluation)</b>                        | Action de distinguer des élèves montrant différents niveaux de compétence.  |
| <b>Discipline</b>  | Domaine d'apprentissage ou champ d'étude scolaire. Manière d'organiser les connaissances à des fins d'enseignement. (Pour des raisons pratiques liées à l'évaluation, les disciplines sont généralement appelées « matières » dans le contexte du PEI et du Programme du diplôme.)  |
| <b>Échantillonnage dynamique</b>                           | Amélioration du processus de révision de notation permettant une meilleure utilisation des contrôles de qualité.  |
| <b>Échantillonnage envoyé pour la révision de notation</b> | Sélection de travaux qui est envoyée à l'IB pour s'assurer que la notation est conforme à la norme attendue.  |
| <b>Effet de remous</b>                                     | Impact que les parties ultérieures d'un processus ont sur les parties antérieures. Dans le contexte de l'éducation, l'effet de remous désigne généralement la façon dont l'apprentissage et l'enseignement sont modifiés par l'évaluation.  |
| <b>Élément</b>   | Plus petite subdivision appropriée d'une évaluation. Partie isolée, mais entière, d'une évaluation qui est jaugée conformément à ce qui est décrit, indiqué et attendu dans les sections « Résumé du programme » ou « Description détaillée de l'évaluation externe » des guides pédagogiques de l'IB.  |
| <b>Élève</b>   | Personne suivant un cours ou un programme de l'IB.  |
| <b>Envoi</b>   | Processus par lequel l'élève (ou l'établissement au nom de l'élève) rend la version définitive d'un travail à l'IB.   |
| <b>Épreuve</b>   | Voir <b>Épreuve d'examen</b> .  |
| <b>Épreuve d'examen</b>                                    | Ensemble de tâches et de questions que l'élève doit réaliser et traiter lors d'un examen. Ce terme peut désigner un examen passé sur l'écran d'un appareil informatique ou un examen sur papier.  |
| <b>Épreuve modifiée</b>                                    | Modifications apportées à une évaluation pour permettre aux élèves ayant des besoins particuliers de réaliser cette évaluation sur un pied d'égalité avec les autres élèves n'ayant pas les mêmes besoins. Il peut s'agir, par exemple, de modifier la taille ou le style de la police de caractères. Ces ajustements ne doivent pas modifier la nature de la question posée. |
| <b>Évaluation</b>  | Toute tâche effectuée par des élèves pour montrer leurs capacités. Ces tâches peuvent être produites et traitées par l'établissement scolaire ou être envoyées à l'IB, et consistent en un recueil de preuves permettant d'émettre des jugements sur l'apprentissage et l'enseignement.   |
| <b>Évaluation critériée</b>                                | Processus d'évaluation selon lequel les niveaux sont déterminés à l'aide de critères précédemment convenus. La norme de notation est donc fixe et non variable, et ne dépend pas des accomplissements de la cohorte dans son intégralité.   |
| <b>Évaluation électronique</b>                             | Évaluation du PEI comptant trois éléments : l'évaluation sur ordinateur, le portfolio électronique et le projet personnel.  |
| <b>Évaluation externe</b>                                  | Évaluation organisée et notée par l'IB, et non par l'enseignant ou enseignante de l'élève.  |
| <b>Évaluation finale</b>                                   | Évaluation sommative du travail à la fin de la période d'enseignement.  |
| <b>Évaluation formative</b>                                | Évaluation continue destinée à fournir des informations qui orienteront l'enseignement et permettront à l'élève d'améliorer son travail.  |

| Terme   | Définition   |
|---|--|
| <b>Évaluation interdisciplinaire</b>                  | Combinaison ou regroupement d'au moins deux domaines d'apprentissage ou champs d'étude scolaire dans une même évaluation. Dans le Programme du diplôme, une matière interdisciplinaire se conforme aux exigences de deux groupes par l'entremise d'une seule matière. Dans le cadre du PEI, l'étude interdisciplinaire peut être développée au sein des groupes de matières et entre eux. L'évaluation externe interdisciplinaire du PEI fait toujours intervenir plusieurs groupes de matières.   |
| <b>Évaluation interne (EI)</b>                        | Évaluation effectuée par les enseignantes et enseignants au sein de l'établissement.   |
| <b>Évaluation sommative</b>                           | Évaluation visant à déterminer la capacité ou le niveau atteint par les élèves, qui est généralement effectuée à la fin d'un programme d'études ou d'une unité de travail.   |
| <b>Évalué en externe</b>                              | Travail entièrement évalué par l'IB.   |
| <b>Évalué en interne</b>                              | Travail qui est évalué par l'enseignant ou enseignante des élèves. L'IB demande aux établissements d'envoyer un échantillonnage de travaux évalués en interne pour la révision de notation.  |
| <b>Examen</b>   | Forme d'évaluation comportant une ou plusieurs tâches de types divers (questions à réponse brève, questions à réponse développée, exercices de résolution de problèmes ou questions d'analyse ; parfois, exercices pratiques ou tâches orales) que les élèves doivent effectuer individuellement, et ce, en un temps donné et sous surveillance étroite.   |
| <b>Examen sur ordinateur du PEI</b>                   | Évaluation numérique formelle, en temps limité et s'appuyant sur de nombreux médias, qui est spécifiquement conçue pour le PEI. Ce type d'examen est réalisé dans un environnement numérique sécurisé et consiste en des tâches spécifiques aux matières que les élèves effectuent sur un appareil.<br><br>Une connexion à Internet n'est pas requise pour faire passer les examens, mais elle est vivement conseillée pour envoyer les réponses en ligne. Elle permet en outre d'enregistrer automatiquement les réponses des élèves et de réduire la charge administrative. La plupart des établissements choisissent de faire passer les évaluations en ligne pour ces raisons. |
| <b>Examens numériques</b>                             | Concernant le PEI, des examens sur ordinateur sont organisés depuis plus de dix ans. Ils sont conçus spécifiquement pour un environnement numérique et comprennent des tâches interactives s'appuyant sur de nombreux médias, qui sont effectuées dans un environnement d'examen sécurisé.<br><br>S'agissant du Programme du diplôme et du POP, les examens numériques consistent en des évaluations officielles de dernière année, passées à l'aide d'un appareil dans des conditions d'examen sécurisées. Ce système sera progressivement adopté, en reproduisant numériquement les examens traditionnels sur papier.  |
| <b>Examineur ou examinatrice</b>                      | Personne qui attribue des points au travail produit par l'élève pour l'évaluation externe.   |
| <b>Examineur ou examinatrice en chef</b>              | Personne la plus haut placée dans la hiérarchie de la correction des examens, à qui il incombe de veiller au respect des normes de notation au fil du temps et entre les disciplines d'un groupe de matières.  |
| <b>Examineur principal ou examinatrice principale</b> | Personne chargée de diriger l'évaluation dans une composante. Elle établit la norme d'évaluation et participe aussi, en général, à la rédaction de l'évaluation.   |

| Terme   | Définition   |
|---|--|
|   | Dans le cadre du PEI, son rôle diffère légèrement du même rôle dans les autres systèmes d'examen. Cette personne est responsable d'une discipline donnée et est chargée de diriger l'équipe de conception de l'évaluation, d'établir et de faire respecter les normes de notation, et de servir de mentor aux chefs d'équipe.  |
| <b>Examineur superviseur ou examinatrice superviseuse</b> | Examineur ou examinatrice ayant de l'expérience qui aide l'examineur principal ou examinatrice principale.   |
| <b>Facilité d'application</b>                             | Mesure dans laquelle toute évaluation ou tâche constitue une charge pour l'élève ou l'établissement. La longueur des évaluations, l'équipement ou le matériel requis pour réaliser l'évaluation ou le nombre d'évaluations requises dans un programme sont autant d'exemples de facteurs devant être pris en considération pour déterminer la facilité d'application.  |
| <b>Facteur de révision de notation</b>                    | Coefficient mathématique appliqué aux notes (par critère) d'un enseignant ou d'une enseignante pour les rendre conformes à la norme définitive.  |
| <b>Familiarisation</b>                                    | Processus par lequel les élèves apprennent à naviguer dans l'environnement propre aux examens numériques de l'IB et à se servir des outils associés grâce à une simulation des systèmes exempte de contenu. Il incombe aux établissements scolaires de veiller à ce que les élèves aient régulièrement accès aux outils de familiarisation en amont de leurs examens numériques.<br><br>Les environnements de familiarisation sont disponibles sur le Centre de ressources pédagogiques à la fois pour le PEI et pour le Programme du diplôme et le POP. |
| <b>Fiabilité</b>  | Mesure dans laquelle l'élève se verra attribuer la même note chaque fois que son travail sera évalué. Cette fiabilité peut concerner la notation d'examineurs ou examinatrices multiples ou celle d'une seule et même personne.  |
| <b>Fichier de réponses de l'élève</b>                     | Dans le cadre des examens sur ordinateur du PEI, une « réponse d'élève » fait référence au travail de l'élève – éventuellement sous forme multimédia – qui est réalisé lors de l'examen. Le processus de création de la réponse fait partie de l'objet évalué. La réponse de l'élève est envoyée sous forme de fichier électronique, d'où le nom usuel de « fichier de réponses de l'élève ».  |
| <b>Fraude</b>   | Toute pratique contrevenant aux principes de l'intégrité intellectuelle (par exemple, le plagiat).   |
| <b>Groupe d'éléments de question</b>                      | Une ou plusieurs questions connexes au sein d'une épreuve d'examen considérées comme un groupe pour la notation. Les examinateurs et examinatrices notent ce groupe d'éléments de question plutôt que l'examen dans son intégralité. Cette approche permet une notation plus fiable que les approches qui consistent à noter la copie dans son ensemble.   |
| <b>Inscription des élèves</b>                             | Procédure entreprise par un coordonnateur ou une coordonnatrice de programme pour inscrire des élèves à une évaluation de l'IB.  |
| <b>Intégrité intellectuelle</b>                           | Ensemble de valeurs et de compétences, qui encourage l'intégrité personnelle et les bonnes pratiques en matière d'apprentissage, d'enseignement et d'évaluation.   |
| <b>Jugement</b>   | Comparaison faite entre le travail de l'élève et un critère d'évaluation.  |
| <b>Langue d'usage</b>                                     | Langue dans laquelle l'élève effectue l'évaluation.  |

| Terme                                   | Définition  |
|---|---|
| <b>Langues de travail</b>               | Langues utilisées par l'IB pour communiquer avec ses parties prenantes et dans lesquelles il s'engage à fournir toute une gamme de services pour la mise en œuvre de ses programmes. Actuellement, les langues de travail de l'IB sont l'anglais, l'espagnol et le français.  |
| <b>Liste de questions</b>               | Ensemble de questions accompagnées d'informations sur le thème concerné et le degré de difficulté attendu. Les informations fournies dans une liste de questions peuvent être utilisées pour créer les épreuves d'examen. À l'heure actuelle, l'IB n'utilise pas de liste de questions.   |
| <b>Marge de tolérance</b>               | Écart avec la note définitive de l'examinateur principal ou examinatrice principale qui, selon l'IB, est assez minime pour montrer que la notation de l'examinateur ou examinatrice respecte bien la norme. Les marges de tolérance sont nécessaires, car la notation relève du jugement, et même les examinateurs et examinatrices ayant de l'expérience attribueront des notes légèrement différentes en recarrant un même travail. Les marges de tolérance varient en fonction du nombre de points, du type de question et de la matière.  |
| <b>Matériel de soutien pédagogique</b>  | Renseignements complémentaires destinés à aider le personnel enseignant à comprendre les exigences d'un cours de l'IB. Il a pour but de donner des conseils pratiques et de faciliter la compréhension et la mise en œuvre de la théorie présentée dans les guides pédagogiques.  |
| <b>Matière pilote</b>                   | Matière en phase d'essai. En cas de succès, l'accès à cette matière est généralisé.   |
| <b>Mauvaise administration</b>          | Action d'une école du monde de l'IB qui enfreint les règlements de l'IB et peut menacer l'intégrité des examens et des évaluations de l'organisation. Un acte de mauvaise administration peut avoir lieu avant, pendant ou après l'évaluation ou l'examen.  |
| <b>Mauvaise conduite</b>                | Comportement (délibéré ou fortuit) procurant ou susceptible de procurer un avantage déloyal à l'élève qui l'adopte ou à d'autres élèves dans une ou plusieurs composantes d'évaluation. Un comportement susceptible de porter préjudice à un ou une autre élève est également considéré comme une mauvaise conduite.  |
| <b>Mode verrouillage</b>                | Environnement d'examen sécurisé qui empêche les élèves d'accéder aux applications, sites Web et fichiers sans lien avec l'examen numérique pendant son déroulement.<br><br>PEI : la fonctionnalité de verrouillage est intégrée aux examens et s'active automatiquement lorsque l'évaluation est lancée.<br><br>Programme du diplôme et POP : l'appareil se verrouille en ouvrant le navigateur Safe Exam Browser sous Windows et Mac ou en activant le mode Kiosk sur tout Chromebook fourni par l'établissement. Les établissements sont chargés de configurer correctement les appareils des élèves.           |
| <b>Modèle d'assurance de la qualité</b> | Approche adoptée par l'IB pour s'assurer que les élèves reçoivent des résultats corrects à l'issue de l'évaluation. L'examinateur principal ou examinatrice principale établit la norme de notation correcte pour chaque question, et chaque examinateur ou examinatrice doit reproduire cette norme. Pour les évaluations notées en externe, ceci s'effectue en prodiguant des conseils aux examinateurs et examinatrices durant la normalisation, en vérifiant leur compréhension de la norme au moyen de réponses de validation et en contrôlant régulièrement leur notation au moyen de réponses de contrôle. |

| Terme   | Définition  |
|---|---|
| <b>Mot-consigne</b>                               | Terme utilisé dans une question pour expliquer l'objectif d'évaluation évalué.  |
| <b>Niveau</b>                                     | Niveau donné lorsque le travail envoyé pour l'évaluation reflète le descripteur correspondant. Les niveaux sont indiqués dans la colonne de gauche des critères d'évaluation.   |
| <b>Normalisation</b>                              | Processus collaboratif par lequel une norme d'évaluation commune est établie entre les réviseurs et réviseuses de notation ou les examinateurs et examinatrices.  |
| <b>Normalisation interne</b>                      | Processus par lequel l'ensemble des enseignantes et enseignants d'une matière d'un même établissement s'assure d'appliquer la même norme de notation.   |
| <b>Norme</b>                                      | Niveau d'accomplissement attendu pour obtenir une note, une note finale ou un résultat en particulier.  |
| <b>Notation électronique</b>                      | Processus par lequel les examinateurs et examinatrices notent le matériel d'examen directement sur l'écran de leur appareil.  |
| <b>Note définitive</b>                            | Note attribuée par l'examineur principal ou examinatrice principale à un travail en particulier envoyé pour l'évaluation. Il s'agit de la note que tout autre examinateur ou toute autre examinatrice doit chercher à reproduire en corrigeant ce travail. (Voir <b>Modèle d'assurance de la qualité.</b> )   |
| <b>Note finale</b>                                | Description de l'accomplissement de l'élève. Les notes finales pour les travaux envoyés pour l'évaluation vont de 1 (la plus faible) à 7 (la plus élevée). La note finale représente le jugement émis par l'IB sur les qualités générales mises en évidence par l'élève. Elle est cohérente d'une année à l'autre et d'une matière à l'autre. Une note finale décrit la qualité de l'accomplissement de l'élève et elle doit signifier la même chose pour chaque examen, année et matière. Elle tient compte de la difficulté relative de la composante d'une session à l'autre et de tout changement en matière de normes. |
| <b>Objectif spécifique</b>                        | Un des énoncés décrivant les compétences, les connaissances et la compréhension qui seront évaluées.  |
| <b>Outil d'évaluation</b>                         | Méthode de recueil d'informations sur l'accomplissement et la compréhension de l'élève.   |
| <b>Outil de familiarisation</b>                   | Simulation sans contenu de l'environnement propre aux examens numériques. Cette version permet aux élèves de s'entraîner à naviguer sur la plateforme, en explorant les outils et en comprenant la présentation des examens, sans aucun contenu spécifique à la matière.  |
| <b>Outil de vérification de la bande passante</b> | Outil en ligne utilisé par les membres de la direction des établissements proposant le Programme du diplôme et le POP pour confirmer que la bande passante Internet de leur structure est suffisante pour prendre en charge un grand nombre d'élèves passant simultanément les examens numériques du Programme du diplôme et du POP. Pour s'assurer d'une configuration optimale, il est recommandé d'utiliser cet outil bien avant les sessions d'examens.   |
| <b>Pertinence du construit</b>                    | Mesure dans laquelle l'évaluation permet effectivement de contrôler les compétences et les connaissances qu'elle est supposée évaluer.  |
| <b>Plagiat</b>                                    | Présentation, intentionnelle ou non, des idées, des propos ou du travail d'une autre personne sans mentionner correctement, clairement et explicitement les sources correspondantes.  |

| Terme   | Définition   |
|---|--|
| <b>Point(s)</b>   | Valeur numérique reflétant la qualité de la réponse de l'élève dans une tâche particulière. Les points sont attribués en fonction de critères, de bandes de notation ou d'un barème de notation pour indiquer dans quelle mesure la réponse est correcte. L'attribution des points varie en fonction des questions et des examens.   |
| <b>Portfolios électroniques</b>                                 | Système ou processus permettant aux établissements de charger pour les élèves du PEI les examens ou les travaux de cours évalués en interne afin que la notation soit révisée en externe par l'IB.   |
| <b>Prévisibilité</b>  | Capacité à évaluer ce qui se passera ou à quel moment une chose se passera. Dans le contexte de l'évaluation, cela fait référence à l'aptitude des établissements scolaires à anticiper les questions qui seront posées dans une épreuve d'examen et à quel moment.  |
| <b>Procédure pour l'attribution d'une note finale manquante</b> | Mécanisme visant à attribuer une note finale à un ou une élève du PEI lorsque l'IB ne peut calculer une note finale exacte ou juste à partir du travail effectué par l'élève en question. Ce processus est pertinent dans les cas où le manque de preuves est dû aux actions de l'IB ou de tiers (à l'exclusion de l'établissement) et où il ne serait pas raisonnable de demander à l'élève de recommencer l'évaluation.              |
| <b>Procédure pour note manquante</b>                            | Mécanisme visant à attribuer une note à un ou une élève du Programme du diplôme ou du POP lorsque l'IB ne peut calculer une note exacte ou juste à partir du travail effectué par l'élève en question. Ce processus est pertinent dans les cas où le manque de preuves est dû aux actions de l'IB ou de tiers (à l'exclusion de l'établissement) et où il ne serait pas raisonnable de demander à l'élève de recommencer l'évaluation. |
| <b>Programme « Apportez votre appareil personnel »</b>          | Pratique selon laquelle les élèves utilisent leurs appareils personnels pour effectuer un travail scolaire. Les établissements scolaires donnent les indications nécessaires pour que les parents et les personnes exerçant une tutelle puissent se procurer un appareil adapté à un usage scolaire.   |
| <b>Publication des résultats</b>                                | Processus par lequel les élèves reçoivent les notes finales attribuées par l'IB en fonction de leurs évaluations.  |
| <b>Question</b>   | Tâche ou activité utilisée pour permettre à l'élève de montrer sa compétence dans une matière.   |
| <b>Question à choix multiple</b>                                | Question dans laquelle l'élève doit choisir la bonne réponse dans une liste de réponses possibles fournie avec la question.  |
| <b>Réclamation concernant les résultats</b>                     | Réexamen de la notation entrepris à la demande d'un établissement.   |
| <b>Recorrection</b>   | Processus consistant à recorriger les réponses des élèves attribuées à un examinateur ou une examinatrice (réponses complètes ou groupes d'éléments de question, en fonction de la structure de l'évaluation) lorsque sa notation est jugée incohérente ou qu'elle s'écarte considérablement de la norme requise. Cela se produit généralement à la suite d'un échec à la révision de notation.  |
| <b>Reprise</b>  | Seconde tentative ou tentative ultérieure consistant à se représenter à un ou plusieurs examens dans l'espoir d'obtenir le certificat du PEI ou le diplôme de l'IB, ou d'augmenter le total des points sur un certificat déjà reçu.  |
| <b>Réponse à l'évaluation</b>                                   | Terme utilisé pour décrire tout matériel produit par l'élève en réponse à une tâche d'évaluation.  |

| Terme                                  | Définition   |
|--|--|
| <b>Réponse atypique</b>                | Réponse à une tâche qui est très différente des réponses généralement reçues. Il peut s'agir, par exemple, d'un travail incomplet, d'un travail non conforme, d'un travail contenant des réponses inattendues, d'un travail problématique ou d'une fraude.   |
| <b>Réponse d'entraînement</b>          | Exemple de travail envoyé pour l'évaluation qui est choisi et noté pendant la normalisation, puis fourni aux examinateurs et examinatrices afin de leur présenter la norme de notation à adopter.  |
| <b>Réponse de contrôle</b>             | Copie qui a été notée par l'examineur principal ou examinatrice principale et qui est incluse de manière aléatoire dans un paquet de copies assignées à un examinateur ou une examinatrice pour notation. Les points attribués à cette réponse de contrôle par l'examineur ou examinatrice sont comparés à ceux attribués par l'examineur principal ou examinatrice principale afin de s'assurer que l'examineur ou examinatrice respecte la norme de notation établie. Dans le cadre du processus de révision de notation, des réponses de contrôle sont incluses de la même façon dans l'échantillonnage dynamique.                |
| <b>Réponse de validation</b>           | Exemple de travail envoyé pour l'évaluation qui est sélectionné par l'examineur principal ou examinatrice principale, puis utilisé pour vérifier que les examinateurs et examinatrices ont compris la norme de notation à appliquer avant de noter les copies d'examen de la session en cours.   |
| <b>Réponses aux examens numériques</b> | Le travail qui est réalisé dans le cadre des évaluations numériques de l'IB porte un nom différent en fonction du programme concerné.<br>S'il s'agit du PEI, ce travail est qualifié de « fichier de réponses de l'élève ».<br>Concernant le Programme du diplôme et le POP, c'est l'appellation « réponse de l'élève » qui a été choisie.<br>Ces deux termes font référence au produit numérique final qui est réalisé lors de l'examen et envoyé pour l'évaluation. Il s'agit de l'équivalent d'une « copie » dans un examen sur papier.   |
| <b>Résultats de cours</b>              | Principal document de l'IB présentant les résultats aux évaluations. Ce document montre chaque discipline étudiée par l'élève ainsi que la note finale obtenue (allant de 1 à 7 ou de A à E, selon les cas). Il présente également les résultats obtenus à d'autres évaluations pertinentes, notamment la note finale obtenue pour les composantes du tronc commun.<br>Enfin, il indique le nom et le code de l'élève, le numéro de session, la session au cours de laquelle les notes ont été attribuées, la date de délivrance et le nom de l'établissement ayant inscrit l'élève (et s'il s'agit d'un duplicata, le cas échéant). |
| <b>Réunion de normalisation</b>        | Réunion tenue par les examinateurs principaux et examinatrices principales afin de décrire la norme de notation requise et de produire des réponses de contrôle.   |
| <b>Révision de notation</b>            | Processus destiné à s'assurer qu'une norme de notation commune est utilisée. Il consiste notamment à réviser un échantillonnage de travaux envoyés pour l'évaluation et à ajuster, le cas échéant, les notes initiales.  |
| <b>Révision de notation externe</b>    | Voir <b>Révision de notation</b> .   |
| <b>Session</b>                         | Voir <b>Session d'examens</b> .  |
| <b>Session d'examens</b>               | Période pendant laquelle les examens sont passés et notés. Chaque année, l'IB organise deux sessions d'examens : une en mai et une en novembre.  |

| Terme   | Définition  |
|---|---|
| <b>Seuil d'attribution des notes finales</b>            | Point où l'accomplissement de l'élève passe d'une note finale à une autre. Il est souvent utilisé pour indiquer le total des niveaux par critère le plus bas ou le plus haut pour l'attribution d'une note finale donnée.   |
| <b>Spécimens d'examens</b>                              | Exemples d'examens de l'IB, complets et spécifiques aux matières. Ces spécimens présentent du contenu d'évaluation réel et rendent compte du format et du niveau des évaluations finales. Ils sont généralement utilisés pour aider les élèves à s'entraîner dans des conditions semblables aux examens.<br>PEI : des spécimens d'épreuves sont disponibles sur le Centre de ressources pédagogiques.<br>Programme du diplôme et POP : des spécimens seront prochainement disponibles sur le Centre de ressources pédagogiques. Concernant les examens numériques, un accès supplémentaire sera possible en se connectant au système d'examens numériques.  |
| <b>Stratégie d'évaluation</b>                           | Méthode ou approche adoptée pour recueillir des informations relatives à l'apprentissage des élèves (par exemple, des observations, des tâches ouvertes ou des réponses choisies).  |
| <b>Surveillant ou surveillante d'examen</b>             | Personne qui surveille l'environnement dans lequel se déroule l'examen et aide au maintien de sa sécurité.  |
| <b>Surveillant principal ou surveillante principale</b> | Programme du diplôme et POP : personne chargée de la surveillance des examens numériques dans l'établissement, qui doit notamment effectuer les tâches administratives requises dans le système d'examens numériques et régler tous les problèmes techniques avant et pendant la période d'examens. Cette fonction peut être occupée à condition de faire partie du personnel de l'établissement et de se voir attribuer ce rôle dans Mon IB par le coordonnateur ou la coordonnatrice du programme.<br>Afin de garantir le bon déroulement des examens, cette fonction est à dissocier de la coordination du programme, tout en assurant une collaboration étroite. Il peut arriver que la même personne assure à la fois la coordination du programme et la surveillance principale des examens.<br>PEI : ce rôle est facultatif. Le coordonnateur ou la coordonnatrice du programme peut décider de nommer quelqu'un pour effectuer les tâches administratives en rapport avec les examens sur ordinateur. |
| <b>Système d'examens numériques</b>                     | Système en ligne utilisé par les établissements scolaires pour organiser les examens numériques du Programme du diplôme et du POP.  |
| <b>Système d'information de l'IB (IBIS)</b>             | Site Web permettant aux coordonnateurs et coordonnatrices de programme d'accomplir des tâches administratives et d'obtenir des informations de l'IB par l'intermédiaire d'un serveur protégé par mot de passe.  |
| <b>Tâche d'évaluation</b>                               | Activité ou série d'activités auxquelles participent les élèves en vue de l'évaluation.   |
| <b>Test</b>   | Ce terme peut être utilisé au sens générique pour désigner un examen, même si les articles de recherche sur l'évaluation peuvent lui conférer un sens spécifique.<br>Parfois, l'évaluation globale de l'élève est divisée en plusieurs éléments distincts, passés à des moments différents. Ces éléments portent le nom de « composantes d'évaluation » ou de « composantes », et peuvent prendre la forme d'examens oraux, d'épreuves d'examen individuelles ou de tâches d'évaluation interne.  |

| Terme   | Définition   |
|---|--|
| <b>Totalisation</b>                             | Processus consistant à additionner les points et les notes pour obtenir un résultat final.   |
| <b>Validité</b>                                 | Terme décrivant l'adéquation d'une évaluation ou du but pour lequel les résultats de l'évaluation sont utilisés.   |
| <b>Vérificateur de compatibilité</b>            | Application autonome utilisée dans le cadre des examens sur ordinateur du PEI et des examens numériques du Programme du diplôme et du POP pour confirmer qu'un appareil répond aux exigences techniques minimales pour faire passer un examen numérique. Le vérificateur doit être exécuté sur chaque appareil destiné à servir pour l'évaluation. Aucun examen ne doit être passé sur un appareil non validé par le vérificateur de compatibilité.  |
| <b>Vérificateur de la configuration requise</b> | Outil utilisé pour vérifier qu'un appareil respecte les exigences techniques pour exécuter les examens numériques du Programme du diplôme et du POP. Il aide les établissements scolaires à tester les appareils en avance et à confirmer qu'ils pourront assurer une expérience de l'évaluation en toute fluidité et sécurité.<br><br>Ce vérificateur est utilisé lors de la préparation des examens, et non le jour des examens, en complément des informations publiées au sujet de la configuration requise. |

## Ressources imprimables

Tout au long de cette publication, des références sont régulièrement faites à des ressources imprimables. Pour pouvoir les consulter plus facilement, toutes ces ressources sont énumérées ci-après.

- « Les protagonistes du cycle d'évaluation et leurs responsabilités » (PDF)
- « La taxonomie de Bloom » (PDF)
- « Le développement d'un état d'esprit éthique » (PDF)
- « Mise en œuvre des aménagements à des fins d'accès et d'inclusion » (PDF)
- « Notation des travaux d'évaluation interne : le rôle du personnel enseignant dans le processus de révision de notation » (PDF)
- « La révision de notation reposant sur l'échantillonnage dynamique » (PDF)
- « Notes finales prévues : guide à l'intention du personnel enseignant » (PDF)
- « Éventail de copies évaluées en interne requis par le modèle d'assurance de la qualité » (PDF)
- « Évaluation sommative et formative » (PDF)
- « Les principes de l'IB en matière d'évaluation » (PDF)
- « La chaîne de validité » (PDF)