

# **Interprétation et utilisation des résultats des évaluations et des tests provinciaux**

Un document d'appui à  
l'intention des enseignants,  
des administrateurs  
et des conseillers

---

***Interprétation et utilisation  
des résultats des évaluations et  
des tests provinciaux***

*Un document d'appui à l'intention  
des enseignants, des administrateurs  
et des conseillers*

**2009**  
Éducation Manitoba

---

## Données de catalogage avant publication – Éducation Manitoba

371.26'097127 Interprétation et utilisation des résultats des évaluations et des tests provinciaux : un document d'appui à l'intention des enseignants, des administrateurs et des conseillers.

Bibliographie : p. Références – 1

ISBN-13 : 978-0-7711-4037-2

1. Tests et mesures en éducation. 2. Tests et mesures en éducation — Manitoba. 3. Élèves — Évaluation. 4. Rendement scolaire — Évaluation. I. Manitoba. Éducation Manitoba.

Tous droits réservés © 2009, le gouvernement du Manitoba, représenté par le ministre de l'Éducation.

Éducation Manitoba  
Division des programmes scolaires  
Winnipeg (Manitoba), Canada

Nous nous sommes efforcés de citer les sources originales et de respecter la Loi sur le droit d'auteur. Si vous décelez des cas où cela n'a pas été fait, veuillez en aviser le ministère de l'Éducation du Manitoba. Les erreurs et omissions seront corrigées dans une édition future.

Les références aux sites Web citées dans ce document sont sous réserve de modifications.

Vous pouvez commander des exemplaires imprimés de cette ressource (numéro d'article 96516) du Centre des manuels scolaires du Manitoba au : <[www.mtbb.mb.ca](http://www.mtbb.mb.ca)>.

Le présent document est également affiché sur le site Web du ministère de l'Éducation du Manitoba, au : <[www.edu.gov.mb.ca/m12/eval/publications.html](http://www.edu.gov.mb.ca/m12/eval/publications.html)>.

Note: Dans le présent document, la forme masculine a été employée dans le seul but d'alléger le texte.

*This document is available in English.*

## **Table des matières**

---

**Introduction** 1

**Politique d'évaluation des élèves de 3<sup>e</sup> année** A1

But A1

Description A1

Données recueillies A2

Utilisation des résultats A2

Mises en garde A4

**Politique d'évaluation des élèves des années intermédiaires** B1

But B1

Description B1

Données recueillies B2

Utilisation des résultats B3

Mises en garde B5

**Tests basés sur les normes de 12<sup>e</sup> année** C1

But C1

Description C1

Données recueillies C2

Utilisation des résultats C2

Mises en garde C5

**Annexe 1 : Explication des termes** 1

**Annexe 2 : Exactitude et cohérence de la correction** 1

**Annexe 3 : Gestion des observations aberrantes en statistique** 1

**Références** 1

## Introduction

---

Le ministère de l'Éducation (« Ministère ») du Manitoba s'est engagé à mettre en œuvre des politiques et des pratiques d'évaluation favorisant un meilleur apprentissage des élèves. L'évaluation\* du progrès des élèves vers l'atteinte des résultats d'apprentissage visés fournit des renseignements cruciaux qui aident les enseignants, les parents\*\*, les administrateurs et les élèves eux mêmes à planifier l'amélioration de l'apprentissage. L'évaluation fait partie intégrante du processus d'apprentissage et doit se faire avant, pendant et après les activités d'enseignement. Bien des moyens sont employés, comme le dialogue, l'observation en classe, les portfolios, les tests préparés par l'enseignant et les évaluations à grande échelle. Puisque l'évaluation sous toutes ses formes joue un rôle général d'information dans le processus d'enseignement, il est crucial que la façon d'utiliser l'information issue de ces divers types d'évaluations soit compatible avec le but et la portée de la stratégie d'évaluation employée. (Ces enjeux sont abordés dans le document provincial *Repenser l'évaluation en classe en fonction des buts visés : l'évaluation au service de l'apprentissage, l'évaluation en tant qu'apprentissage, l'évaluation de l'apprentissage.*)

Les données d'évaluation peuvent être utilisées à de multiples fins, comme pour fournir une rétroaction aux élèves, planifier l'enseignement, confirmer le rendement, communiquer avec les parents, évaluer les programmes et allouer des ressources. L'information recueillie par diverses méthodes d'évaluation peut aider les enseignants, les élèves et les administrateurs à échanger et à mettre l'accent sur la réussite des élèves, ainsi qu'à planifier et à surveiller leurs progrès. Toutefois, pour utiliser l'information recueillie lors d'évaluations, on doit savoir dans quel but cette évaluation a été conçue et déterminer la fiabilité et la validité des données obtenues ainsi que leur interprétation.

Le présent document porte principalement sur les données résultant des politiques d'évaluation provinciales. **Son but est d'aider les enseignants, les administrateurs et les conseillers à interpréter et à utiliser les données obtenues suite à des évaluations provinciales lorsqu'ils planifient l'enseignement ou évaluent la programmation des écoles et de la division.** À mesure qu'évoluent les politiques d'évaluation provinciales, il pourrait y avoir de nouvelles sections ajoutées à ce document pour favoriser l'utilisation appropriée des données qui en découlent.

---

\* L'annexe 1 fournit une explication de certains termes utilisés en évaluation.

\*\* Dans le présent document, le terme « parents » s'applique aux parents et aux tuteurs. Le terme « parents » est utilisé tout en reconnaissant que dans certains cas seul un parent prend part dans l'éducation de l'enfant.

Il convient de noter que l'on fait référence à des politiques provinciales. Comme les politiques sont susceptibles d'être changées, le lecteur doit vérifier si les énoncés de politiques sont encore valables et demander des explications ou des instructions en cas de doute ou de modification.

Bien que le présent document porte avant tout sur les données obtenues en conformité avec les politiques d'évaluation provinciales, les données provenant d'autres sources, surtout d'évaluations faites au jour le jour en salle de classe, revêtent la plus grande importance aux fins de l'orientation de l'enseignement et de la diffusion des résultats. Incidemment, bon nombre des idées exprimées dans ce document peuvent s'appliquer à d'autres types d'évaluation.

## **Politique d'évaluation des élèves de 3<sup>e</sup> année**

### **But**

Le but principal de la Politique d'évaluation des élèves de 3<sup>e</sup> année est de fournir des renseignements aux parents au sujet des connaissances et des habiletés de base de leurs enfants en lecture et en notions de calcul au début de la 3<sup>e</sup> année, et en lecture en français au début de la 4<sup>e</sup> année pour les élèves en immersion française. À partir de ces données, les enseignants et les parents pourront discuter, réfléchir, planifier et collaborer ensemble afin d'améliorer l'apprentissage de l'enfant\*.

### **Description**

L'Évaluation des élèves de 3<sup>e</sup> année se fait en salle de classe et permet d'évaluer les connaissances et les habiletés de base des élèves. La politique provinciale décrit les compétences de base à évaluer, mais les enseignants choisissent ou élaborent diverses stratégies afin d'évaluer le rendement des élèves. Les données sont recueillies par l'enseignant de la 3<sup>e</sup> année (ou par l'enseignant de la 4<sup>e</sup> année dans le programme d'immersion française) et par celui de la 2<sup>e</sup> année ou par des spécialistes du milieu scolaire, au besoin. Les enseignants disposent d'un continuum (fourni par le Ministère dans le document de politique) pour la lecture et les notions de calcul pour évaluer la performance des élèves.

Comme l'évaluation est intégrée aux activités quotidiennes en salle de classe, les enseignants exercent leur jugement professionnel en s'inspirant du continuum et des programmes d'études, afin de déterminer les méthodes de collecte des données d'évaluation. L'exercice du jugement professionnel signifie que le Ministère ne s'attend pas, pas plus qu'il ne considère approprié, que tous les élèves soient soumis à une évaluation approfondie de chaque compétence pour les fins visées par cette politique. De plus, le gouvernement provincial ne fournit pas ou ne prescrit pas d'outils ni de stratégies d'évaluation précises. Dans bien des cas, les enseignants sont en mesure de préparer les rapports requis sur les niveaux de performance des élèves à partir de leurs travaux quotidiens et de leurs conversations avec les élèves. Les enseignants peuvent également utiliser des stratégies afin d'obtenir la participation des élèves à une réflexion sur leur apprentissage.

---

\* On peut consulter le document de politique au [www.edu.gov.mb.ca/m12/eval/eva\\_oblig\\_3-4\\_lec.html](http://www.edu.gov.mb.ca/m12/eval/eva_oblig_3-4_lec.html).

## **Données recueillies**

Les écoles transmettent les données d'évaluation de chaque élève à ses parents pour chaque compétence de base en lecture et en notions de calcul, en indiquant si l'enfant

- a besoin d'un soutien continu
- a besoin d'être guidé pour répondre aux attentes
- répond aux attentes

Les écoles compilent aussi le nombre total d'élèves performant à ces niveaux et envoient les données à leur division qui regroupe tous les résultats et les communique au Ministère (les écoles indépendantes envoient leurs résultats directement au Ministère). Le Ministère prépare un rapport provincial indiquant le pourcentage d'élèves, à la grandeur de la province, qui performant à chacun des trois niveaux pour chaque compétence d'après l'évaluation faite par chaque enseignant de ses élèves.

## **Utilisation des résultats**

### **Enseignants et directeurs d'école**

Les résultats d'évaluation peuvent servir de base de réflexion et de discussion entre les éducateurs relativement aux besoins en matière d'apprentissage des élèves.

- Ils aident à la planification globale de l'enseignement en aidant à cerner les tendances relatives au rendement.
- Ils peuvent aider à cerner les secteurs d'enseignement convenant à certains élèves ou petits groupes d'enfants.
- Ils peuvent aider à déterminer la nécessité d'interventions spécialisées afin de répondre à des besoins en matière d'apprentissage particuliers.
- Ils fournissent des renseignements utiles pour la prise de décisions sur l'allocation des ressources ou l'examen de programmes d'intervention.

Le niveau de rendement des élèves reflète en partie leur développement cumulatif et leur rendement qui découlent de l'enseignement des années précédentes. Il convient donc de partager cette information avec les autres membres du personnel de l'école pour favoriser les efforts de collaboration afin d'améliorer l'apprentissage des élèves.

### **Parents**

En connaissant le niveau de performance actuel de l'enfant ainsi que ses forces et ses besoins sur le plan cognitif dans des champs de compétence précis, les parents peuvent contribuer davantage à l'apprentissage chez leur enfant. Les résultats de l'évaluation fournissent un point de départ pour des discussions productives avec l'enseignant.

### **Administrateurs d'école ou de division, et commissaires**

L'examen des données d'évaluation contenues dans les rapports provinciaux, divisionnaires et scolaires permet de

- dégager des tendances générales concernant les niveaux de rendement des élèves dans les compétences prescrites suivant les stratégies choisies par l'enseignant
- utiliser les tendances observées afin d'orienter les discussions entre enseignants et administrateurs concernant les répercussions sur l'amélioration de l'apprentissage des élèves

Les données de l'Évaluation des élèves de 3<sup>e</sup> année peuvent être intégrées aux plans et aux rapports scolaires et divisionnaires à la condition d'indiquer clairement que ces résultats ont pour seul but d'aider à orienter les enseignants dans leur travail en classe avec les élèves et dans leurs discussions avec les parents durant l'année scolaire. Comme le but de l'Évaluation des élèves de 3<sup>e</sup> année est d'éclairer la pratique future de l'enseignement en salle de classe, il faut être prudent lorsqu'on utilise ces résultats à d'autres fins. Voir la section des Mises en garde ci-après.

### **Le Ministère ou autres organismes**

Les tendances observées dans les niveaux de performance des élèves conformément à la Politique d'évaluation des élèves de 3<sup>e</sup> année servent de point de discussion entre le Ministère et les éducateurs, à partir duquel il est possible de répertorier les besoins particuliers pouvant être pris en charge à l'échelle provinciale. Il peut s'agir de plans visant le perfectionnement professionnel, la détermination de programmes d'intervention, l'examen des programmes d'études et l'adaptation de l'enseignement. Par exemple, la subvention provinciale pour intervention précoce en matière d'alphabétisation et la subvention pour les notions de calcul des années primaires sont le résultat de l'analyse des résultats de l'Évaluation des élèves de 3<sup>e</sup> année.

## Mises en garde

Les données recueillies à partir de l'Évaluation des élèves de 3<sup>e</sup> année sont destinées principalement à être utilisées par les enseignants dans leurs communications avec les parents et dans la prise de décisions touchant l'enseignement. Les enseignants qui mènent l'évaluation sont les mieux placés pour interpréter les données en fonction de la méthode d'évaluation et du contexte dans lequel elle s'est déroulée en classe. Parallèlement, l'agrégation et la diffusion des données peuvent donner lieu à des discussions enrichissantes au sujet de l'enseignement au niveau des années primaires, à la condition que les personnes engagées dans la discussion reconnaissent les limites de ces données et les valident en regard des renseignements provenant d'autres sources. Les données de l'Évaluation des élèves de 3<sup>e</sup> année représentent une importante source d'information sur leur rendement parce qu'elles sont recueillies par les enseignants. Toutefois, il ne faut pas les utiliser isolément pour tirer des conclusions importantes sur le rendement global des élèves, comme pour comparer des groupes ou des tendances au fil du temps, mais plutôt s'en servir comme une source de données parmi plusieurs autres.

Comme chaque langue d'enseignement présente des caractéristiques uniques qui ne trouvent pas toujours facilement d'équivalents et que par conséquent il y a des différences entre les programmes d'études, les méthodes d'enseignement et les processus d'évaluation, il n'est pas conseillé de faire des comparaisons entre les divers programmes scolaires.

## **Politique d'évaluation des élèves des années intermédiaires**

---

### **But**

Le but premier de la Politique d'évaluation des élèves des années intermédiaires\* est d'améliorer l'apprentissage et l'engagement des élèves par des méthodes d'évaluation en classe qui sensibilisent les élèves et leur donnent confiance en matière d'apprentissage. La politique est conçue de manière à permettre aux enseignants d'appliquer à de telles fins les pratiques d'évaluation formative (*évaluation au service de l'apprentissage* et *en tant qu'apprentissage*).

Le deuxième but est de recueillir des renseignements de nature sommative sur les niveaux de rendement que les élèves des années intermédiaires ont atteints à la fin janvier dans des domaines clés appelés compétences. Ces compétences clés sont le sens des nombres, les habiletés mathématiques et l'engagement en 7<sup>e</sup> année, et la compréhension en lecture et l'écriture de textes informatifs en 8<sup>e</sup> année. Les publics auxquels s'adressent ces renseignements sommatifs sont

- les enseignants, les élèves et les parents qui peuvent utiliser ces renseignements directement et immédiatement dans la prise de décisions au sujet de l'apprentissage en fonction du niveau de chaque élève
- l'école, la division scolaire et le Ministère qui peuvent utiliser les données cumulatives pour faire le suivi des tendances en vue de prendre des décisions plus éclairées en ce qui concerne, par exemple, les ressources à prévoir pour favoriser l'amélioration de l'apprentissage des élèves

### **Description**

L'évaluation des élèves des années intermédiaires se fait en salle de classe et permet d'évaluer certaines compétences clés. La politique provinciale décrit les compétences à évaluer, mais les enseignants choisissent ou élaborent diverses stratégies afin d'évaluer le rendement des élèves tout en cherchant à les faire participer au processus (*évaluation au service de l'apprentissage* et *en tant qu'apprentissage*). Les enseignants utilisent l'information ainsi obtenue (au cours des premiers mois de l'année scolaire) pour déterminer le niveau de

---

\* On peut consulter le document de politique *Évaluation des compétences de base en mathématiques, en compréhension en lecture, en écriture de textes informatifs et sur l'engagement de l'élève au niveau des années intermédiaires* à l'adresse [www.edu.gov.mb.ca/frpub/me/enonce\\_pol/index.html](http://www.edu.gov.mb.ca/frpub/me/enonce_pol/index.html).

rendement le plus récent de chaque élève aux fins de la préparation de rapports (à l'intention des parents et du Ministère) lors des deux dernières semaines de janvier. Les enseignants établissent les niveaux en se servant des indicateurs de rendement à mi-parcours, illustrés par des copies types de travaux d'élèves, qui se trouvent dans les documents d'appui fournis par le Ministère.

Comme l'évaluation est intégrée aux activités quotidiennes d'apprentissage en salle de classe, les enseignants exercent leur jugement professionnel en s'inspirant des documents d'appui et des programmes d'études, afin de déterminer les méthodes de collecte des données d'évaluation. Cette étape devrait se dérouler au cours des premiers mois de l'année scolaire. Ce processus d'évaluation repose sur l'utilisation de la rétroaction descriptive et l'établissement d'objectifs et de critères avec la participation des élèves, ce qui les aide à progresser dans les champs de compétence et ce qui donne finalement lieu à la production d'un rapport sommatif à l'intention des parents sur le rendement de chaque élève en date des deux dernières semaines de janvier selon les données les plus récentes et les plus stables sur la performance. Les écoles font également rapport au Ministère des données sur la performance pour chaque élève. Le Ministère communique aux divisions scolaires un sommaire de la rétroaction des divisions pour les aider dans leur utilisation des données aux fins de rapports et d'analyses.

Cette évaluation s'applique à tous les élèves de la 7<sup>e</sup> année et de la 8<sup>e</sup> année inscrits dans les divisions scolaires et les écoles indépendantes subventionnées quel que soit le programme spécifique en place. Ce n'est que dans certains cas exceptionnels où l'on ne peut pas évaluer les élèves (p. ex. arrivée tardive dans l'école en provenance d'une autre province) qu'il peut y avoir des exemptions.

### **Données recueillies**

Les écoles transmettent les données d'évaluation de chaque élève à ses parents pour les compétences en mathématiques à la 7<sup>e</sup> année et en langues à la 8<sup>e</sup> année en fonction des niveaux de performance suivants :

- n'atteint pas le niveau de performance à mi-parcours de la 7<sup>e</sup>/8<sup>e</sup> année
- atteint presque le niveau de performance à mi-parcours de la 7<sup>e</sup>/8<sup>e</sup> année
- atteint le niveau de performance à mi-parcours de la 7<sup>e</sup>/8<sup>e</sup> année

Les rapports à l'intention des parents comportent des cases de commentaires que les enseignants et les élèves peuvent utiliser pour communiquer des informations contextuelles sur le rendement et les plans d'apprentissage, par exemple, les résultats d'apprentissage.

Les élèves qui sont au niveau « atteint... » ont soit atteint, soit dépassé les niveaux de performance à mi-parcours pour la compétence. Les élèves qui sont au niveau « atteint presque... » ont besoin d'un supplément d'attention, de soutien ou d'effort pour atteindre les résultats d'apprentissage de fin d'année. Les élèves qui sont au niveau « n'atteint pas... » à mi-parcours risquent de ne pas atteindre les résultats d'apprentissage de fin d'année sans une intervention substantielle destinée à renforcer les connaissances et les habiletés associées à ces compétences.

Dans le cas de certains élèves, les niveaux de performance à mi-parcours se situeront sous le niveau le plus bas des rapports (sous « n'atteint pas... »), le plus souvent en raison d'une déficience cognitive ou d'une situation d'apprentissage d'une langue additionnelle, auxquels cas un plan d'apprentissage personnalisé est prévu. Dans de tels cas, on utilise seulement les cases de commentaires des rapports pour communiquer avec les parents au sujet du contexte de l'élève et de son progrès en regard des compétences.

Pour les compétences liées à l'engagement de l'élève (7<sup>e</sup> année), les niveaux figurant sur les rapports sont

- en émergence (manifeste parfois la compétence)
- en croissance (manifeste très souvent la compétence)
- acquis (manifeste presque toujours la compétence)

Dans les cas où l'on observe une variation importante quant à l'engagement de l'élève envers son apprentissage (pour une ou plusieurs matières), on se sert de la catégorie *variable*.

## Utilisation des résultats

### Enseignants et directeurs d'école

Conformément au but premier de la politique, l'un des résultats clés à obtenir est l'augmentation graduelle de la conscientisation des élèves et de leur sens des responsabilités par rapport à l'apprentissage, au moyen de pratiques d'évaluation formative. Les résultats obtenus pour la compétence *engagement de l'élève* en 7<sup>e</sup> année peuvent servir d'indicateur important de l'efficacité des pratiques d'évaluation formative rattachées à la politique. Voici d'autres indicateurs :

- améliorations quant à la réussite de l'élève
- amélioration générale relativement à la dynamique de la classe, par exemple, présence plus régulière, comportement moins perturbateur, plus de temps consacré au travail et plus grande coopération entre les élèves

Un dialogue immédiat et continu entre les élèves et les enseignants au sujet de l'apprentissage des élèves, qui doit découler des pratiques d'évaluation formative, facilite la planification d'expériences d'apprentissage qui favoriseront la réussite.

Les résultats sommatifs sur la performance (transmis aux parents et au Ministère) figurant dans l'évaluation s'avèrent utiles comme outils de réflexion et de discussion pour les raisons qui suivent.

- Les rapports présentés aux parents procurent une autre occasion d'obtenir la participation des parents pour ce qui est d'assurer que leurs enfants développent les compétences clés.
- Après compilation, les résultats fournissent des renseignements utiles pour la prise de décisions sur la planification de l'enseignement, l'affectation des ressources et l'examen des programmes d'intervention.

Le niveau de rendement des élèves reflète en partie leur développement cumulatif et leur rendement suivant l'enseignement des années précédentes. Il convient donc de partager cette information avec tous les autres membres du personnel de l'école et d'en discuter pour favoriser les efforts de collaboration qui améliorent l'apprentissage des élèves.

### **Parents**

En connaissant le niveau de performance actuel de l'enfant en regard des compétences clés, les parents peuvent contribuer davantage à l'apprentissage chez leur enfant. Les résultats de l'évaluation fournissent aussi un point de départ pour des discussions productives avec l'enseignant.

### **Administrateurs d'école ou de division, et commissaires**

L'examen des données d'évaluation contenues dans les rapports provinciaux, divisionnaires et scolaires permet de

- dégager des tendances générales concernant les niveaux de rendement des élèves dans les compétences prescrites
- utiliser les tendances observées afin d'orienter les discussions entre enseignants et administrateurs concernant les répercussions sur l'amélioration de l'apprentissage des élèves

Les données de l'Évaluation des élèves des années intermédiaires doivent être intégrées aux plans et aux rapports scolaires et divisionnaires pour aider à orienter les enseignants dans leur travail en classe avec les élèves et dans leurs discussions avec les parents durant l'année scolaire. Comme le but de l'Évaluation des élèves des années intermédiaires est de promouvoir les pratiques d'évaluation formative et de fournir des renseignements en vue de la prise de décisions touchant l'enseignement, il faut être prudent lorsqu'on utilise ces résultats à d'autres fins. Voir la section des Mises en garde ci-après.

### **Le Ministère ou autres organismes**

Les tendances observées dans les niveaux de performance des élèves servent de point de discussion entre le Ministère et les éducateurs, à partir duquel il est possible de répertorier les besoins particuliers pouvant être pris en charge à l'échelle provinciale. Il peut s'agir de plans visant le perfectionnement professionnel, la détermination de programmes d'intervention et l'examen des programmes d'études.

### **Mises en garde**

Les données recueillies à partir des évaluations des élèves des années intermédiaires sont produites et utilisées par les enseignants pour alimenter la prise de décisions touchant l'enseignement. En ce qui concerne la compilation des données et l'analyse des tendances, il importe de reconnaître que les enseignants peuvent recourir à diverses stratégies d'évaluation qui peuvent influencer sur la façon dont les données devraient être interprétées.

Parallèlement, la compilation et la diffusion des données peuvent donner lieu à des discussions enrichissantes au sujet de l'enseignement au niveau des années intermédiaires, à la condition que les personnes engagées dans la discussion les valident en regard des renseignements provenant d'autres sources. Les données de l'Évaluation des élèves des années intermédiaires représentent une importante source d'information sur leur rendement parce qu'elles sont recueillies par les enseignants. Lorsqu'elles sont correctement pondérées en fonction d'autres sources d'information, ces données peuvent permettre de tirer des conclusions importantes sur le rendement des élèves.

Lorsqu'on compare des groupes linguistiques, il importe de reconnaître que chaque langue présente des caractéristiques uniques et que les programmes d'études, les méthodes d'enseignement et les processus d'évaluation varient en conséquence. Les comparaisons doivent donc être faites avec précaution, particulièrement en ce qui touche les compétences en lecture et en écriture.

## Tests basés sur les normes de 12<sup>e</sup> année

### But

Les tests basés sur les normes sont des évaluations sommatives visant à fournir des informations pertinentes au sujet des connaissances et des habiletés acquises par les élèves à la fin d'un cycle d'enseignement par rapport aux résultats d'apprentissage énoncés dans les programmes d'études provinciaux. Cette information aide à confirmer le rendement des élèves et les enseignants peuvent s'en servir pour évaluer des stratégies d'enseignement afin d'améliorer l'apprentissage des élèves.

### Description

Une description des tests basés sur les normes et du programme d'évaluation se trouve dans le document *Politiques et modalités pour les tests basés sur les normes* et dans les bulletins d'information accessibles dans toutes les écoles de la 12<sup>e</sup> année ainsi que dans Internet\*. Les tests basés sur les normes sont harmonisés avec les programmes d'études dans toute la mesure possible, conformément aux paramètres d'évaluation à grande échelle. Le niveau de performance d'un élève dans un test représente un portrait « instantané » de ses connaissances et de ses habiletés par rapport à un échantillon représentatif de résultats d'apprentissage énoncés dans le programme d'études et visés dans le test.

Les tests basés sur les normes sont élaborés à l'échelle centrale par des enseignants spécialistes de la matière visée et sont mis à l'essai afin de s'assurer que les questions et les tâches reflètent assez fidèlement les résultats d'apprentissage visés (c.-à.-d., la validité du contenu). Les tests sont administrés vers la fin de chaque année scolaire ou du semestre, et ce sont les enseignants qui les corrigent à l'échelle locale selon des barèmes de notation, des grilles d'évaluation et des copies types que fournit le Ministère.

Les écoles communiquent aux élèves la note obtenue au test immédiatement après la correction à l'échelle locale et doivent indiquer sur le bulletin scolaire de l'élève la note obtenue au test basé sur les normes. Les tests basés sur les normes représentent 30 % de la note finale d'un élève dans le cours.

---

\* On peut consulter ces documents à l'adresse  
<[www.edu.gov.mb.ca/m12/eval/publications.html](http://www.edu.gov.mb.ca/m12/eval/publications.html)>

## Données recueillies

Après la correction à l'échelle locale des tests basés sur les normes, les écoles ou les divisions scolaires communiquent les résultats de chaque élève au Ministère qui les compile et retourne des profils de performance pour la province, la division et l'école, indiquant notamment

- la note moyenne
- le taux de réussite
- les notes sur des parties de tests par rapport à un type de question en particulier ou à un groupe de résultats d'apprentissage curriculaires apparentés

Le Ministère met aussi à la disposition de chaque instance (division scolaire ou école indépendante) les données contenant les résultats de chaque élève à chaque question du test. Le Ministère résume également des renseignements au sujet de la participation des élèves tels que le nombre d'élèves exemptés du test et le nombre d'élèves absents et en fait rapport à l'échelle de l'école, de la division et de la province.

Après la correction à l'échelle locale, un échantillon de cahiers de test (à l'exception de Mathématiques du consommateur) recueilli auprès de chaque instance est corrigé à nouveau à l'échelle centrale par le Ministère. Le Ministère fait rapport aux instances scolaires sur l'exactitude et la cohérence de la correction à l'échelle locale afin d'aider les instances à évaluer leurs méthodes de correction à cette échelle. L'annexe 2 fournit de l'information sur l'interprétation des données de ces rapports.

## Utilisation des résultats

### Enseignants

Les résultats aux tests basés sur les normes constituent un indicateur du rendement des élèves en fonction des résultats d'apprentissage et des normes énoncés dans le programme d'études. Par conséquent, il convient de les utiliser pour

- intégrer aux évaluations et aux rapports sur les élèves
- planifier l'enseignement en servant à déterminer les résultats d'apprentissage du programme d'études pour lesquels le rendement des élèves était inférieur ou supérieur
- fournir une base de discussion entre les enseignants et les administrateurs concernant les répercussions sur l'amélioration de l'apprentissage des élèves

Dans ses rapports concernant les tests basés sur les normes en mathématiques, le Ministère inclut des commentaires relatifs aux erreurs communes (à l'échelle provinciale) afin d'appuyer les discussions touchant les résultats aux tests basés sur les normes.

Étant donné que les tests basés sur les normes reflètent les résultats d'apprentissage énoncés dans les programmes d'études et que la correction se fait en fonction de critères (comme les grilles d'évaluation) communs (provinciaux) basés sur le programme d'études et qu'elle est appuyée par des séances de formation offertes par le Ministère, les enseignants peuvent utiliser les résultats à ces tests pour déterminer le degré de cohérence entre leurs normes d'évaluation et celles du programme d'études pour les résultats d'apprentissage examinés dans les tests.

Les enseignants peuvent aussi voir se dessiner des tendances au fil du temps dans les résultats aux tests basés sur les normes (analyse des tendances). Bien que les tests soient différents à chaque occasion d'évaluation, chaque test est préparé en fonction des mêmes spécifications et tous les efforts possibles sont déployés pour assurer un niveau de difficulté général équivalent dans les différents tests. Tout changement apporté aux spécifications de test est signalé dans les bulletins d'information envoyés aux écoles chaque automne.

Il convient de faire des comparaisons entre les groupes et avec les résultats provinciaux (analyse comparative) à la condition que les modalités d'administration et de correction des tests sont scrupuleusement suivies.

### **Élèves ou parents**

En tant que source d'information pertinente sur les connaissances et les habiletés des élèves en fonction d'un échantillon représentatif de résultats d'apprentissage visés, les tests basés sur les normes sont considérés et utilisés à juste titre comme des outils de confirmation des apprentissages, pour la production de rapports ou pour leur intégration à des portfolios. Mais ces utilisations ne sont possibles que si chaque élève peut démontrer en toute liberté ses connaissances et son savoir faire durant le test, et que si les directives concernant la correction sont scrupuleusement suivies.

### **Administrateurs d'école ou de division, et commissaires**

La performance des élèves dans les tests basés sur les normes reflète un rendement et un progrès cumulatifs qui découlent de l'enseignement dispensé les années précédentes ainsi que durant l'année scolaire en cours. En tant que mesures valides et fiables du rendement des élèves, les données regroupées à partir des tests basés sur les normes peuvent être communiquées au personnel de l'école afin de favoriser les efforts de collaboration dans le but d'améliorer l'apprentissage des élèves à l'aide de méthodes appropriées dans le contexte local. Les meilleurs véhicules pour ce faire sont les plans et les rapports scolaires et divisionnaires qui peuvent servir de base de discussions appuyées par des faits sur des aspects tels que l'affectation des ressources.

La rétroaction fournie par le Ministère relativement à la cohérence des corrections (rapport sur le processus de correction) devrait être utilisée par les autorités locales afin d'évaluer les méthodes de correction à l'échelle locale et lors de l'interprétation des résultats aux tests. À titre d'exemple, une instance scolaire peut constater dans ce rapport que ses corrections sont généralement trop sévères, et que cette tendance est plus évidente dans certains types de questions ou dans certains domaines du programme d'études. Cette information est particulièrement importante pour l'analyse comparative et l'analyse des tendances. Il est également important de noter les taux de participation des élèves lorsqu'on examine les données au fil du temps ou dans une évaluation comparative; cette information est résumée par le Ministère et retournée aux instances scolaires.

Les tests basés sur les normes qui sont administrés et corrigés de façon appropriée peuvent servir à

- l'évaluation d'un programme lorsqu'il importe d'avoir des données comparables avant et après la mise en œuvre, par exemple, d'un programme d'intervention
- l'analyse des tendances — l'examen des résultats au fil du temps par rapport aux résultats à l'échelle provinciale
- l'analyse comparative — l'examen des résultats de l'école ou de la division scolaire par rapport aux résultats divisionnaires ou provinciaux afin de déterminer, par exemple, des domaines précis du programme d'études qui sont des forces relatives afin de miser sur ces aspects (en partageant des stratégies d'enseignement, par exemple)

### **Le Ministère ou autres organismes**

Les données obtenues à partir de tests basés sur les normes reflètent le rendement des élèves par rapport à un échantillon de résultats d'apprentissage et de normes du programme d'études. Ces données peuvent servir dans divers contextes où ce genre d'information est nécessaire, notamment l'étude de questions comme les différences liées au sexe ou les variations régionales du rendement, les modèles récurrents et les besoins liés à la mise en œuvre du programme d'études, et la détermination des domaines potentiels dans lesquels un perfectionnement professionnel ou une révision du programme d'études s'impose.

### **Mises en garde**

Les tests basés sur les normes sont administrés et corrigés à l'échelle locale conformément aux modalités prescrites. En cas de non-respect de ces modalités, la fiabilité et la validité des résultats sont compromises.

Les tests basés sur les normes ne représentent qu'une seule mesure du rendement des élèves. Bien que ces tests aient été élaborés avec soin afin d'assurer la concordance avec le programme d'études, ainsi que l'équité et l'accessibilité pour tous les élèves qui ont été correctement classés dans le cours en question, ce ne sont pas tous les élèves qui performant au mieux en situation d'examen final. Pour déterminer exactement tout rendement scolaire individuel, il faut donc utiliser diverses stratégies d'évaluation.

Comme les tests basés sur les normes sont corrigés à l'échelle locale, souvent par les enseignants des élèves, il est important de consulter le rapport sur le processus de correction fourni par le Ministère (aux directeurs généraux de division scolaire et aux directeurs d'école indépendante) lorsqu'on interprète les résultats aux tests pour s'assurer qu'il n'y a pas eu de grave erreur de correction ou de biais majeur. Nous encourageons fortement les écoles et les divisions scolaires qui organisent les séances de correction à s'assurer que

- les tests basés sur les normes sont corrigés en un endroit central, par un groupe d'enseignants formés et travaillant ensemble
- les tests sont distribués au hasard entre les enseignants-correcteurs pour la correction
- des vérifications occasionnelles de l'exactitude des corrections sont menées durant le processus de correction (p. ex., double correction et discussion sur les discordances)
- la correction est faite à simple insu (identité de l'élève protégée par l'anonymat)

Lorsqu'on utilise les résultats aux tests basés sur les normes pour évaluer les profils de rendement au fil du temps, il importe de reconnaître un certain nombre de facteurs atténuants qui altèrent la performance moyenne d'une classe, d'une école ou d'une division. Ces facteurs comprennent les suivants :

- *Toute variation des politiques locales d'exemption des élèves et le taux d'absence d'élèves lors du test* — Les politiques et les modalités d'exemption sont établies à l'échelle provinciale, mais ce sont les écoles qui doivent les appliquer. L'intention de la politique d'exemption est de permettre à des élèves de ne pas participer à un test basé sur les normes lorsqu'ils ont des besoins spéciaux pour lesquels il n'y a pas d'adaptation disponible permettant que l'évaluation se fasse de façon juste et équitable. L'absence d'élèves à un test compromet l'interprétation des résultats regroupés, surtout si ces absences sont peut être des stratégies permettant de se soustraire au test\*.
- *Une légère variabilité du contenu et du niveau de difficulté d'un test à l'autre* — Une variation dans la moyenne d'un groupe allant jusqu'à environ cinq pour cent ne devrait pas être considérée comme étant importante puisque les divers tests ne sont pas précisément équivalents en termes de contenu et de niveau de difficulté. Cependant, si l'on note une augmentation (ou une diminution) constante dans les tests administrés (telle que trois augmentations consécutives de 3 % à 5 %), ou si un changement local (p. ex., à l'échelle de l'école) de la performance va à l'encontre des tendances générales (p. ex., résultats provinciaux), cette variation indique probablement un changement significatif de la performance du groupe.
- *Une variabilité naturelle entre les groupes d'élèves soumis au test à un moment particulier* — Lorsqu'on regroupe et interprète les résultats de groupes d'environ 30 élèves ou moins, on doit tenir compte d'une variabilité « naturelle » aléatoire et incontrôlable de la performance du groupe. Par exemple, la présence même d'un seul élève ayant une note exceptionnellement haute ou basse peut provoquer une grande variation de la moyenne d'un petit groupe qui n'a pas de lien avec des changements (au fil du temps) ou des différences (entre les groupes) dans des facteurs qui peuvent influencer sur la performance. L'annexe 3 fournit d'autres détails sur ces observations aberrantes en statistique.

---

\* Le Ministère prépare des rapports pour les écoles et les divisions scolaires sur les résultats de leur correction à l'échelle locale des tests basés sur les normes, ajoutant des informations sur le nombre et la proportion d'élèves exclus des rapports. Ce rapport se fonde aussi sur des données fournies par les écoles et comprend des points tels que les exemptions et les absences à l'échelle de l'école, de la division et de la province.

- *Une variabilité systémique dans les groupes d'élèves soumis au test à un moment particulier* — Des changements dans les cours offerts ou dans les politiques et les pratiques relatives à l'inscription des élèves aux cours et aux programmes peuvent avoir des répercussions importantes sur les résultats du test. Par exemple, si un cours de mathématiques particulier est offert dans une école durant une année donnée mais pas l'année suivante, il y aura des répercussions sur les tendances en matière d'inscription et les types d'interprétations pouvant être faites au moment d'effectuer les comparaisons de la performance d'une année à l'autre pour les mathématiques.

Il faut utiliser des méthodes statistiques spéciales lorsqu'on compare les moyennes de petits groupes, comme une classe d'élèves. Pour les études de cette nature, votre organisation devrait demander l'aide d'un conseiller en statistiques si elle ne possède pas l'expertise nécessaire.

## Annexe 1 : Explication des termes

---

**Continuum** : Désigne les descripteurs généralement présentés sous forme de tableau et illustrant ce que les apprenants savent et peuvent faire à divers stades de l'apprentissage ou du développement, que ce soit en lecture, en écriture, en mathématiques, etc. Le continuum peut être utilisé pour

- suivre les progrès des élèves de façon continue
- rehausser la fiabilité du suivi du développement des élèves
- aider à la planification et à la différenciation de l'enseignement
- établir des objectifs d'apprentissage
- faire rapport aux parents

Le continuum peut être un continuum de *développement*, ce qui signifie qu'il reflète des caractéristiques développementales observables d'un processus d'apprentissage. D'autres outils similaires mettent l'accent sur certaines compétences au niveau du contenu, qui reflètent une séquence d'apprentissage prescrite ou prévue dans les documents du programme d'études, on parle alors de *portée* et de *séquence*.

**Copie type** : Exemple de travail d'élèves ou documentation des processus utilisés pour préciser les critères d'évaluation des élèves ou pour illustrer des stratégies et clarifier les messages concernant l'évaluation.

**Évaluation\*** : L'évaluation peut avoir deux buts. Un type d'évaluation consiste à générer et à recueillir des données selon des méthodes variées dans un but descriptif. L'autre type consiste à utiliser les données dans le but de prendre des *décisions basées sur des valeurs* au sujet des élèves, comme la décision de faire passer ou échouer un élève, la nécessité de remédiation ou l'énoncé de conclusions au sujet du niveau de rendement d'un élève (p. ex., notes du bulletin scolaire). La **fidélité** et la **validité** sont des considérations importantes peu importe le but de l'évaluation.

*L'évaluation au service de l'apprentissage* se fait tout au long du processus d'apprentissage. Elle comprend les stratégies conçues afin de déterminer les besoins des élèves en matière d'apprentissage, de planifier l'enseignement, de fournir une rétroaction descriptive aux élèves et de suivre de près leurs progrès vers l'atteinte des résultats d'apprentissage visés.

---

\* Remarque : Le mot « évaluation » peut être traduit en anglais par « assessment » ou « evaluation ». Ce dernier est parfois traduit par « mesure » en français selon le contexte. Voir la définition de ce mot.

*L'évaluation en tant qu'apprentissage* désigne les stratégies visant à engager les élèves dans l'établissement d'objectifs et dans l'autorégulation de leurs progrès.

*L'évaluation au service de l'apprentissage et l'évaluation en tant qu'apprentissage* entrent dans la catégorie de l'**évaluation formative**. Elles ont pour but d'appuyer et d'améliorer l'apprentissage plutôt que d'attribuer des notes aux élèves.

*L'évaluation de l'apprentissage*, aussi appelée **évaluation sommative**, s'entend d'une évaluation qui est généralement administrée à la fin d'une unité ou d'un cycle prescrit et qui tente de déterminer dans quelle mesure et à quel niveau les résultats d'apprentissage visés ont été atteints. L'évaluation sommative est souvent utilisée pour attribuer des notes aux élèves, pour faire rapport aux parents et pour mesurer l'efficacité de la programmation.

Lorsque le but de l'évaluation est de décrire le rendement de l'élève, on parlera des types d'évaluation suivants :

- **l'évaluation normative** : Le rendement d'un élève est décrit en fonction de son rang dans la population scolaire (la classe, le groupe d'âge, l'année, etc.).
- **l'évaluation critériée** : Le niveau de rendement d'un élève est décrit par rapport à des **normes** de performance établies.
- **l'évaluation par rapport à une performance antérieure de l'élève** : Le niveau de rendement d'un élève est décrit par rapport à ses progrès au fil du temps.

**Évaluation critériée** : voir **Évaluation**

**Évaluation normative** : voir **Évaluation**

**Évaluation formative** : voir **Évaluation**

**Évaluation par rapport à une performance antérieure de l'élève** : voir **Évaluation**

**Évaluation sommative** : voir **Évaluation**

**Fidélité** : Indique si l'on a recueilli *suffisamment* de données au sujet des connaissances ou des habiletés d'un élève pour que cette information soit crédible, et si les observations, les notes ou les résultats obtenus seraient les mêmes si l'information avait été recueillie selon une méthode différente ou interprétée par une autre personne (enseignant); en d'autres termes, c'est la *cohérence* de la mesure.

Les facteurs cruciaux de la fidélité comprennent notamment

- la conception des instruments d'évaluation
- le soin apporté à la correction
- la clarté de la grille d'évaluation ou des critères de correction
- la quantité de données d'évaluation recueillies
- la diversité des types d'occasions pour les élèves de démontrer leurs connaissances et leur savoir-faire

Lorsque cette notion est élargie de façon à inclure l'auto-évaluation des élèves, ceux-ci doivent connaître suffisamment les critères (ce qui est évalué ou fait l'objet d'un suivi et les descripteurs du niveau de performance, s'il y a lieu) et avoir assez d'occasions de pratiquer et assez de rétroaction sur leur auto-évaluation pour que les résultats soient « fidèles ».

**Généralisabilité** : Désigne le degré auquel il est approprié d'interpréter les données d'évaluation par-delà le contexte dans lequel ces données ont été recueillies. Pour généraliser des données d'évaluation, il faut respecter certaines modalités établies durant la collecte et l'interprétation de ces données. Par exemple, les résultats d'un test administré à un échantillon d'élèves ne peuvent être généralisés pour tous les élèves d'une division scolaire, à moins que cet échantillon ne soit représentatif de la population étudiante de la division et que les élèves aient été choisis de façon aléatoire. De même, on ne peut généraliser les résultats du rendement en lecture d'un élève à moins que les évaluations effectuées n'aient mesuré de façon approfondie tous les aspects relatifs à la lecture. Enfin, la possibilité de généralisation est directement liée au moment de la collecte des données et à la comparabilité des données obtenues à différents endroits (p. ex., en salle de classe). Si l'on prévoit faire une analyse comparative, les modalités d'évaluation (les éléments évalués et la façon de les évaluer) et les critères d'inclusion des élèves (c.-à-d. qui sera soumis à l'évaluation) doivent être communs ou appliqués généralement à toutes les sources de données.

**Grille d'évaluation** : Ensemble de critères à partir desquels les progrès, le rendement ou le travail d'un élève seront évalués. Il s'agit souvent d'une description détaillée de la qualité du travail à divers niveaux de rendement. Cette grille peut être élaborée en collaboration avec les élèves et est très utile quand elle est accompagnée de **copies types**.

**Mesure** : Ce mot comporte plusieurs définitions même dans le contexte de l'évaluation. Dans le présent document, on lui donne le sens suivant : Opération qui consiste à décrire quantitativement les observations d'élèves. Lorsqu'on se prononce sur la compétence d'un élève à partir de ces observations, on parle alors « d'évaluation ».

**Norme** : Dans un contexte d'évaluation, c'est un niveau fixe de rendement décrit, par exemple, dans les programmes d'études (résultats d'apprentissage et normes).

**Test basé sur les normes** : Au Manitoba, l'expression « test basé sur les normes » sert à indiquer que nos tests provinciaux de la 12<sup>e</sup> année sont basés sur des critères spécifiques aux programmes d'études (les résultats d'apprentissage) ainsi que des **normes**. Il s'agit de tests **sommatifs** et **critériés**.

**Test ou évaluation standardisé** : Instrument conçu, administré et corrigé conformément aux modalités établies afin de pouvoir effectuer, dans toute la mesure du possible, des comparaisons entre les élèves (évaluation comparative) et sur une échelle temporelle (analyse des tendances). Les tests standardisés sont souvent des tests **normatifs**.

**Validité** : Aborde la question de *ce que* l'évaluation mesure et de *comment* l'information est utilisée. Il peut arriver que les données d'évaluation soient très fidèles (précises et approfondies), tout en n'étant pas valides lorsque

- les instruments ou les stratégies utilisés pour recueillir les données (y compris les lignes directrices ou les critères de correction) ne témoignent pas de façon précise des connaissances et des habiletés que l'on veut évaluer (*validité du contenu*) comme le prévoient les programmes d'études
- la stratégie utilisée ne fait pas appel uniquement aux compétences ou processus intellectuels que l'on veut évaluer (*validité de construction*), par exemple, évaluation de la compréhension en lecture au moyen d'un travail de rédaction très élaboré
- l'information est utilisée d'une façon qui ne convient pas au but ou à la nature de l'évaluation entraînant des conséquences imprévues (*validité conséquentielle*)

## Annexe 2 : Exactitude et cohérence de la correction

Les tests provinciaux de la 12<sup>e</sup> année sont corrigés à l'échelle locale par les enseignants qui utilisent les lignes directrices concernant la correction fournies par le Ministère. Des échantillons de cahiers de test sont corrigés à nouveau à l'échelle centrale par le Ministère. On effectue ensuite des comparaisons entre la correction à l'échelle locale et celle à l'échelle centrale pour produire des rapports sur l'*exactitude* et la *cohérence* de la correction de chaque instance à l'échelle locale. **Il est important de tenir compte des résultats des rapports sur le processus de correction lorsque vous étudiez les tendances de performance au test provincial de la 12<sup>e</sup> année dans votre instance scolaire.**

Les rapports présentent deux types d'information sur la correction à l'échelle locale : l'*exactitude* de la correction et la *cohérence* de la correction. On détermine l'*exactitude* pour savoir si la moyenne des résultats des cahiers de test échantillonnés de votre instance scolaire est la même que celle des cahiers corrigés à l'échelle centrale, ou si on y trouve un biais dans un sens ou dans l'autre. L'*exactitude* est évaluée grâce à un test statistique commun, appelé un test t, et elle est présentée dans un tableau du rapport, à la colonne intitulée « Exactitude de la correction ». Le test t compare la différence des résultats entre la correction à l'échelle locale et la correction à l'échelle centrale des cahiers de test échantillonnés. Si on note que les résultats des cahiers corrigés à l'échelle locale ont une tendance constante à être plus élevés ou plus bas que les résultats des cahiers corrigés à l'échelle centrale, le résultat du test t sera alors *statistiquement significatif*, ce qui signifie que ce qui a été observé dans l'échantillon représente vraisemblablement la population où l'échantillon a été pris.

Voici des exemples de ce tableau et de son interprétation :

Exemple 1 : Exactitude de la correction : moyenne des résultats du test

Nombre d'élèves qui ont participé au test	Taille de l'échantillon	Notes moyennes des tests de l'échantillon		Exactitude de la correction
		Correction à l'échelle locale	Correction à l'échelle centrale	
55	7	63,3 %	61,2 %	=

L'analyse est fondée sur un échantillon de sept cahiers de test. Le signe égal (=) indique qu'il n'y a aucune tendance statistiquement significative montrant que la correction à l'échelle locale était constamment généreuse ou sévère. La correction à l'échelle locale était soit constamment exacte, soit variable, certains cahiers de tests étant corrigés avec exactitude, certains avec générosité et certains avec sévérité. On traite de la *cohérence* de la correction en détail un peu plus loin.

Exemple 2 : Exactitude de la correction : moyenne des résultats du test

Nombre d'élèves qui ont participé au test	Taille de l'échantillon	Notes moyennes des tests de l'échantillon		Exactitude de la correction
		Correction à l'échelle locale	Correction à l'échelle centrale	
55	7	62,7 %	58,8 %	+

Le signe plus (+) indique qu'il y a une tendance statistiquement significative de correction généreuse à l'échelle locale. Même si la correction à l'échelle locale était *inexacte*, elle pouvait tout de même être *cohérente* (dans ce cas elle était constamment généreuse), auquel cas le problème comprend probablement une tendance dans la correction qui peut être déterminée et éliminée, comme une interprétation particulière d'une clé de correction ou d'un barème de notation.

Exemple 3 : Exactitude de la correction : moyenne des résultats du test

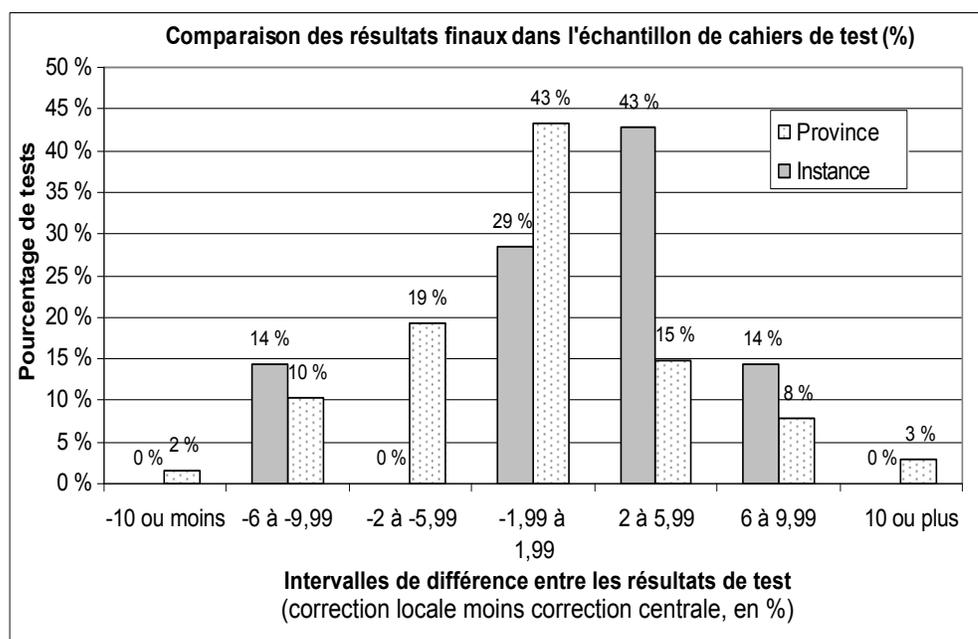
Nombre d'élèves qui ont participé au test	Taille de l'échantillon	Notes moyennes des tests de l'échantillon		Exactitude de la correction
		Correction à l'échelle locale	Correction à l'échelle centrale	
357	22	64,4 %	66,2 %	-

Le signe moins (-) indique qu'il y a une tendance statistiquement significative de correction sévère à l'échelle locale, bien qu'en moyenne celle-ci soit faible ( $64,4 - 66,2 = 1,8$  %). L'échantillon plus grand\* (22) fait de sorte que le test statistique (test t) soit sensible même aux petits degrés de biais. Le fait que la différence (1,8 %) est très faible **et** statistiquement significative signifie que la correction à l'échelle locale était très cohérente.

\* On utilise un échantillon plus grand lorsqu'il faut s'assurer que l'échantillon est représentatif d'un grand nombre d'écoles ou de classes.

La *cohérence* de la correction réfère au fait que les correcteurs sont cohérents dans leur façon de corriger, qu'ils soient exacts, généreux ou sévères. La *cohérence* de la correction est présentée sous forme de graphique dans le rapport sur le processus de correction, au moyen d'un diagramme à barres, tel qu'illustré dans les exemples qui suivent.

Exemple 1: Cohérence de la correction : résultats du test



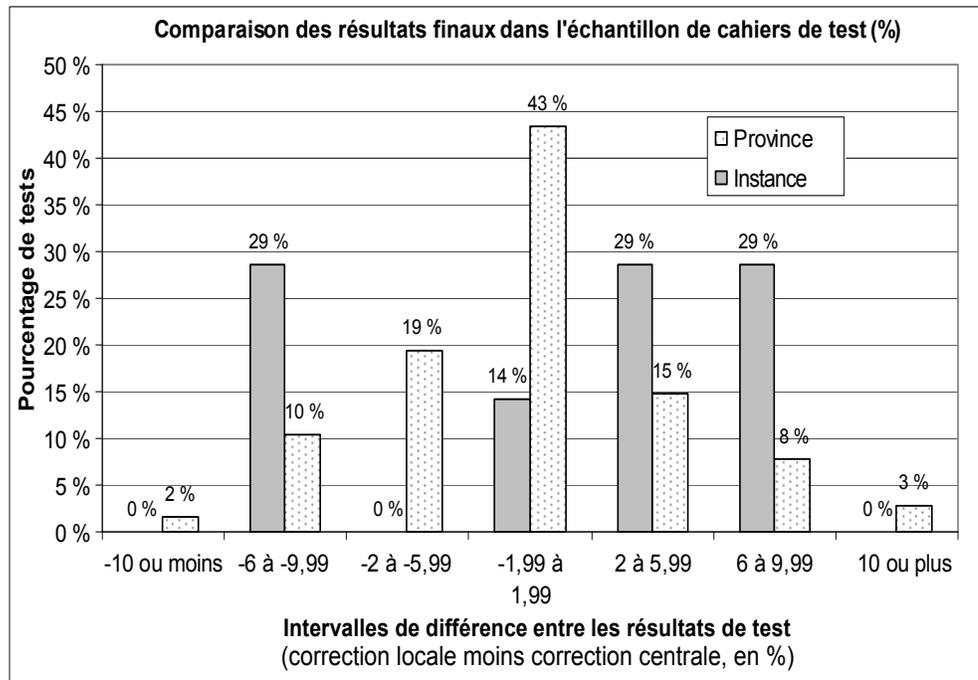
Instance :  $n = 0$      $n = 1$      $n = 0$      $n = 2$      $n = 3$      $n = 1$      $n = 0$

Dans l'ensemble de la province, 43 % des cahiers de test échantillonnés ont reçu les mêmes résultats à l'échelle locale qu'à l'échelle centrale, +/- 1,99 point de pourcentage, alors que pour 15 % des cahiers de test, les résultats à l'échelle locale ont dépassé les résultats à l'échelle centrale de 2 à 5,99 points de pourcentage. Il n'y avait pas de biais constant vers des résultats plus élevés ou plus bas à l'échelle locale, comme l'indique le motif de cloche formé par les bandes représentant la province (grandes barres au centre et barres de plus en plus petites de chaque côté), la barre la plus longue (43 %) représentant la différence la plus faible entre les résultats de test (entre -1,99 % et 1,99 %).

Les barres grises (instances) indiquent que les correcteurs à l'échelle locale ont eu tendance à être généreux. Cette situation est illustrée par le fait que les bandes grises les plus longues sont situées au centre ou à sa droite, 57 % (43 % + 14 %) des cahiers de test échantillonnés ayant reçu une note plus élevée à l'échelle locale qu'à l'échelle centrale. En même temps, le motif de cloche formé par les barres grises indique un degré de *cohérence* dans la correction. Les valeurs «  $n$  » sous le tableau indiquent le nombre de cahiers de test échantillonnés tombant dans chaque catégorie de différence entre les résultats de test pour l'instance.

Un exemple de variabilité de la correction dans l'instance (barres grises) est illustré ci-dessous.

Exemple 2: Cohérence de la correction : résultats du test



Instance : n = 0      n = 2      n = 0      n = 1      n = 2      n = 2      n = 0

Pour l'instance, la correction des cahiers de test échantillonnés (barres grises) à l'échelle locale était variable. Cette situation est illustrée par l'absence de motif de cloche. Certains cahiers de test ont été corrigés de façon sévère (29 % dans la catégorie -6 % à -9,99 %) et certains ont été corrigés de façon généreuse. Les causes probables comprennent le fait qu'il puisse y avoir plusieurs interprétations des critères de correction, une inattention aux critères, ou le fait que les correcteurs aient été influencés par des facteurs non associés à la qualité des travaux des élèves.

En général, la taille des échantillons utilisés dans ces analyses est petite, il faut donc être prudent au moment de l'interprétation. Comme toutes les analyses faites à partir d'échantillons, il est possible que l'échantillon ne reflète pas adéquatement l'ensemble de la correction dans votre instance. Le mieux est d'examiner les résultats de ce rapport en tenant compte des facteurs qui se rapportent à la correction dans votre instance et des résultats des rapports antérieurs sur le processus de correction.

## **Annexe 3 : Gestion des observations aberrantes en statistique**

---

Une observation aberrante est une donnée (une note de test, par exemple) dont la valeur se situe à l'extérieur de la distribution ou du profil général des autres données. La plupart du temps, il est facile d'expliquer une observation aberrante. S'il s'agit d'une erreur dans la collecte des données, il suffit de la corriger. La recherche d'observations aberrantes est une pratique commune lorsqu'on veut vérifier l'exactitude des données avant d'amorcer toute analyse.

Si l'observation aberrante est une donnée exacte, il n'est peut-être pas approprié de l'éliminer. Mais sa présence peut altérer les résultats des analyses, entraînant des erreurs d'interprétation et des décisions mal fondées. Ce pourrait être le cas, par exemple, lorsqu'un élève obtient un résultat anormalement faible à un test : ce résultat peut abaisser indûment la note finale de l'élève, donnant une fausse indication de ses connaissances et habiletés générales s'il y a trop peu de notes pour arriver à la note finale.

Une méthode de gestion des observations aberrantes dans le calcul de la moyenne (pour un groupe ou un élève en particulier) consiste à utiliser la note **médiane** plutôt que la **moyenne**. La médiane, tout comme la moyenne, mesure la tendance centrale d'une série de notes, mais elle est déterminée différemment de la moyenne. La médiane est la note située au centre d'une série de notes classées de la plus faible à la plus forte.

On trouvera ci-dessous la moyenne et la médiane d'une liste de sept notes de tests qui sont toutes exprimées en pourcentage et classées de la plus faible à la plus forte, pour des raisons pratiques.

20 %, 65 %, 70 %, 75 %, 80 %, 80 %, 85 %

moyenne = 67,9 %

médiane = 75 % (la note centrale de la liste)

(Pour un nombre de notes pair, la médiane est la moyenne des deux notes centrales.)

À supposer que ces notes aient été obtenues à la suite de tests sommatifs pour un élève, la médiane pourrait constituer une représentation plus exacte de la performance générale de l'élève. La médiane est un outil qui permet aux utilisateurs de données de résoudre les observations aberrantes.

Une description claire et détaillée des principes et des modalités permettant de résumer les notes des élèves et de faire rapport à ce sujet se trouve dans le document *The Mindful School: How to Grade for Learning*, de Ken O'Connor.

## Références

---

- LEGENDRE, Renald. *Dictionnaire actuel de l'éducation*, 3<sup>e</sup> éd., Montréal, Guérin, 2005.
- MANITOBA. MINISTÈRE DE L'ÉDUCATION, DE LA CITOYENNETÉ ET DE LA JEUNESSE. *Grade 3 Assessment in Reading, Lecture and Numeracy and Grade 4 Assessment in French Immersion Lecture = Évaluation de Reading, de la lecture et des notions de calcul des élèves de 3<sup>e</sup> année et évaluation de la lecture des élèves de 4<sup>e</sup> année d'immersion française*, Winnipeg, Manitoba, Le Ministère, 2004. Accessible en ligne : [www.edu.gov.mb.ca/k12/assess/publications.html](http://www.edu.gov.mb.ca/k12/assess/publications.html).
- MANITOBA. MINISTÈRE DE L'ÉDUCATION, DE LA CITOYENNETÉ ET DE LA JEUNESSE. *Politiques et modalités pour les tests basés sur les normes*, Winnipeg, Manitoba, Le Ministère, Publication annuelle. Accessible en ligne : [www.edu.gov.mb.ca/k12/assess/publications.html](http://www.edu.gov.mb.ca/k12/assess/publications.html).
- MANITOBA. MINISTÈRE DE L'ÉDUCATION, DE LA CITOYENNETÉ ET DE LA JEUNESSE. *Repenser l'évaluation en classe en fonction des buts visés : l'évaluation au service de l'apprentissage, l'évaluation en tant qu'apprentissage, l'évaluation de l'apprentissage*, Winnipeg, Manitoba, Le Ministère, 2006. Accessible en ligne : [www.edu.gov.mb.ca/frpub/me/docs/repenser\\_eval/index.html](http://www.edu.gov.mb.ca/frpub/me/docs/repenser_eval/index.html).
- MANITOBA. MINISTÈRE DE L'ÉDUCATION ET DE LA FORMATION PROFESSIONNELLE. *Le succès à la portée de tous les apprenants : Manuel concernant l'enseignement différentiel*, Winnipeg, Manitoba, Le Ministère, 1996.
- MANITOBA. MINISTÈRE DE L'ÉDUCATION ET DE LA FORMATION PROFESSIONNELLE. *Senior 4 English Language Arts: A Foundation for Implementation*, Winnipeg, Manitoba, Le Ministère, 2000.
- MANITOBA. MINISTÈRE DE L'ÉDUCATION ET DE LA JEUNESSE. *Vocabulaire de la planification de l'éducation — ÉBAUCHE*, Winnipeg, Manitoba, Le Ministère, mars 2003. [En ligne], [www.edu.gov.mb.ca/frpub/fc/praction/vocabulaire.pdf](http://www.edu.gov.mb.ca/frpub/fc/praction/vocabulaire.pdf).
- O'CONNOR, Ken. *The Mindful School: How to Grade for Learning*. Arlington Heights, IL: Skylight Training and Publishing Inc., 1999.