

Résumé du rapport technique de l'OQRE

pour les tests de 2006-2007

*Tests en lecture, écriture et mathématiques,
cycle primaire (de la 1^{re} à la 3^e année) et cycle moyen (de la 4^e à la 6^e année),
Test de mathématiques, 9^e année, et
Test provincial de compétences linguistiques*



En vue de remplir son mandat, l'OQRE effectue quatre tests annuels à l'échelle de la province auprès des élèves de langue française et de langue anglaise.

Les évaluations à grande échelle visent à mesurer le rendement des élèves à un certain moment, dans des conditions standardisées.

Les devis des tests de l'OQRE servent à élaborer les items à choix multiple et à réponse construite pour chaque test, afin que le test présente les mêmes caractéristiques chaque année.

Introduction

Les évaluations à grande échelle de l'OQRE évaluent dans quelle mesure les élèves satisfont aux attentes et contenus d'apprentissage décrits dans *Le curriculum de l'Ontario*. Les tests contiennent des items à réponse construite qui demandent une réponse écrite ainsi que des items à choix multiple. Chaque année, on prépare un document-cadre pour chaque test afin de décrire de manière détaillée de quelle façon sera élaboré le test.

Après leur administration, les tests produisent des résultats au niveau des élèves, des écoles et des conseils scolaires. Les résultats des tests procurent des renseignements utiles qui appuient la planification de l'amélioration par les écoles, les conseils scolaires et le ministère de l'Éducation de l'Ontario.

Ce résumé a pour objet de décrire certaines des caractéristiques techniques et l'expertise professionnelle qui ont servi à assurer l'exactitude, la validité et l'intégrité psychométrique des tests suivants, administrés durant l'année scolaire 2006-2007 :

- Tests en lecture, écriture et mathématiques, cycle primaire (de la 1^{re} à la 3^e année) et cycle moyen (de la 4^e à la 6^e année);
- Test de mathématiques, 9^e année;
- Test provincial de compétences linguistiques (TPCL).

Méthodes utilisées pour produire les tests de l'OQRE

Élaboration des items

Il est essentiel de veiller à ce qu'on tienne compte des mêmes attentes et contenus d'apprentissage du curriculum de l'Ontario et que le degré de difficulté des tests soit similaire chaque année. Pour ce faire, les nouveaux items sont fondés sur une version détaillée des devis. L'OQRE recrute et forme des éducatrices et éducateurs spécialisés dans les domaines de la littérature (lecture et écriture en français et en anglais) et des mathématiques, venant de tous les coins de la province, afin d'appliquer les devis à leur travail dans le cadre des comités d'élaboration des items.

Chaque année, de nouveaux items sont élaborés et mis à l'essai pour devenir des items de la composante test dans le test de l'année suivante.

On accorde une grande attention à la sélection d'items appropriés à l'âge et à l'année d'études des élèves ainsi qu'aux attentes et contenus d'apprentissage décrits dans le curriculum.

Les items mis à l'essai sont intégrés aux tests provinciaux de chaque année.

Avant de pouvoir être inclus dans un test, un item doit respecter des critères stricts, être mis à l'essai, passer une révision de sensibilité et être révisé et édité. Pour les items à choix multiple, les équipes examinent la clarté et l'intégralité de l'amorce, l'exactitude de la réponse saisie et la plausibilité des trois leurres. Pour les items de lecture et de mathématiques à réponse construite ainsi que les tâches d'écriture, les équipes tiennent compte de la correspondance des items à leurs grilles de notation pour déterminer si les items vont obtenir la gamme de réponses attendues.

Les membres du comité d'évaluation du contenu et du comité de sensibilité mettent l'accent sur les questions d'équité et d'accessibilité pour le plus grand nombre d'élèves en Ontario. Pour ce faire, ils tiennent compte des questions d'équité en éducation (notamment les questions de multiculturalisme, les questions touchant les élèves inscrits à un programme d'ALF/ de PDF et les élèves ayant des besoins particuliers) afin qu'aucun groupe particulier d'élèves ne soit avantagé ou désavantagé par un des items figurant dans le test.

Mise à l'essai

Étant donné que les items de la composante mise à l'essai présentent le même format que les items de la composante test, les élèves ne savent pas à quel type d'item ils répondent. Cette technique évite le manque de motivation qui se produit chez les élèves lorsqu'ils répondent à de nouveaux items mis à l'essai en dehors du test provincial. Les résultats obtenus pour les items mis à l'essai ne sont pas utilisés pour établir les résultats au niveau des élèves, des écoles, des conseils scolaires ou de la province.

La mise à l'essai du matériel du test permet d'assurer que les items sélectionnés pour de futurs tests provinciaux sont appropriés et justes sur le plan psychométrique pour tous les élèves. La mise à l'essai permet aussi de fournir des données qui sont utilisées pour mettre en équivalence le test de chaque année avec celui de l'année précédente de façon à pouvoir effectuer une comparaison valide des résultats. La qualité de la comparaison des données d'une année à l'autre dépend de cette mise en équivalence.

Un certain nombre de directives doivent être suivies.

Élaboration des tests

L'OQRE tient compte des directives qui suivent pour sélectionner les items devant faire partie de ses tests provinciaux :

- Les items doivent être équitables et le degré de difficulté du test doit se comparer à celui des tests précédents.
- Les items doivent être conformes au devis et être équilibrés quant aux points suivants : le contenu de la matière, la représentation des filles et des garçons ainsi que les aspects de la démographie provinciale (urbain/rural, nord/sud).
- Avec le temps, l'ensemble des attentes et contenus d'apprentissage mesurables d'un groupe d'attentes doivent être inclus dans un test.

Élaboration des questionnaires

Les informations contextuelles sont utilisées pour analyser les données par groupe et mettre l'accent sur les recommandations portant sur l'amélioration de l'apprentissage des élèves.

L'OQRE élabore des questionnaires à l'intention des élèves, du personnel enseignant et des directions d'école pour recueillir des données et des informations contextuelles sur les pratiques d'enseignement et les attitudes liées à la littératie et aux mathématiques. Les questions portent sur un vaste éventail de facteurs associés au rendement des élèves.

Administration des tests

Les guides d'administration décrivent les procédures qui assurent l'uniformité et l'équité des tests pour tous les élèves.

Pour chaque test, il existe un guide d'administration qui décrit les rôles et responsabilités des enseignantes et enseignants et donne des directives sur la préparation des élèves au test et ce qu'il faut leur dire ainsi que sur leur participation (c.-à-d. adaptations pour les élèves ayant des besoins particuliers, dispositions particulières pour les élèves inscrits à un programme d'ALF/de PDF, exemptions et reports).

Une autre façon pour l'OQRE d'assurer l'intégrité des procédures d'évaluation est d'affecter des observatrices et observateurs dans des écoles choisies de façon aléatoire durant les tests. En outre, l'organisme analyse statistiquement les fichiers de données des réponses d'élèves pour repérer des patrons de réponses inhabituels.

Des procédures de notation rigoureuses assurent la fidélité des résultats des tests.

Notation

Les items à choix multiple sont notés à l'aide d'un système de reconnaissance optique de marques. Les réponses écrites aux items à réponse construite en lecture et en mathématiques et aux tâches d'écriture sont notées soigneusement et systématiquement par des éducatrices et éducateurs de l'Ontario. Une échelle descriptive d'appréciation générique pour chaque type d'item à réponse construite décrit le travail correspondant aux différents niveaux de rendement et maintient l'uniformité entre les items et d'une année à l'autre. Une grille de notation spécifique aux items pour chaque item à réponse construite, pour chaque matière, se fonde sur les grilles de notation génériques. Les copies types illustrent les descripteurs pour les points bruts de la grille de notation.

Sélection des copies types

L'objectif de la sélection des copies types, qui est la première étape du processus de sélection, est de définir et d'illustrer les divers niveaux de rendement pour chaque code des grilles de notation. Cette tâche est accomplie par un comité composé d'éducatrices et d'éducateurs ayant une expertise dans le système éducatif de l'Ontario et qui s'entendent sur les réponses d'élèves représentant chaque point brut d'une grille de notation. Cet encodage des réponses d'élèves est ensuite utilisé pour sélectionner les copies types pour la grille de notation et pour former les correctrices et correcteurs qui noteront les réponses des élèves.

Une formation poussée et des procédures de contrôle produisent des résultats présentant le plus haut niveau de validité.

Chaque correctrice ou correcteur suit une formation poussée pour acquérir une compréhension claire et commune du matériel de notation, afin que chaque chef de notation, adjointe ou adjoint au chef de notation et correctrice ou correcteur interprète et applique le matériel de notation de la même façon.

Voici d'autres mesures de protection et procédures de contrôle de la qualité :

- tests de qualification pour les correctrices et correcteurs;
- supervision des correctrices et correcteurs par les chefs de notation;
- examen quotidien et activités de calibrage;
- notation anonyme;
- procédures pour signaler une preuve d'élèves à risque, de contenu inapproprié et d'ingérence par l'enseignante ou l'enseignant;
- comparaisons quotidiennes avec des copies de validation afin d'éviter toute dérive dans la notation;
- surveillance du taux d'exactitude des correctrices et correcteurs;
- normes de productivité;
- vérification des incohérences dans la notation;
- normes de rendement pour les chefs de notation et leurs adjointes et adjoints.

Toutes ces mesures visent à assurer une notation uniforme du travail des élèves peu importe la correctrice ou le correcteur qui effectue la notation.

Mesure de la fidélité et de la validité des correctrices et correcteurs

Les mesures de fidélité interjuges et de validité de notation servent à mesurer l'uniformité du rendement des correctrices et correcteurs. On compare les résultats attribués à la réponse d'un élève donné pour voir s'ils sont exactement les mêmes, séparés d'un point brut (adjacente) ou de plus d'un point brut. Au moins deux correctrices et correcteurs notent chaque réponse d'élève pour le TPCL. Si les deux scores ne sont pas identiques ou adjacents, une correctrice spécialiste ou un correcteur spécialiste va trancher en ce qui a trait au score.

La fidélité interjuges est déterminée par la comparaison des scores attribués par des pairs de correctrices et correcteurs à des réponses d'élèves choisis de façon aléatoire.

Pour la fidélité et la validité, l'OQRE a établi les normes industrielles ci-dessous :

- **accord précis de 60 % pour les grilles de notation à six points;**
- **accord précis de 70 % pour les grilles de notation à quatre points;**
- **accord précis de 75 % pour les grilles de notation à trois points;**
- **accord précis et adjacent de 95 %.**

Fidélité interjuges

Pour les tests des cycles primaire et moyen, l'objectif de l'OQRE en matière de concordance parfaite ou adjacente a été atteint pour 70 % des items de lecture en français et pour 85 % des items de lecture en anglais, pour aucune tâche d'écriture en français et pour 92 % des tâches d'écriture en anglais ainsi que pour 75 % des items de mathématiques en français et pour 100 % des items de mathématiques en anglais. L'objectif d'accord précis et adjacent de l'OQRE a été atteint pour tous les items de lecture, pour 58 % des tâches d'écriture en français et pour 75 % des tâches d'écriture en anglais, ainsi que pour 63 % des items de mathématiques en français et pour tous les items de mathématiques en anglais.

Pour le Test de mathématiques, 9^e année, les pourcentages d'accord précis et d'accord précis et adjacent ont dépassé les objectifs de l'OQRE pour tous les items dans les deux langues officielles, à l'exception d'un accord précis pour un item en français dans le test de mathématiques du cours appliqué.

Pour les items de lecture du TPCL, les pourcentages d'accord précis n'ont pas atteint l'objectif de l'OQRE. Les pourcentages d'accord précis et adjacent ont atteint l'objectif pour 75 % des items en français et pour 25 % des items en anglais. Pour les tâches d'écriture dans le cadre du TPCL, les pourcentages d'accord précis ont atteint l'objectif uniquement pour les tâches d'écriture à réponse courte en anglais. Les pourcentages d'accord précis et adjacent ont atteint l'objectif pour 50 % des tâches d'écriture en français et pour 67 % des tâches d'écriture en anglais.

La validité de la notation est mesurée par la comparaison du degré d'accord entre les scores attribués par les correctrices et correcteurs et ceux attribués par un panel d'experts.

Validité

Pour les tests des cycles primaire et moyen, l'objectif de l'OQRE en matière d'accord précis a été atteint pour 55 % des items de lecture en français et pour 20 % des items de lecture en anglais, toutes les tâches d'écriture en français et aucune des tâches d'écriture en anglais ainsi que pour tous les items de mathématiques, sauf pour un item en français. L'objectif de l'OQRE pour un accord précis et adjacent a été atteint pour tous les items de lecture en français et pour 70 % des items de lecture en anglais, pour 83 % des tâches d'écriture en français et pour la moitié des tâches d'écriture en anglais ainsi que pour tous les items de mathématiques.

Pour le Test de mathématiques, 9^e année, les pourcentages d'accord précis et d'accord précis et adjacent ont dépassé les objectifs de l'OQRE pour tous les items du test, dans les deux langues.

Pour le TPCL, les pourcentages d'accord précis ont atteint l'objectif pour la moitié des items de lecture en français et aucun des items de lecture en anglais ainsi qu'uniquement pour les tâches d'écriture courte. Les pourcentages d'accord précis et adjacent ont atteint l'objectif pour la moitié des items de lecture et pour 67 % des tâches d'écriture.

Dans les cas où les scores attribués par les correctrices et correcteurs ont différencié de ceux attribués par le panel, la tendance observée de notes plus élevées et moins élevées que celles attribuées par le panel était semblable, à quelques variantes près dans les items.

Notation de la mise à l'essai

La notation d'items à réponse construite de la composante mise à l'essai est réalisée après la notation des items de la composante test. Les correctrices et correcteurs les plus fiables et productifs en notation de la composante test effectuent la notation des items de la composante mise à l'essai afin d'assurer l'uniformité d'une année à l'autre et de réduire le temps requis pour former les correctrices et correcteurs. Le personnel de notation de la mise à l'essai est formé à l'aide d'items mis à l'essai et de tâches, de grilles de notation spécifiques aux items et d'exigences en matière de notation, afin de produire des données spécifiques aux items et aux tâches qui sont valides et fidèles pour la conception des tests provinciaux et la mise en équivalence.

Mise en équivalence et analyse des résultats des élèves

La mise en équivalence est un processus qui permet de convertir les mesures d'une échelle à une autre, un peu comme l'on convertit les mesures de température des degrés Fahrenheit aux degrés Celsius.

L'OQRE emploie divers modèles, conceptions et analyses statistiques, afin d'assurer la comparabilité des résultats des tests d'une administration à l'autre. Le processus de mise en équivalence met les résultats d'élèves de deux années adjacentes sur une échelle commune. Cela permet de contrôler les petites différences dans le degré de difficulté des tests d'une année à l'autre et de veiller à ce que les élèves d'une année donnée ne soient pas avantagés au détriment des élèves d'une autre année.

La théorie de la réponse à l'item définit la relation entre une ou un élève ayant un niveau de capacité donné et la probabilité que l'élève réponde correctement à un item donné.

Les seuils de rendement sont ensuite déterminés pour indiquer les niveaux de rendement utilisés dans la communication des résultats de l'OQRE (quatre niveaux de rendement pour le cycle primaire, le cycle moyen et la 9^e année ainsi que réussite ou non réussite du TPCL). Les élèves qui reçoivent un résultat donné dans l'année en cours auront démontré le même niveau d'habileté et de connaissances que les élèves des années antérieures à qui on a attribué ce niveau.

On retrouve beaucoup de similarités, mais aussi quelques différences dans les procédures de mise en équivalence pour les tests des cycles primaire et moyen, le test de 9^e année et le TPCL.

L'OQRE utilise la théorie de la réponse à l'item (TRI) pour le calibrage des items, la notation et la mise en équivalence. Les modèles généraux de la TRI, utilisés par l'OQRE, qui servent à estimer les paramètres d'item et de capacité pour les items à choix multiple et à réponse construite, sont le modèle logistique à trois paramètres et le modèle à crédit partiel généralisé, respectivement. Le modèle logistique à trois paramètres est modifié en fixant le paramètre de pseudo-chance à 0,20 pour tous les items à choix multiple. Pour le TPCL, l'OQRE utilise un modèle de Rasch modifié dont le paramètre de pseudo-chance est fixé à 0,20.

Le plan du groupe non équivalent d'items communs à paramètre fixe-avant a été utilisé pour mettre en équivalence les tests de l'OQRE administrés en 2006-2007 et en 2005-2006. Les tests provinciaux de l'année scolaire 2006-2007 ont été créés à partir d'items mis à l'essai en 2005-2006 et ces items ont été utilisés pour établir un lien entre les tests. Les items des deux années scolaires ont été calibrés à l'échelle pour chaque test provincial de l'année scolaire 2006-2007, puis les paramètres de ces items ont été utilisés pour noter de nouveau les réponses d'élèves de l'année scolaire 2005-2006. Enfin, on a déterminé les valeurs thêta correspondant au pourcentage d'élèves pour chaque niveau de rendement de l'année scolaire 2005-2006 et elles ont été appliquées à l'étendue des résultats de l'année scolaire 2006-2007 pour déterminer le pourcentage des élèves pour chaque niveau de rendement de cette année.

Les procédures de mise en équivalence et la détermination des seuils de rendement ont été réalisées séparément pour les versions française et anglaise des tests.

Les analyses ont été réalisées à deux reprises : une fois par le personnel de l'OQRE et une autre fois par un sous-traitant externe. Les résultats ont ensuite été comparés, puis les différences ont été corrigées, garantissant ainsi l'exactitude des résultats.

Une fois la notation terminée, des analyses statistiques et psychométriques ont été réalisées à l'aide de la théorie classique des scores et de la théorie de la réponse à l'item.

Diverses statistiques de test ont été calculées, notamment le coefficient de fidélité alpha de Cronbach, les erreurs-types de mesure, les courbes caractéristiques de test et les fonctions informations de test, pour évaluer le degré de précision des scores attribués aux tests de l'OQRE de 2007. Dans l'ensemble, les résultats de ces mesures indiquent que des niveaux de précision satisfaisants ont été obtenus. Les coefficients de fidélité varient entre 0,77 et 0,89 pour les tests des cycles primaire et moyen, entre 0,81 et 0,86 pour le Test de mathématiques, 9^e année, et entre 0,85 et 0,87 pour le TPCL

Publication des résultats

Les résultats décrivent le rendement de l'élève à des étapes importantes de son éducation.

Les résultats des tests sont publiés au niveau des élèves, des écoles, des conseils scolaires et de la province. Les résultats fournissent de l'information utile permettant de soutenir la planification de l'amélioration à tous ces niveaux.

Pour les tests des cycles primaire et moyen et le test de 9^e année, une échelle à quatre niveaux est utilisée pour communiquer le rendement de l'élève en lecture, en écriture et en mathématiques.

Les niveaux de rendement utilisés pour publier les résultats des tests des cycles primaire et moyen et de la 9^e année proviennent du curriculum de l'Ontario, qui établit le niveau 3 comme étant la norme provinciale. Les niveaux 1 et 2 indiquent un rendement inférieur à la norme provinciale et le niveau 4 indique un niveau supérieur à cette norme.

Pour le TPCL, deux niveaux de rendement sont communiqués : réussite et non réussite.

Quant au TPCL, les élèves sont classés en deux catégories, soit réussite ou non réussite. Les élèves ayant réussi ont atteint le niveau de compétences linguistiques prévu à la fin de la 9^e année. Ils ont donc satisfait à l'exigence linguistique permettant d'obtenir leur diplôme. Les élèves n'ayant pas réussi reçoivent une rétroaction afin de les aider à travailler en vue d'atteindre la norme.

Tous les rapports sont conçus avec la plus grande attention en ce qui a trait à l'exactitude, la clarté et l'efficacité.

Les rapports comprennent :

- **des tableaux des résultats;**
- **des diagrammes des résultats;**
- **des renseignements directeurs pour l'interprétation des résultats.**

Les rapports à l'intention des conseils scolaires et des écoles sont affichés sur le site Web public de l'OQRE à l'adresse suivante : www.oqre.on.ca. Si le nombre d'élèves pour une école ou un conseil scolaire donné est inférieur à 15, les résultats ne sont pas affichés afin d'éviter la possibilité de pouvoir identifier les résultats individuels des élèves. Les rapports détaillés incluant les résultats de tous les élèves et les fichiers de données sont affichés dans la section sécurisée du site Web de l'OQRE, qui est accessible au personnel des conseils scolaires et des écoles à l'aide d'un code d'identification et d'un mot de passe.

Les parents et les tuteurs et tutrices reçoivent les Rapports individuels de l'élève, qui donnent les résultats globaux pour chaque test en lecture, en écriture et en mathématiques des cycles primaire et moyen ainsi que pour le test de mathématiques, cours appliqué ou théorique, de 9^e année. Les rapports publiés à l'intention des écoles, des conseils scolaires et de la province fournissent des informations contextuelles; les résultats pour chaque matière pour l'école, le conseil scolaire et la province; les résultats selon le sexe de l'élève, pour les élèves bénéficiant d'un soutien ou d'un programme d'ALF/de PDF et les élèves ayant des besoins particuliers; les résultats au fil du temps; les forces et les points devant être améliorés relativement au curriculum; les résultats sur l'information des items pour chaque élève, école, conseil scolaire et pour la province ainsi que les résultats issus des questionnaires à l'intention des élèves, du personnel enseignant et des directions d'école. Les résultats provenant des questionnaires à l'intention du personnel enseignant et des directions d'école ne sont pas publiés au niveau des écoles.

De même, pour le TPCL, les résultats individuels d'élèves sont communiqués aux élèves tout comme les résultats des élèves admissibles pour la première fois et admissibles antérieurement aux niveaux de l'école, du conseil scolaire et de la province sont communiqués aux écoles et aux conseils scolaires.

En outre, l'OQRE publie un rapport provincial sur chaque test, versions française et anglaise, pour les intervenantes et intervenants en milieu scolaire et le public. Ces rapports sont accessibles au public sur le site Web de l'OQRE.