

Evaluation externe 3^e année

**RESULTATS
ET
COMMENTAIRES**

Dossier pour les enseignants

Avril 1996

Ministère de l'Education, de la Recherche et de la Formation
Pilotage de l'Enseignement en Communauté française

L'épreuve standardisée d'évaluation externe a été organisée par les Inspections de l'enseignement fondamental à la demande de la **Cellule de pilotage** du Ministère de l'Éducation, de la Recherche et de la Formation.

Cette épreuve a été conçue par Dominique LAFONTAINE, Chercheuse à l'Université de Liège, Jacques GREGOIRE, Professeur à l'Université Catholique de Louvain, Geneviève DE COCK, Assistante à l'Université Catholique de Louvain, Valérie PIRON, Inspectrice générale et Gilberte CHRISTIAEN, Josette DERMOUCHAMPS, André DEROUBAIX, Alain FILLEUL, Freddy GEORGES, Jean-Marie HENDRICKS, Francis HERMAL, André JOACHIM, Alain LAURENT, Michel MOSSIAT, Victor NIZET, Inspecteurs de l'enseignement primaire.

Table des matières

PREMIERE PARTIE. INFORMATIONS GENERALES	4
1. Contexte	4
2. Rappel des objectifs généraux du test	4
3. Caractéristiques de l'échantillon	5
4. Comparez les résultats de vos élèves à ceux de l'échantillon	6
DEUXIEME PARTIE. ANALYSE DES RESULTATS	
PAR SITUATION	8
Introduction	8
1. Résultats à la situation 1	9
2. Résultats à la situation 2	14
3. Résultats à la situation 3	22
4. Résultats à la situation 4	30
5. Analyse des résultats à l'épreuve de lecture de mots isolés	35
TROISIEME PARTIE. ANALYSE DES RESULTATS	
POUR L'ENSEMBLE DE L'EPREUVE	38
1. Distribution des résultats à l'ensemble de l'épreuve	38
2. Comment situer votre classe par rapport à l'ensemble des classes	40
3. Comment situer votre classe par rapport à des classes qui lui « ressemblent » ?	42
4. Résultats en fonction des caractéristiques des élèves	44
5. Niveaux de compétence des élèves	45
6. Evaluation par les titulaires et les inspecteurs du « niveau » des compétences mises en œuvre dans l'épreuve	50
7. Les résultats au test correspondent-ils au "niveau" de la classe apprécié par le titulaire	52
QUATRIEME PARTIE. RESULTATS DU	54
QUESTIONNAIRE AU TITULAIRE	54
1. Caractéristiques de l'école et de la classe	54
2. Appréciation de la difficulté de l'épreuve par situation	56
BILAN ET PERSPECTIVES	57
GLOSSAIRE	58

PREMIERE PARTIE. INFORMATIONS GENERALES

1. Contexte

Au début du second semestre de cette année scolaire 1995-1996, une opération d'évaluation externe s'est déroulée en troisième année primaire à l'initiative de la Cellule de Pilotage du Ministère de l'Education, de la Recherche et de la Formation. Il s'agissait de la deuxième opération de ce genre, après l'évaluation effectuée en 5e primaire en 19941

Tous les élèves de troisième ont donc passé une même épreuve pluridisciplinaire au cours de la même semaine. L'épreuve, d'après les échos que nous avons recueillis dans les questionnaires aux titulaires, a été unanimement appréciée. Les élèves, les enseignants, les chefs d'établissements et les inspecteurs ont, chacun à leur niveau, activement contribué à faire de l'opération un succès et nous tenons, avant toute autre considération, à les en remercier chaleureusement. En dépit des difficultés d'organisation, les opérations se sont déroulées au mieux : aucune grille ou questionnaire ne manque à l'appel, les documents sont soigneusement complétés, des commentaires et coups de téléphone utiles ont été reçus par les Universités. Voici autant de motifs de satisfaction pour la cellule organisatrice de l'opération.

Lors de la mise en place de l'opération, nous nous sommes engagés à vous fournir des résultats le plus rapidement possible. Telle est la fonction du présent document. Vous y trouverez une description des résultats des élèves appartenant à un échantillon représentatif de classes tirées aléatoirement au sein de la Communauté française, tous réseaux confondus. La possibilité vous est offerte de comparer les résultats de vos élèves à ceux de cet échantillon. Pour ce faire, vous devez suivre la procédure décrite au point 4 de la Ire partie.

Les résultats de l'échantillon ont fait l'objet d'une analyse par question, par groupe de questions et pour l'ensemble de l'épreuve. Cette analyse vous apporte un premier éclairage sur les compétences que maîtrisent (ou ne maîtrisent pas) les élèves au sortir du premier degré. D'autres analyses, plus fouillées, vous parviendront au mois de septembre. Elles seront accompagnées de pistes didactiques.

2. Rappel des objectifs généraux du test

Comme expliqué lors de la présentation de l'épreuve, l'évaluation externe en troisième année primaire avait pour objectif principal de **faire le point, en début de 2e cycle, sur des compétences acquises à l'issue du 1er cycle**. Cette épreuve organisée à l'articulation des deux cycles permet en outre d'ouvrir des perspectives de travail pour le 1er et le 2e cycles.

Une épreuve ayant cette finalité ne vise donc pas, comme le font classiquement beaucoup d'épreuves, à discriminer les élèves, et à recréer la *courbe de Gauss* (beaucoup d'élèves moyens, une minorité d'élèves faibles et une minorité d'élèves forts). A ce type d'épreuves, des résultats en courbe de Gauss seraient inquiétants. C'est au contraire une *courbe en forme de J2* que l'on attend (la majorité des élèves réussissant très bien l'épreuve, et une minorité d'élèves la réussissant moins bien), puisqu'en principe tous les élèves devraient avoir acquis les compétences évaluées par l'épreuve (cette affirmation sera nuancée au fil du document).

1 Pour information, une épreuve d'évaluation externe en français et en mathématique aura lieu à l'automne 1996 en Ire année du secondaire.

2 Ces termes sont expliqués dans le glossaire en fin de document.

3. Caractéristiques de l'échantillon

Toutes les classes de troisième ont passé l'épreuve et leurs titulaires en ont corrigé individuellement les épreuves.

Les Universités ont de leur côté traité les résultats d'un échantillon représentatif d'élèves.

Cet échantillon répond à deux critères

1. il doit être suffisamment important (en nombre d'élèves et de classes); le nombre de classes à tester dépend notamment des disparités de réussite entre écoles (connues par des tests antérieurs³). Plus les classes sont différentes l'une de l'autre, plus il faut en tester;
2. il doit être représentatif de la diversité de notre système éducatif, diversité géographique et diversité de réseau.

En tenant compte de ces différents critères, un échantillon de 155 écoles a été sélectionné, où les différents réseaux et provinces sont représentés dans des proportions équivalentes à celles qu'ils occupent dans l'ensemble de la population scolaire. Ainsi, par exemple, on a déterminé qu'il fallait quatre écoles pour le réseau de la Communauté française dans la province de Brabant. Ces quatre écoles sont tirées au sort parmi l'ensemble des écoles du réseau de la Communauté française dans le Brabant.

Les 155 écoles de l'échantillon se répartissent comme suit

Réseau /Province	Communauté	Libre	Officiel subventionné	Total
Brabant	4	22	18	44
Hainaut	5	21	22	48
Liège	3	13	18	34
Luxembourg	2	4	6	12
Namur	2	9	6	17
Total	16	69	70	155
Soit en %	10 %	45 %	45 %	100 %

Les écoles retenues dans l'échantillon ont, via leur Inspecteur/trice, communiqué les grilles de correction et les questionnaires aux Universités. Parmi ces documents, un second tirage au sort a eu lieu, visant à sélectionner **une seule classe par école** (parmi les différentes classes des différentes implantations reçues). Quand une école comportait plusieurs petites classes de moins de 5 élèves, les différentes classes ont été retenues.

On se retrouve ainsi avec un échantillon de 159 classes, composé de 2915 élèves

C'est sur la base de ces 159 classes qu'ont été établis les résultats qui vont être présentés.

³ Notamment par l'épreuve d'évaluation externe organisée en 1994 en Se primaire et les études internationales de l'I.E.A. (Association Internationale pour l'Evaluation du Rendement Scolaire). Ces indices de disparité, pour la Communauté française de Belgique, sont assez importants, comparés à ceux d'autres pays.

La moyenne d'âge des élèves est de 8 ans 10 mois (écart type : 9 mois). Quatre-vingt pour cent des élèves sont à l'heure, 15 % sont en retard scolaire, 1,9 % des élèves sont "avancés".

4. Comparez les résultats de vos élèves à ceux de l'échantillon

Un des objectifs de l'évaluation externe est de vous donner l'occasion de situer les résultats de vos élèves par rapport à ceux d'un échantillon représentatif d'élèves dans la Communauté française. Pour réaliser cette comparaison, quelques traitements simples doivent être réalisés à partir des grilles de correction de votre classe:

1. Comparaison des résultats par question

- Pour chaque question, calculez le nombre d'élèves par code. Par exemple, à la question 1.1.1, quatre codes différents pouvaient être attribués aux réponses. Il vous suffit de calculer combien d'élèves ont obtenu le code 2, le code 1, le code 0 ou le code 9.
- Une fois calculées les fréquences par code, celles-ci doivent être transformées en pourcentage. Par exemple, à la question 1.1.1, il s'agit de calculer à quel pourcentage de votre classe correspond le nombre d'élèves ayant obtenu le code 2.
- Vous pouvez alors reporter les pourcentages dans la colonne " Votre classe " au sein du tableau de résultats des différentes questions. Vous aurez ainsi la possibilité de comparer les pourcentages que vous observez dans votre classe à ceux qui ont été observés dans l'échantillon de référence.

2. Comparaison des résultats à certains groupes de questions et à l'ensemble du test

- Pour réaliser cette comparaison, il est d'abord nécessaire de transformer les codes en scores dichotomiques: 1 pour une réponse correcte et 0 pour une réponse incorrecte. Le score 1 correspond au code 2 et le score 0 correspond à tous les autres codes (1,7,9 et 0)⁴
- Une fois cette transformation effectuée, il suffit d'additionner les résultats de chaque élève aux questions considérées. Par exemple, si vous voulez obtenir le score " lecture de l'heure ", vous devez additionner les scores aux questions 2.6, 2.7, 2.10.1, 2.10.2 et 2.10.3. **Si vous souhaitez calculer un score global pour l'ensemble de l'épreuve, vous devez additionner les résultats à toutes les questions, sauf ceux obtenus à la question 4.11.** Les raisons de l'exclusion de cette question sont explicitées dans la discussion des résultats de la situation 4.
- La fréquence des scores globaux est alors calculée pour l'ensemble de votre classe. Par exemple, vous calculerez le nombre d'élèves dont le score global " lecture de l'heure " est 5, 4, 3, 2, 1 ou 0.
- Une fois les fréquences calculées, celles-ci sont transformées en pourcentage.
- Vous pouvez enfin reporter les pourcentages dans la colonne " Votre classe " au sein du tableau de résultats ad hoc. Vous aurez ainsi la possibilité de comparer les pourcentages que vous observez dans votre classe à ceux qui ont été observés dans l'échantillon de référence.

⁴ Cette transformation, il faut le préciser, correspond à une correction assez "sévère", puisque les réussites partielles (code 1 à certaines questions) et l'absence de réponses découlant d'une erreur antérieure sont assimilées à des erreurs.

3. Comparaison des résultats obtenus dans différentes classes d'un même établissement

- lorsque plusieurs classes de 3e sont organisées dans un même établissement, il peut être intéressant d'évaluer l'hétérogénéité des performances entre celles-ci. Les mêmes calculs que ceux décrits ci-dessus doivent alors être réalisés pour l'ensemble des élèves de troisième, toutes classes confondues. Les pourcentages ainsi calculés sont reportés dans la colonne " Votre école " de chaque tableau. Il est dès lors possible de comparer les résultats de chaque classe par rapport aux résultats de l'ensemble des classes de la même école.

DEUXIEME PARTIE. ANALYSE DES RESULTATS PAR SITUATION

Introduction

Dans cette partie, les résultats sont présentés et discutés par situation. Pour chacune de celles-ci, une première section présente les pourcentages de résultats par question. Tous les tableaux sont organisés de la même manière. Le titre renseigne le numéro de la question et son énoncé tels qu'ils apparaissent dans la grille de correction. La première colonne mentionne de manière synthétique les différentes réponses possibles suivies du code correspondant entre parenthèses. Dans la deuxième colonne apparaissent les pourcentages de sujets de l'échantillon ayant donné la réponse correspondante. Les deux colonnes suivantes " Votre classe " et " Votre école " sont vides et peuvent être complétées par les pourcentages observés dans votre propre classe et dans l'ensemble des classes de 3e de votre école (voir la procédure décrite ci-dessus).

Une seconde section présente les résultats pour certains regroupements de questions. Ces regroupements varient d'une situation à l'autre en fonction de leur pertinence. Pour calculer les résultats à un ensemble de questions, les codes ont été transformés en scores dichotomiques: 1 pour une réponse correcte et 0 pour une réponse incorrecte. Le score 1 correspond au code 2 et le score 0 correspond à tous les autres codes (1,7,9 et 0). Les scores ont ensuite été additionnés pour l'ensemble des questions considérées. Le score maximum varie donc d'un groupe de questions à l'autre selon le nombre de questions prises en compte.

Enfin, une troisième section est réservée à la discussion des résultats présentés dans la ou les section(s) précédente(s). Certains liens sont également faits avec des résultats observés dans d'autres situations. Cette discussion ne constitue qu'une première analyse des performances des élèves. Elle permet toutefois déjà de mettre en évidence leurs caractéristiques majeures. Une analyse plus détaillée des erreurs commises par les élèves est en cours de réalisation et sera présentée dans le document pédagogique qui vous parviendra au mois de septembre.

1. Résultats à la situation 1

1.1 Présentation des résultats par question

Question 1.1.1. Copie du nom, prénom, adresse.

	Echantillon	Votre classe	Votre école
aucun élément ne manque (2)	81,7		
il manque un élément (1)	4,1		
il manque plus d'un élément (0)	8,8		
n'a rien écrit à droite (9)	5,4		

Question 1.1.2. Copie avec majuscules.

	Echantillon	Votre classe	Votre école
emploi correct (2)	78,5		
emploi irrégulier (1)	12,3		
aucune majuscule (0)	2,6		
n'a rien écrit à droite (9)	6,7		

Question 1.1.3. Copie et orthographe.

	Echantillon	Votre classe	Votre école
orthographe correcte (2)	70,0		
une erreur ou substitution (1)	15,3		
deux erreurs ou plus (0)	8,1		
n'a rien écrit à droite (9)	6,6		

Question 1.1.4. Adresse de la carte postale.

	Echantillon	Votre classe	Votre école
adresse et signature (2)	64,1		
adresse ou signature (1)	25,0		
ni adresse, ni signature (0)	8,4		
n'a rien écrit à gauche (9)	2,5		

Question 1.1.5. Contenu de la carte postale.

	Echantillon	Votre classe	Votre école
un des deux renseignements demandés (2)	81,6		
recopie du contenu (1)	8,5		
copie hors sujet (0)	6,9		
n'a rien écrit à gauche (9)	3,0		

Question 1.1.6. Adresse de la piscine.

	Echantillon	Votre classe	Votre école
adresse correcte (2)	58,7		
adresse incorrecte (1)	9,4		
pas d'adresse (0)	28,8		
n'a rien écrit à gauche (9)	3,1		

Question 1.1.7. Heure d'ouverture de la piscine.

	Echantillon	Votre classe	Votre école
heure d'ouverture correcte (2)	52,3		
heure d'ouverture incorrecte (1)	32,2		
pas d'heure (0)	12,6		
n'a rien écrit à gauche (9)	2,9		

Question 1.2.1. Prix d'un bain adulte.

	Echantillon	Votre classe	Votre école
prix correct (2)	98,6		
prix incorrect (1)	1,1		
pas de réponse (0)	0,3		

Question 1.2.2. Prix d'une leçon de natation.

	Echantillon	Votre classe	Votre école
prix correct (2)	97,6		
prix incorrect (1)	2,1		
pas de réponse (0)	0,3		

Question 1.2.3. Prix de l'abonnement.

	Echantillon	Votre classe	Votre école
prix correct (2)	97,4		
prix incorrect (1)	2,0		
pas de réponse (0)	0,6		

Question 1.3. Heure d'ouverture de la piscine.

	Echantillon	Votre classe	Votre école
heure d'ouverture correcte (2)	86,9		
heure d'ouverture incorrecte (1)	12,0		
pas de réponse (0)	1,0		

Question 1.4.1. Adresse de Julie.

	Echantillon	Votre classe	Votre école
adresse correcte (2)	82,3		
adresse incorrecte (1)	14,4		
pas de réponse (0)	3,4		

Question 1.4.2. Adresse de Laurent.

	Echantillon	Votre classe	Votre école
adresse correcte (2)	87,8		
adresse incorrecte (1)	9,5		
pas de réponse (0)	2,7		

Question 1.4.3. Adresse de la piscine.

	Echantillon	Votre classe	Votre école
adresse correcte (2)	82,7		
adresse incorrecte (1)	13,5		
pas de réponse (0)	3,9		

1.2 Discussion

- **Items 1.1.1. à 1.1.3. : copie de la partie droite de la carte postale.**

Ce groupe d'items est relativement bien réussi. En situation de copie, 80,2 % des élèves n'oublient aucun élément des nom et adresse modèles. Quelques élèves (4 %) oublient un élément. Parmi les élèves oubliant plus d'un élément et n'ayant rien écrit à droite, il faut en outre compter un certain nombre d'élèves ayant écrit d'autres nom et adresse ou encore des élèves ayant écrit à gauche. Des codes n'étant pas prévus pour ces situations, les enseignants ont été contraints de se glisser dans les cases prévues (ces cas sont toutefois marginaux).

Les majuscules sont correctement utilisées par 77 % des élèves. On note un emploi irrégulier chez 12 % des élèves. Très rares (2,5 %) sont en revanche ceux qui oublient toutes les majuscules (sauf à Laurent).

Les résultats pour l'orthographe ne sont pas entièrement satisfaisants : seuls 69 % des élèves copient les noms et adresses sans commettre d'erreur. Quinze pour-cent des élèves commettent une erreur ou substitution et 8 % deux erreurs ou davantage. En situation de copie, ce résultat apparaît comme un peu décevant.

- **Items 1.1.4. à 1.1.7. : rédaction de la carte postale.**

En présence d'un modèle et d'une consigne orale, 63 % des élèves réussissent pleinement et 24,5 % partiellement à écrire une carte postale en identifiant le destinataire (Laurent) et l'auteur de la carte (Julie). Seuls 8 % des élèves produisent un écrit dont ces deux éléments sont absents.

Par ailleurs, 80 % des élèves semblent avoir bien compris la tâche qui leur était assignée : en s'appuyant sur le modèle, ils fournissent au moins une des deux informations qu'il fallait aller rechercher à d'autres endroits du document. Seuls 8 % des élèves n'arrivent pas à se dégager du contenu du modèle (qu'ils recopient); 7 % des élèves sont par ailleurs complètement hors du sujet, n'ayant vraisemblablement pas compris la consigne.

Sur le plan du contenu de la carte, 58 % des élèves se montrent capables de fournir l'adresse correcte de la piscine, en consultant la page en regard. La majorité des élèves qui fournissent l'information le font correctement. Seuls 9 % des élèves donnent une adresse incorrecte, mais ils sont 28 % à ne pas fournir l'information demandée.

Les résultats relatifs aux heures d'ouverture sont un peu différents. Moins d'élèves (51 %) fournissent l'information correcte; moins nombreux sont les élèves (12 %) qui ne la fournissent pas. En revanche, près d'un tiers des élèves fournissent une information peu ciblée, "trop large" par rapport à la consigne, en donnant toutes les heures d'ouverture du samedi, voire celles des autres jours. Ce comportement, s'il ne s'accorde pas précisément avec la consigne, ne peut toutefois être considéré comme inadéquat : fournir trop d'informations, dans un cas comme celui-ci, ne nuit pas à l'efficacité du message.

- **Items 1.2. à 1.4. : réponses aux questions portant sur des compétences isolées.**

Pour rappel, ce groupe de questions visait à tenter d'évaluer, pour les trois éléments d'information intervenant dans la carte postale, dans quelle mesure le fait de poser des questions hors contexte, isolément, ou dans le contexte de la carte postale, influençait les résultats.

Les résultats à ce groupe de questions sont dans l'ensemble excellents. Pour les prix, la réussite dépasse 95 %. Pour l'heure d'ouverture de la piscine, le taux de bonnes réponses est de 85 %. Pour les adresses enfin, la réussite dépasse également les 80 %.

Ces résultats sont intéressants car ils montrent à quel point le contexte dans lequel une compétence est évaluée (un contexte de vie simulé, une question de recherche d'informations plus ciblée, une question ouverte, une Q.C.M.) peut influencer le degré de réussite. Ceux-ci semblent aussi montrer que les taux de réussite enregistrés dans la situation carte postale, sur le plan du contenu (items 1.1.6. et 1.1.7.) ne tiennent pas à une incompétence à repérer et fournir les informations demandées, puisque la majorité des élèves savent le faire lorsqu'on le leur demande expressément, mais une difficulté à accomplir cela dans un contexte comportant de multiples exigences simultanées.

La maîtrise de compétences à l'état isolé ne garantit donc pas que ces compétences soient mises en œuvre dans des tâches qui font intervenir en parallèle plusieurs de ces compétences. C'est pourquoi il est souhaitable de confronter régulièrement l'élève à de telles tâches, en le guidant dans l'analyse et la planification de ces tâches plus complexes (Que dois-je faire? Par où commencer?...). Clairement, la réussite du type de tâche illustré ici représente autre chose que l'addition de compétences isolées.

2. Résultats à la situation 2

2.1 Présentation des résultats par question

Question 2.1. Profondeur de l'eau

	Echantillon	Votre classe	Votre école
Profondeur correcte (2)	63,5		
Profondeur incorrecte (1)	36,1		
pas de réponse (0)	0,4		

Question 2.2. Nombre de longueurs

	Echantillon	Votre classe	Votre école
Nombre correct (2)	62,0		
Nombre incorrect (1)	37,5		
Pas de réponse (0)	0,5		

Question 2.3. Phrases correctes

	Echantillon	Votre classe	Votre école
2 phrases correctes (2)	68,5		
Autres phrases entourées (1)	31,1		
pas de réponse (0)	0,4		

Question 2.4. Mesure de la longueur

	Echantillon	Votre classe	Votre école
Mesure correcte (2)	81,1		
Mesure incorrecte (1)	15,9		
pas de réponse (0)	3,0		

Question 2.5. Nombre de mètres.

	Echantillon	Votre classe	Votre école
nombre de mètres correct (2)	71,6		
nombre de mètres incorrect (1)	21,1		
pas de réponse (0)	0,4		
erreur à la 2.4 (9)	6,9		

Question 2.6. Heure d'entrée.

	Echantillon	Votre classe	Votre école
heure d'entrée correcte (2)	91,5		
heure d'entrée incorrecte (1)	8,0		
pas de réponse (0)	0,5		

Question 2.7. Heure de sortie.

	Echantillon	Votre classe	Votre école
heure de sortie correcte (2)	66,9		
heure de sortie incorrecte (1)	32,0		
pas de réponse (0)	1,1		

Question 2.8. Durée⁵.

	Echantillon	Votre classe	Votre école
durée correcte (2)	52,6		
durée incorrecte (1)	20,7		
pas de réponse (0)	1,8		
erreur à la 2.6 ou 2.7 (9)	24,9		

⁵ Pour cet item, certains enfants ont trouvé la durée correcte malgré une erreur à l'une des deux questions précédentes. Le cas n'était pas prévu dans la grille de correction. Certains enseignants ont codé le cas 2, d'autres 9. Nous estimons préférable de trancher en **faveur de l'enfant** et d'utiliser le code 2. Si vous avez procédé autrement, il vous est toujours loisible de recorriger cette question.

Question 2.9.1. Relier Lavabo-20 cm.

	Echantillon	Votre classe	Votre école
lien correct (2)	85,7		
lien incorrect (1)	13,6		
pas de réponse (0)	0,8		

Question 2.9.2. Relier Bassin-2m50.

	Echantillon	Votre classe	Votre école
lien correct (2)	77,3		
lien incorrect (1)	22,0		
pas de réponse (0)	0,7		

Question 2.9.3. Relier Baignoire-60 cm.

	Echantillon	Votre classe	Votre école
lien correct (2)	75,8		
lien incorrect (1)	23,4		
pas de réponse (0)	0,7		

Question 2.10.1. Lecture d'heure (8h).

	Echantillon	Votre classe	Votre école
lecture d'heure correcte (2)	87,7		
lecture d'heure incorrecte (1)	11,9		
pas de réponse (0)	0,4		

Question 2.10.2. Lecture d'heure (10 h 15).

	Echantillon	Votre classe	Votre école
lecture d'heure correcte (2)	76,2		
lecture d'heure incorrecte (1)	23,0		
pas de réponse (0)	0,8		

Question 2.10.3. Lecture d'heure (3h30).

	Echantillon	Votre classe	Votre école
lecture d'heure correcte (2)	79,8		
lecture d'heure incorrecte (1)	19,4		
pas de réponse (0)	0,8		

Question 2.11.1. Durée du film (2h15).

	Echantillon	Votre classe	Votre école
durée correcte (2)	54,1		
durée incorrecte (1)	43,3		
pas de réponse (0)	2,6		

Question 2.11.2. Durée du film (2h30).

	Echantillon	Votre classe	Votre école
durée correcte (2)	57,8		
durée incorrecte (1)	37,5		
pas de réponse (0)	4,7		

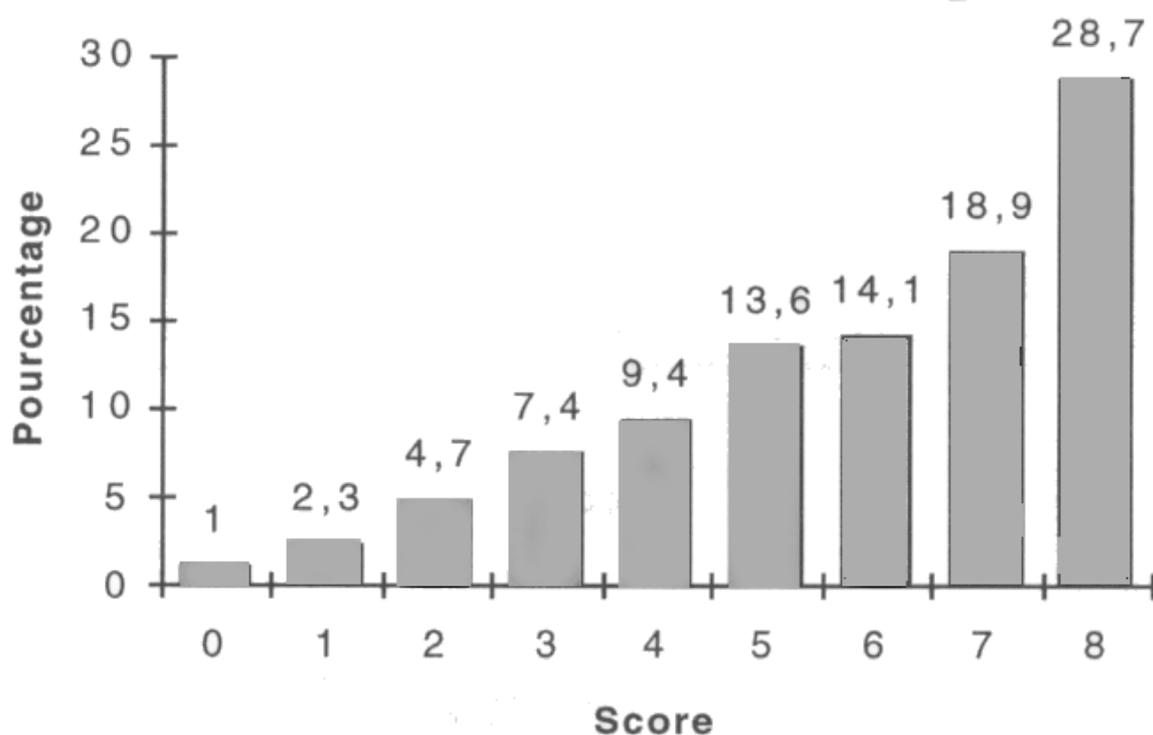
Question 2.11.3. Durée du film (30 minutes).

	Echantillon	Votre classe	Votre école
durée correcte (2)	36,6		
durée incorrecte (1)	57,5		
pas de réponse (0)	5,8		

2.2 Présentation des résultats par groupes de questions

2.2.1 Mesure de l'espace

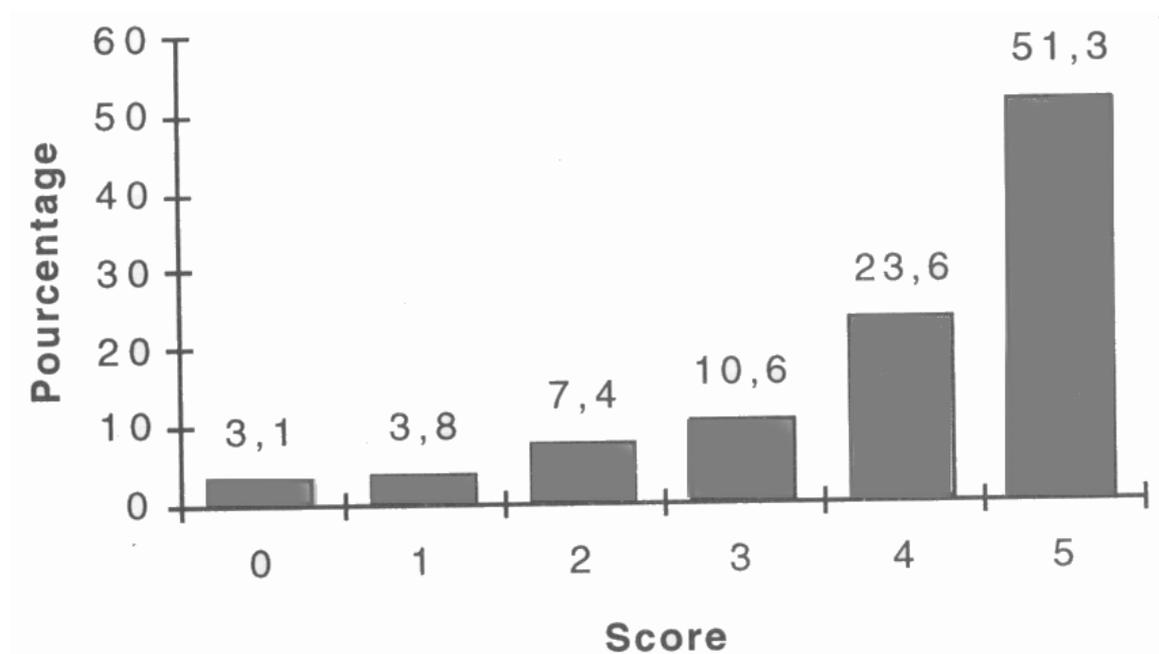
Les résultats des différentes questions relatives à la mesure de l'espace ont été additionnés pour constituer un seul score total. Sont incluses dans ce score les questions 2.1, 2.2, 2.3, 2.4, 2.5, 2.9.1, 2.9.2 et 2.9.3. Pour chaque question, les résultats ont été transformés en 'score dichotomique (1 = réussite; 0 = échec). Par conséquent, le score maximum en « mesure de l'espace » est de 8 points. Le graphique présente le pourcentage de sujets de l'échantillon pour chaque score en "mesure de l'espace". Ces mêmes pourcentages sont repris dans le tableau où vous pouvez reporter les pourcentages observés dans votre classe et dans votre école.



Score	Échantillon	Votre classe	Votre école
0	1,0		
1	2,3		
2	4,7		
3	7,4		
4	9,4		
5	13,6		
6	14,1		
7	18,9		
8	28,7		

2.2.2 Lecture de l'heure

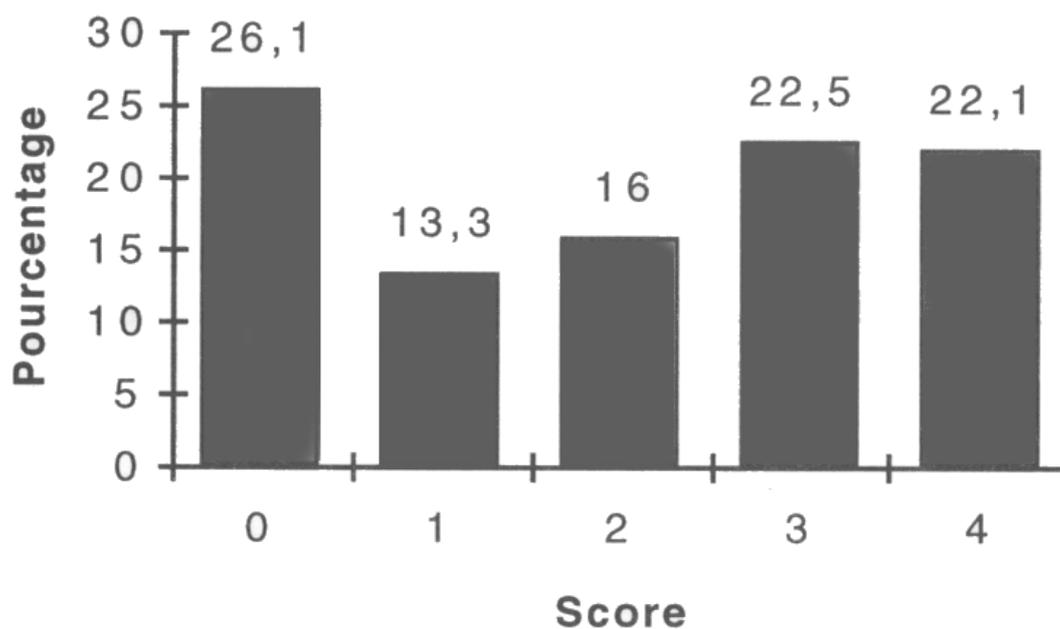
Les résultats des différentes questions relatives à la lecture de l'heure ont été additionnés pour constituer un seul score total. Sont incluses dans ce score les questions 2.6, 2.7, 2.10.1, 2.10.2 et 2.10.3. Pour chaque question, les résultats ont été transformés en score dichotomique (1 = réussite; 0 = échec). Par conséquent, le score maximum en " mesure de l'espace " est de 5 points. Le graphique présente le pourcentage de sujets de l'échantillon pour chaque score en " lecture de l'heure ". Ces mêmes pourcentages sont repris dans le tableau où vous pouvez reporter les pourcentages observés dans votre classe et dans votre école.



Score	Echantillon	Votre classe	Votre école
0	3,1		
1	3,8		
2	7,4		
3	10,6		
4	23,9		
5	51,3		

2.2.3 Calcul de la durée

Les résultats des différentes questions relatives au calcul de la durée ont été additionnés pour constituer un seul score total. Sont incluses dans ce score les questions 2.8, 2.11.1, 2.11.2 et 2.11.3. Pour chaque question, les résultats ont été transformés en score dichotomique (1 = réussite; 0 = échec). Par conséquent, le score maximum en " calcul de la durée " est de 4 points. Le graphique présente le pourcentage de sujets de l'échantillon pour chaque score en " calcul de la durée ". Ces mêmes pourcentages sont repris dans le tableau où vous pouvez reporter les pourcentages observés dans votre classe et dans votre école.



Score	Echantillon	Votre classe	Votre école
0	26,1		
1	13,3		
2	16,0		
3	22,5		
4	22,1		

2.3 Discussion

Les questions présentées dans la situation 2 mettent en œuvre un certain nombre de compétences de base relatives à la mesure de l'espace et du temps. Certaines sont incluses dans le contexte de l'histoire de Laurent et Julie. D'autres sont présentées de manière isolée.

Les questions concernant la mesure de l'espace évaluent la capacité à estimer des grandeurs familières, la connaissance de concepts " longueur " et " largeur " et la capacité de réaliser des opérations simples sur des longueurs exprimées en mètres. Globalement, les résultats sont satisfaisants sans être brillants. Cinquante pour-cent des sujets obtiennent 6 points ou plus (maximum) à l'ensemble des questions évaluant les compétences relatives à la mesure de l'espace.

On constate cependant qu'un nombre assez important d'élèves ne maîtrisent pas ces compétences. Plus d'un tiers des élèves n'estiment pas correctement la profondeur de la piscine. Il est vrai que cette tâche est plus complexe qu'elle n'y paraît au premier abord. L'élève doit connaître sa propre taille et inférer que celle de Julie est proche de la sienne. Sur cette base, il doit ensuite éliminer les deux alternatives non vraisemblables ("2 mètres" et "3 mètres") et constater que "1 mètre" est la seule alternative vraisemblable. Cette question est particulièrement intéressante car elle révèle la relation que le sujet peut établir entre son propre corps et les unités de mesure de l'espace. Cette mise en relation est importante car elle sous-tend la capacité d'estimer de nombreuses mesures de l'espace. Les estimations de profondeur sont un peu mieux réussies dans les questions isolées (entre 77% et 86% de réussite), sans doute parce que les choix proposés sont plus contrastés et que la discrimination entre les mesures est dès lors plus simple à réaliser.

La connaissance de concepts « longueur » et « largeur » est également inférieure à ce que l'on pouvait espérer à ce moment de la scolarité. Près d'un tiers des sujets échouent en effet à la question 2.3. Enfin, les opérations sur les grandeurs donnent lieu à des résultats contrastés: 62% de réussite à la question 2.2 et 81 % de réussite à la question 2.5. Dans le premier cas, une division doit être effectuée alors que dans le second cas, il faut multiplier. La première tâche est d'évidence plus difficile que la seconde. Par ailleurs, certains échecs peuvent s'expliquer par l'incapacité de trouver la longueur correcte (15% d'échecs à la question 2.4), soit par incompréhension de cette notion, soit par incapacité de lire correctement cette mesure sur le dessin.

Les questions évaluant la capacité de lire l'heure sur une horloge analogique sont moyennement réussies, qu'elles se situent en contexte ou qu'elles soient isolées. Le fait que plus de la moitié des élèves obtiennent le maximum de points à cet ensemble de questions ne doit pas masquer les performances insatisfaisantes de 25% des élèves dont les résultats s'échelonnent de 0 à 3 points. La tâche demandée (lecture de l'heure, de la demi-heure et du quart d'heure) est simple et largement entraînée en 2ème année. Par conséquent, on serait en droit d'attendre un pourcentage de réussite supérieur à 80% à l'ensemble des questions.

On observe une certaine gradation dans la tâche demandée: la lecture de l'heure est mieux réussie que la lecture de la demi-heure qui, elle-même, est mieux réussie que la lecture du quart d'heure. La lecture de la demi-heure est sensiblement mieux réussie lorsque l'horloge est présentée de manière isolée que lorsqu'elle est présentée dans le contexte de l'histoire de Laurent et Julie. Dans ce dernier cas, la difficulté semble provenir du repérage de l'horloge à prendre en compte pour déterminer l'heure de sortie des enfants.

Le calcul de la durée représente une difficulté importante pour un grand nombre d'élèves. Il s'agit d'évidence d'une compétence encore en construction. Plus d'un quart des élèves n'obtiennent aucun point aux quatre questions évaluant la durée. Ils ne sont que 22% à obtenir le maximum des points à cet ensemble de questions.

3. Résultats à la situation 3

3.1 Présentation des résultats par question

Question 3.1. Argent de Julie en poche.

	Echantillon	Votre classe	Votre école
argent correct (2)	89,0		
argent incorrect (1)	10,2		
pas de réponse (0)	0,8		

Question 3.2. Prix de la glace de Julie.

	Echantillon	Votre classe	Votre école
Prix correct (2)	81,3		
démarche correcte, prix incorrect ou manquant (7)	1,9		
Prix incorrect (1)	15,4		
Pas de réponse (0)	1,4		

Question 3.3. Coloriage des pièces.

	Echantillon	Votre classe	Votre école
coloriage correct (2)	84,4		
coloriage incorrect (1)	7,4		
pas de réponse (0)	1,5		
erreur à la 3.2. (9)	6,7		

Question 3.4. Reste de Julie.

	Echantillon	Votre classe	Votre école
reste correct (2)	71,5		
démarche correcte, reste incorrect ou manquant (7)	1,7		
reste incorrect (1)	13,5		
pas de réponse (0)	2,3		
erreur à la question 3.1 et/ou 3.2 (9)	11,1		

Question 3.5. La monnaie de retour.

	Echantillon	Votre classe	Votre école
argent correct (2)	52,6		
démarche correcte, argent incorrect ou manquant (7)	9,7		
argent incorrect (1)	33,5		
pas de réponse (0)	4,1		

Question 3.6. Calcul (32+7).

	Echantillon	Votre classe	Votre école
calcul correct (2)	94,7		
calcul incorrect (1)	5,0		
pas de réponse (0)	0,3		

Question 3.7. Calcul (20+15).

	Echantillon	Votre classe	Votre école
calcul correct (2)	97,0		
calcul incorrect (1)	2,7		
pas de réponse (0)	0,3		

Question 3.8. Calcul (23+4).

	Echantillon	Votre classe	Votre école
calcul correct (2)	93,6		
calcul incorrect (1)	6,1		
pas de réponse (0)	0,4		

Question 3.9. Calcul (20+20+20).

	Echantillon	Votre classe	Votre école
calcul correct (2)	97,1		
calcul incorrect (1)	2,6		
pas de réponse (0)	0,4		

Question 3.10. Calcul (13+7+10).

	Echantillon	Votre classe	Votre école
calcul correct (2)	93,4		
calcul incorrect (1)	6,1		
pas de réponse (0)	0,5		

Question 3.11. Calcul (2 x 20).

	Echantillon	Votre classe	Votre école
calcul correct (2)	94,4		
calcul incorrect (1)	4,9		
pas de réponse (0)	0,7		

Question 3.12. Calcul (7 x 5).

	Echantillon	Votre classe	Votre école
calcul correct (2)	92,3		
calcul incorrect (1)	6,9		
pas de réponse (0)	0,8		

Question 3.13. Calcul (18-6).

	Echantillon	Votre classe	Votre école
calcul correct (2)	91,1		
calcul incorrect (1)	7,9		
pas de réponse (0)	1,1		

Question 3.14. Calcul (50-20).

	Echantillon	Votre classe	Votre école
calcul correct (2)	96,7		
calcul incorrect (1)	2,6		
pas de réponse (0)	0,7		

Question 3.15. Calcul (50-35).

	Echantillon	Votre classe	Votre école
calcul correct (2)	66,7		
calcul incorrect (1)	32,3		
pas de réponse (0)	1,0		

Question 3.16. Opérations (20 x 255)

	Echantillon	Votre classe	Votre école
opération correcte (2)	86,0		
opération incorrecte (1)	13,1		
pas de réponse (0)	1,0		

Question 3.17. Opérations (20 - 5)

	Echantillon	Votre classe	Votre école
opération correcte (2)	87,7		
opération incorrecte (1)	11,3		
pas de réponse (0)	1,0		

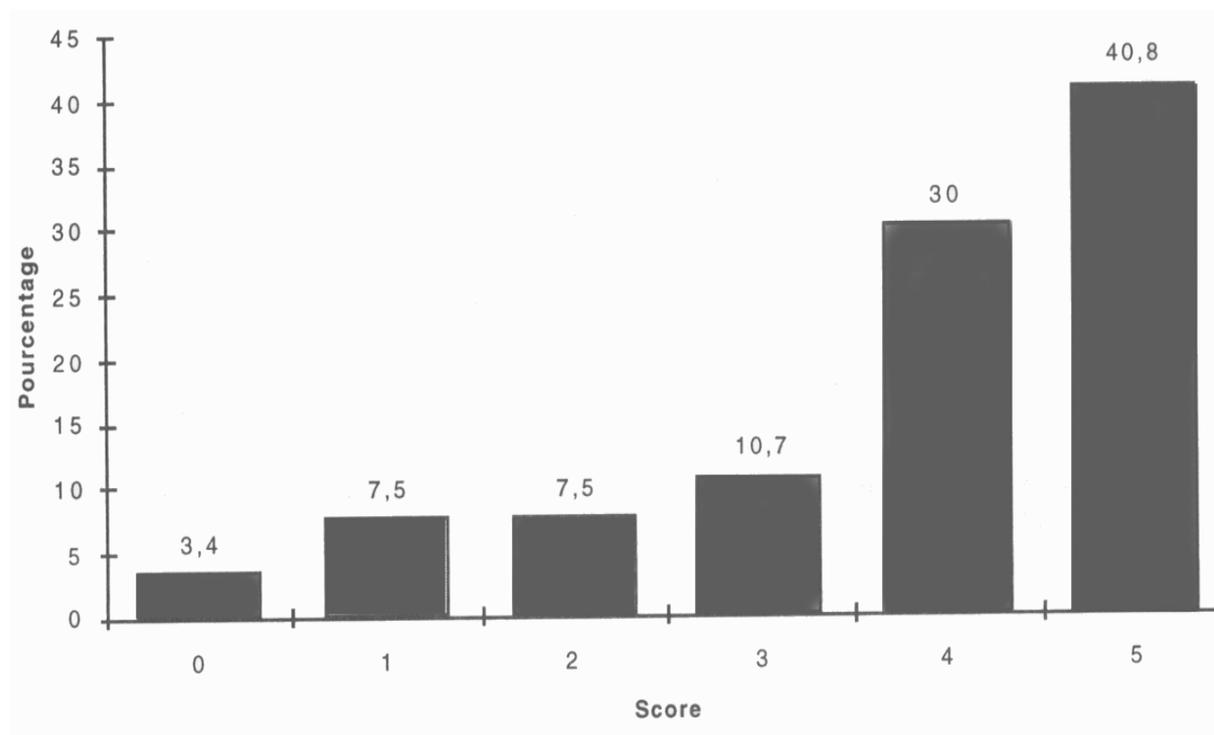
Question 3.18. Opérations (20 + 5)

	Echantillon	Votre classe	Votre école
opération correcte (2)	82,1		
opération incorrecte (1)	16,9		
pas de réponse (0)	1,1		

3.2 Présentation des résultats par groupes de questions

3.2.1 Résultats pour l'ensemble des questions en contexte

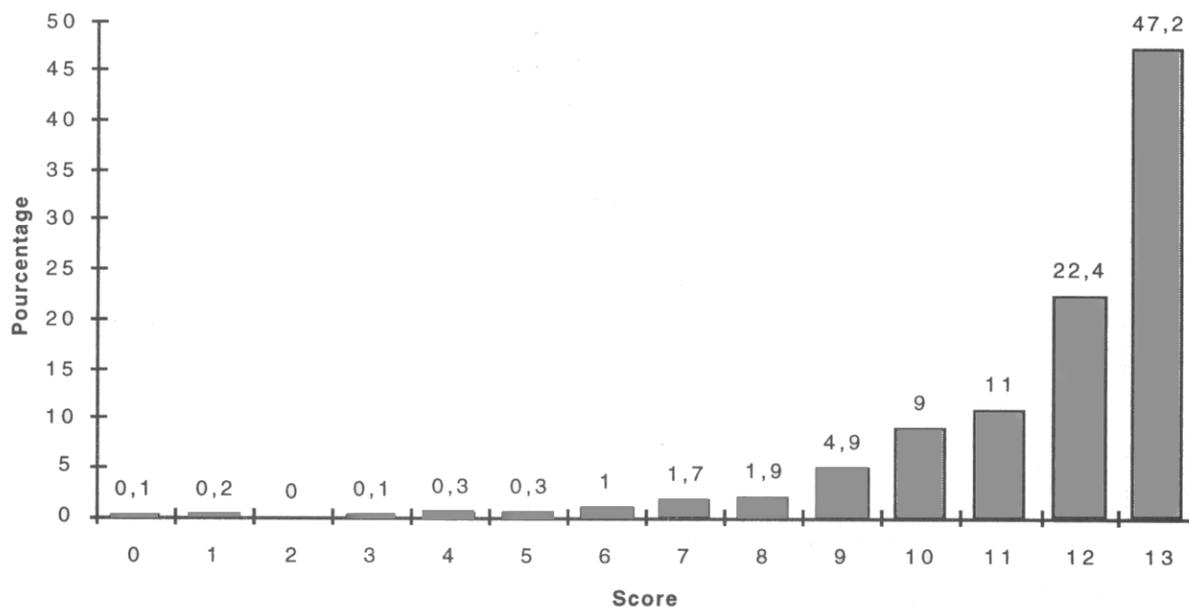
Les résultats des différentes questions évaluant les compétences arithmétiques en contexte ont été additionnés pour constituer un seul score total. Sont incluses dans ce score les questions 3.1, 3.2, 3.3, 3.4, et 3.5. Pour chaque question, les résultats ont été transformés en score dichotomique (1 = réussite; 0 = échec). Par conséquent, le score maximum est de 5 points. Le graphique présente le pourcentage de sujets de l'échantillon pour chaque score. Ces mêmes pourcentages sont repris dans le tableau où l'enseignant peut reporter les pourcentages observés dans sa classe et dans son école.



Score	Échantillon	Votre classe	Votre école
0	3,4		
1	7,5		
2	7,5		
3	10,7		
4	30,0		
5	40,8		

3.2.2 Résultats pour l'ensemble des questions isolées

Les résultats des différentes questions évaluant les compétences arithmétiques isolées ont été additionnés pour constituer un seul score total. Sont incluses dans ce score les questions 3.6, 3.7, 3.8, 3.9, 3.10, 3.11, 3.12, 3.13, 3.14, 3.15, 3.16, 3.17 et 3.18. Pour chaque question, les résultats ont été transformés en score dichotomique (1 = réussite; 0 = échec). Par conséquent, le score maximum est de 13 points. Le graphique ci-dessus présente le pourcentage de sujets de l'échantillon pour chaque score. Ces mêmes pourcentages sont repris dans le tableau ci-dessous où l'enseignant peut reporter les pourcentages observés dans sa classe et dans son école.



Score	Echantillon	Votre classe	Votre école
5 et -	1,0		
6	1,0		
7	1,7		
8	1,9		
9	4,9		
10	9,0		
11	11,0		
12	22,0		
13	47,2		

3.3 Discussion

Les questions de la situation 3 évaluent la maîtrise des opérations d'addition, de soustraction et de multiplication des nombres entiers naturels inférieurs à 100. Ces opérations sont présentées soit de manière isolée, soit dans un contexte où elles sont intégrées dans une procédure de résolution de problème.

Les questions isolées (3.6 à 3.18) sont globalement bien réussies: plus de 80% des élèves obtiennent un score total entre 11 et 13 points. Une analyse plus fine des résultats révèle cependant une certaine variabilité dans la maîtrise des compétences évaluées. Les opérations d'addition et de multiplication ne posent aucune difficulté pour la grande majorité des élèves. Plus de 90% de ceux-ci les réussissent toutes. Par contre, les trois questions demandant le calcul d'une différence conduisent à des résultats plus contrastés. Les calculs " 18-6 " et " 50-20 " sont réussis par plus de 90% des élèves mais < 50-35 > n'est réussi que par 66,7% de ceux-ci.

Comment expliquer un tel écart entre les pourcentages de réussite? Dans les trois cas, le calcul doit être effectué mentalement. Une telle tâche demande une compréhension du système de numération de position en base 10, la compréhension de l'opération de soustraction et la connaissance de certains résultats d'opérations courantes stockés en mémoire. Cette dernière connaissance n'est pas indispensable mais facilite le calcul mental en réduisant la charge cognitive de la tâche. Le calcul de « 18-6 » demande de retirer 6 unités de 8 unités, la dizaine restant inchangée. Le calcul « 50-20 » suppose, quant à lui, de comprendre qu'il s'agit de soustraire 2 dizaines de 5 dizaines et donc d'effectuer « 5 - 2 ». Dans les deux cas, le calcul est simple et le résultat existe souvent dans la mémoire des sujets. Par contre, « 50-35 » présente une difficulté particulière car il n'est pas possible de retirer 5 unités de 0. Comment les sujets procèdent-ils pour résoudre ce problème par calcul mental? Très peu d'élèves appliquent mentalement les algorithmes de calcul écrit⁶. Ils préfèrent utiliser d'autres procédures moins lourdes. Certaines de ces procédures sont enseignées, mais d'autres sont inventées par les sujets eux-mêmes. Dans le cas de « 50-35 », certains sujets transforment le calcul initial en « combien faut-il ajouter à 35 pour arriver à 50 ? ». Il compte alors de 35 à 50 ou, plus simplement, additionne 5 (ce qu'il faut pour passer de 35 à 40) et 10 (ce qu'il faut pour passer de 40 à 50). D'autres procédures efficaces sont également possibles. Par exemple, on peut décomposer 35 en « 30+5 », effectuer « 50-30 » puis « 20-5 »... Toutes ces procédures font appel à la capacité d'exprimer un même nombre de différentes manières. Par ailleurs, elles requièrent beaucoup d'espace mental pour effectuer les différents calculs et retenir les résultats intermédiaires. Le risque d'erreur est dès lors plus important et il n'est pas étonnant que la question 3.5 soit moins bien réussie que les deux précédentes.

La mise en correspondance d'un problème exprimé verbalement et de son expression mathématique est une tâche un peu plus difficile qu'effectuer une opération. Sans doute cette compétence est-elle moins entraînée en classe que le calcul routinier. En effet, lorsque cette tâche est demandée de manière isolée, les pourcentages de réussite varient de 82 à 88%.

Les résultats des problèmes posés dans le contexte de l'histoire de Laurent et Julie (3.1 à 3.5) sont légèrement moins bons que les résultats obtenus aux questions isolées. Ce n'est guère étonnant. L'élève doit en effet transférer des compétences « scolaires » dans une situation de vie. Il doit également coordonner différentes procédures élémentaires au sein d'une procédure plus complexe. Nous avons vu plus haut qu'effectuées isolément, les additions et les multiplications sont réussies par plus de 90% des élèves. Aux questions 3.1 et 3.2, il faut également effectuer une addition ou une multiplication simple. A la première question, la réussite est de 89% et, à la seconde, elle est de 81%. Cette différence de réussites des mêmes opérations s'explique par le fait que, dans les problèmes, il faut en plus traduire la question verbale sous forme d'opération arithmétique et rechercher ailleurs les informations nécessaires à la solution.

⁶ English, L.D. & Halford, G.S. (1995). *Mathematics education*. Mahawah, NJ: Lawrence Erlbaum.

L'opération de soustraction soulève les mêmes difficultés au sein d'un problème qu'effectuée de manière isolée. Ainsi, le calcul de la monnaie que le marchand doit rendre à Laurent (« 100-45 ») n'est réussi que par 53% des élèves. Dix pour-cent des élèves posent bien l'opération, mais ne peuvent l'effectuer correctement. Le calcul de l'argent qu'il reste à Julie (question 3.4) est, lui, mieux réussi car il suffit de retirer 3 dizaines de 4 dizaines. Le fait que ce calcul ne soit malgré tout pas réussi par près de 30% des élèves provient de l'effet cumulatif des erreurs aux questions précédentes. En effet, si l'élève n'a pu calculer correctement l'argent de Julie et le prix de la glace, il ne peut pas non plus déterminer correctement l'argent qui restera une fois la glace payée⁷.

⁷ Du moins en principe. Lors de la correction, il est apparu que des enfants réussissaient à trouver la réponse correcte **malgré une erreur antérieure**. Ce cas n'était en rien prévu dans les critères de correction; certains enseignants ont dans ce cas codé 2, d'autres ont codé 9. Nous estimons préférable de coder **en faveur de l'enfant** (code 2). Si vous avez procédé autrement, il vous est toujours loisible de recorriger cette question.

4. Résultats à la situation 4

4.1 Présentation des résultats par question

Question 4.1. Le premier joueur.

	Echantillon	Votre classe	Votre école
réponse correcte (2)	88,4		
réponse incorrecte (1)	3,6		
pas de réponse (0)	8,1		

Question 4.2.1. Emplacement du pion de Julie après le premier coup.

	Echantillon	Votre classe	Votre école
emplacement correct (2)	96,6		
emplacement incorrect (1)	1,4		
pas de réponse (0)	2,0		

Question 4.2.2. Emplacement du pion de Laurent après le premier coup.

	Echantillon	Votre classe	Votre école
emplacement correct (2)	96,8		
emplacement incorrect (1)	1,3		
pas de réponse (0)	1,9		

Question 4.3. Emplacement du pion de Julie après le deuxième coup.

	Echantillon	Votre classe	Votre école
emplacement correct (2)	61,5		
emplacement incorrect (1)	35,7		
pas de réponse (0)	2,9		

Question 4.4. Emplacement du pion de Laurent après le deuxième coup.

	Echantillon	Votre classe	Votre école
emplacement correct (2)	53,3		
emplacement sur case 13 (7)	16,5		
emplacement incorrect (1)	26,8		
pas de réponse (0)	3,3		

Question 4.5. Nombre de points pour que Julie gagne la partie.

	Echantillon	Votre classe	Votre école
nombre de points correct (2)	58,0		
nombre de point incorrect (1)	13,0		
pas de réponse (0)	1,4		
erreur à la question 4.3 (9)	27,6		

Question 4.6. Possibilité pour Laurent de gagner la partie.

	Echantillon	Votre classe	Votre école
réponse correcte (2)	77,1		
réponse incorrecte (1)	7,9		
pas de réponse (0)	1,3		
erreur à la question 4.4 (9)	13,8		

Question 4.7. Justification pour que Laurent gagne la partie.

	Echantillon	Votre classe	Votre école
Justification précise (2)	28,1		
Justification vague (1)	48,3		
Pas de réponse (0)	5,6		
Erreur à la question 4.6 (9)	18,0		

Question 4.8. Addition de points.

	Echantillon	Votre classe	Votre école
Réponse correcte (2)	85,3		
Réponse incorrecte (1)	13,7		
Pas de réponse (0)	1,0		

Question 4.9. Compréhension de la règle.

	Echantillon	Votre classe	Votre école
compréhension correcte (2)	83,9		
compréhension incorrecte (1)	14,6		
pas de réponse (0)	1,9		

Question 4.10. Emplacement après avoir passé un tour.

	Echantillon	Votre classe	Votre école
emplacement correct (2)	74,4		
emplacement incorrect (1)	24,0		
pas de réponse (0)	1,6		

Question 4.11. Comment faire pour gagner la partie.

	Echantillon	Votre classe	Votre école
réponse correcte (2)	44,6		
réponse incorrecte (1)	51,4		
pas de réponse (0)	4,0		

4.2 Discussion

- **Items 4.1. et 4.2.**

La réussite à ces deux items qui demandent une simple lecture et l'exécution d'une consigne, est excellente: 84,5 % dans le cas du 1er item et 92 % dans le cas du 2e item.

- **Items 4.3. et 4.4.**

La complexité de ces deux items est nettement plus marquée puisqu'il s'agit dans ce cas de dénombrer les points sur deux dés et de faire avancer les pions du nombre de cases voulu en tenant compte de la règle du jeu figurant à la page précédente.

Dans le cas de Julie, aucune règle spéciale n'intervient (pas de case spéciale) et le retour à la règle du jeu est donc superflu. Le taux de réussite est cependant modeste : 58,7 %. Plus d'un tiers des élèves (34 %) commettent des erreurs, qu'il conviendrait d'analyser plus en détails.

Dans le cas de Laurent, où il faut procéder à deux déplacements, le pion tombant sur une case spéciale, le taux de réussite est encore plus faible (51 %). Près de 16 % des élèves ne tiennent pas compte de la règle, en apparence (code 7 : Pion de Laurent sur la case 13), mais à vrai dire, ce sont les erreurs d'une autre nature qui sont les plus nombreuses (25,7 %), le pion de Laurent étant placé sur n'importe quelle case autre que la 4 et la 13.

Le faible niveau de réussite et le nombre élevé d'erreurs paraissent donc davantage liés à la situation de jeu simulé (que plusieurs des enseignants ont jugée trop abstraite) qu'à la difficulté de jouer en consultant les règles à la page précédente. Ceci peut être aisément vérifié en classe en faisant jouer les élèves avec de vrais pions⁹.

- **Items 4.5. à 4.7.**

Contrairement à ce que l'on aurait pu penser, les items supposant une anticipation sont mieux réussis : 55 % des élèves indiquent correctement combien de points Julie doit faire pour gagner et seuls 12 % d'entre eux se trompent vraiment (26 % se sont trompés avant et ne peuvent donc répondre correctement). De même, 73,6 % des élèves estiment que Laurent ne peut gagner; seuls 7,5 % se trompent à cette question.

Le codage des justifications de la réponse (item 4.7) n'était pas parfaitement univoque et les résultats ne sont qu'indicatifs : une majorité d'élèves fourniraient une justification vague (*Laurent est trop loin, il ne peut pas dépasser Julie*), sans opérer de véritable raisonnement; d'autres (26,9 %) produiraient des justifications plus élaborées, prenant en compte différents paramètres dans leur raisonnement (le nombre de points maximal, les positions respectives de Laurent et Julie, les cases spéciales ...). Sur ce plan, il conviendra de poursuivre les analyses en examinant qualitativement les réponses fournies par les enfants.

⁸ Pour rappel, les Universités ne disposent pas des carnets leur permettant de mener ce genre d'analyses.

⁹ Pour l'épreuve, on a écarté à cette solution vu les problèmes techniques qu'elle posait.

- **Items 4.8. à 4.11.**

Ce groupe de questions visait à évaluer isolément les différentes composantes intervenant dans la situation de jeu : dénombrement, compréhension des règles sous différentes formes.

Pour ce qui est du dénombrement (item 4.8), le taux de réussite est satisfaisant (81,5 %). On ne manquera cependant pas d'être étonné par le nombre d'erreurs (13,1 %). Même pour les élèves peu familiers des jeux de société, le résultat correct s'obtient en effet par simple addition. Ce nombre élevé d'erreurs laisse perplexe. Tiendrait-il dans la difficulté de compréhension d'une consigne orale ? Il serait sans doute judicieux d'interroger les enfants qui ont commis des erreurs pour en connaître la source.

Près de 80 % des élèves savent lire une règle de jeu simple et en restituer une partie (item 4.9). Plusieurs erreurs consistent à dire qu'il faut retourner à la case départ", parce que c'est souvent comme cela dans les jeux. Certains élèves répondent donc en utilisant leurs connaissances antérieures (ce qu'ils savent déjà des jeux de société) plutôt que sur base d'une lecture attentive des règles qui leur sont présentées. Il est vraisemblable que ces élèves ignorent ce qui reste sans doute souvent implicite
la réponse à des questions de compréhension de lecture se trouve dans le texte.

La question 4.10 est d'un ordre plus complexe puisqu'elle porte sur une règle qu'il faut restituer à partir de la présentation d'une situation de jeu y correspondant (*J'ai dû passer un tour parce que j'étais tombé sur la case ...*). Le taux de réussite est relativement satisfaisant (71 %); 23 % des élèves commettent des erreurs et là aussi il conviendrait d'analyser plus avant leur nature.

Enfin, l'item 4.11 sollicitait la restitution de l'une des règles du jeu (*Pour gagner la partie, il faut ...*). La correction de cet item a malheureusement posé problème, car les critères fournis étaient insuffisamment précis. Les critères précisait que la bonne réponse était "arriver le premier à la case 25"; plusieurs correcteurs auraient cependant été tentés de considérer la réponse "arriver à la case 25" comme correcte également. C'est pourquoi les résultats ne sont guère interprétables. Le taux de bonnes réponses complètes (42,6 %) est à l'évidence surestimé. Très rares sont, en effet, les élèves qui fournissent ce type de réponses. Pour votre classe, il serait intéressant de procéder à un nouveau codage en distinguant les réponses correctes et complètes, les réponses correctes mais incomplètes, et les réponses incorrectes.

5. Analyse des résultats à l'épreuve de lecture de mots isolés

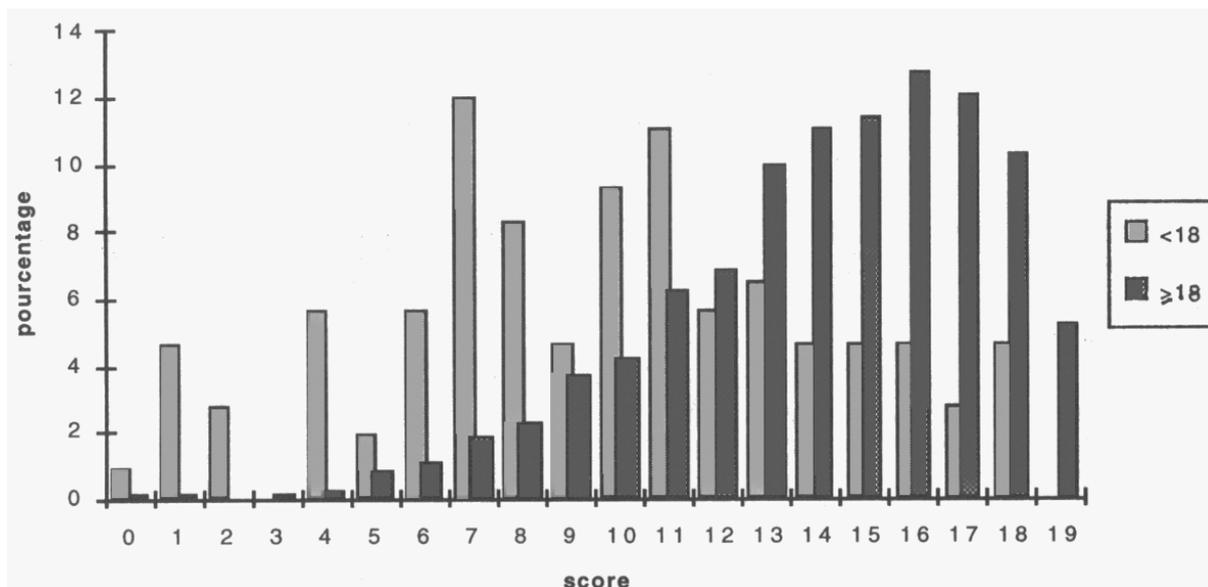
Nombre de mots	Pourcentage de sujets	Pourcentage cumulé
0	0,1	0,1
1	0,1	0,1
2	0,0	0,2
3	0,1	0,3
5	0,1	0,5
6	0,4	0,8
7	0,4	1,3
8	0,1	1,4
9	0,1	1,5
10	0,4	1,9
11	0,2	2,0
12	0,1	2,2
13	0,2	2,4
14	0,3	2,7
15	0,3	2,9
16	0,3	3,2
17	0,9	4,1
18	1,7	5,8
19	8,0	13,8
20	86,2	100,0

Beaucoup d'entre vous, dans les questionnaires, se sont émus de la trop grande facilité de l'épreuve. C'est en effet l'impression que l'on peut éprouver. La facilité ou la difficulté d'une épreuve ne se détermine toutefois qu'en relation avec l'objectif poursuivi. Or, l'objectif poursuivi et annoncé était de diagnostiquer les élèves présentant une lenteur de lecture de mots isolés telle qu'elle les mette en grande difficulté dans l'épreuve. Une majorité d'élèves savent évidemment reconnaître bien plus de 20 mots dans le temps imparti (2 minutes) mais cela, nous ne cherchions pas à l'évaluer.

Au début du second semestre en 3e année, on constate ainsi qu'une majorité d'élèves (86,2 %) lisent les 20 mots en 2 minutes. La proportion d'élèves qui n'y parviennent pas (près de 14 %) est cependant loin d'être négligeable.

On peut s'interroger sur le lien existant entre la réussite à cette épreuve et les items de l'épreuve qui nécessitent une lecture autonome de l'élève (19 items)¹⁰. Le graphique suivant illustre de manière éloquent cette relation. Sur ce graphique, les élèves qui ont obtenus un score à l'épreuve de lecture de mots inférieur à 18 sont représentés en gris clair; ceux qui ont obtenu un score égal ou supérieur à 18 en gris foncé.

¹⁰ Items 1.1.5, 1.1.6, 1.1.7, 1.2.1, 1.2.2, 1.2.3, 1.3, 1.4.1, 1.4.2, 1.4.3, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 4.10 et 4.11.



le décalage entre les deux courbes apparaît clairement. Les lecteurs "lents" à l'épreuve de reconnaissance de mots obtiennent des scores aux items de lecture "autonome" sensiblement inférieurs aux lecteurs "rapides". La majorité des lecteurs "rapides" obtiennent de bons scores aux items de lecture (la courbe prend l'allure d'un J).

Pour tenter de déterminer quel est le profil des élèves lecteurs « lents », les résultats à l'épreuve de reconnaissance de mots isolés ont été mis en relation avec les renseignements recueillis grâce à la fiche Elèves.

Il apparaît ainsi que les garçons sont légèrement sur-représentés dans ce groupe, composé à 57 % de garçons et 43 % de filles; cette différence n'est toutefois pas significative sur le plan statistique

La proportion d'élèves ne parlant pas le français à la maison est plus importante dans le groupe de lecteurs « lents », la répartition entre groupes diffère significativement de ce que laisserait attendre le hasard. Alors que les élèves ne parlant pas le français à la maison représentent 10 % de l'échantillon, ils constituent 34 % des élèves du groupe de lecteurs « lents ».

	Lecteurs « lents »	Lecteurs « rapides »
Ne parlent pas le français à la maison	34 %	9 %
Parlent le français à la maison	66 %	91 %
	100 %	100 %

Lorsque l'on compare les résultats de deux groupes ou la répartition entre deux groupes, il convient de vérifier, grâce à un test statistique, si les différences observées sont significatives. Cette "significativité" s'exprime dans une formule qui exprime la probabilité que cette différence soit due au hasard. On considère ainsi que si l'on a moins d'une chance sur 100 que la différence observée soit due au hasard, celle-ci est significative. Le fait qu'une différence soit statistiquement significative est fortement lié à la taille de l'échantillon. Ainsi, une différence de même ampleur (par exemple un score de 45 et un score de 47) sera très significative si l'échantillon comporte 3000 élèves et non significative si l'échantillon n'en comporte que 100.

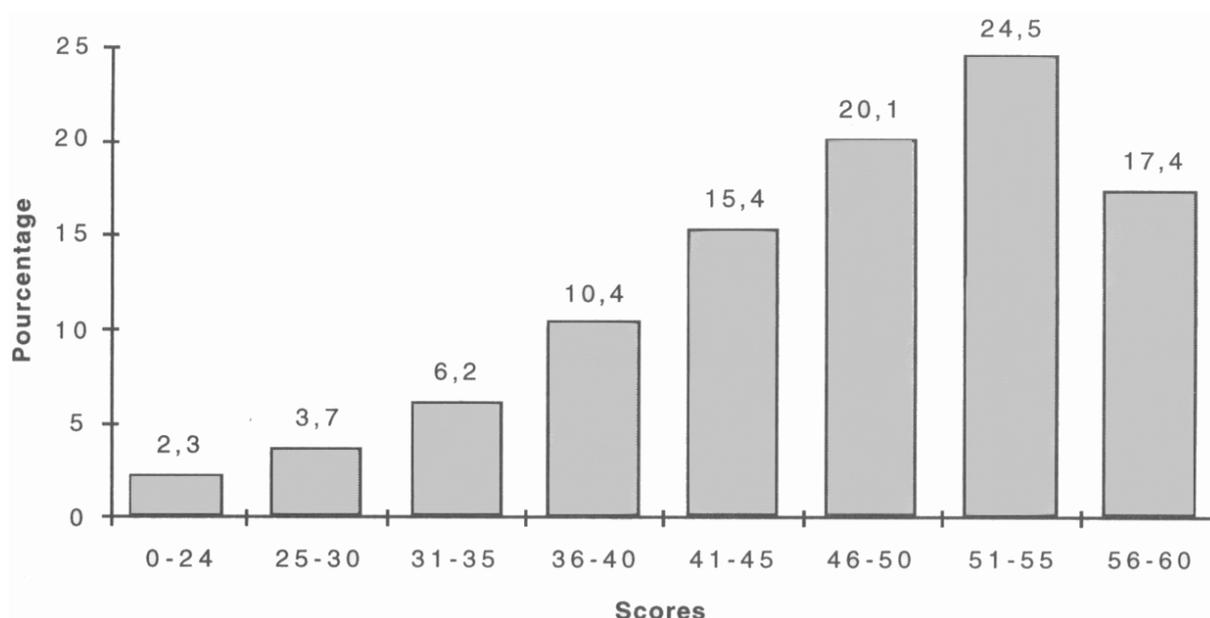
Enfin, la proportion d'élèves en retard scolaire parmi les lecteurs « lents » diffère elle aussi significativement du hasard. Alors que les élèves en retard scolaire représentent 16 % de l'échantillon, ils constituent 35 % des élèves du groupe de lecteurs « lents ».

	Lecteurs « lents »	Lecteurs « rapides »
Elèves en retard scolaire	35 %	14 %
Elèves à l'heure ou avancés	65 %	86 %
	100 %	100 %

TROISIEME PARTIE. ANALYSE DES RESULTATS POUR L'ENSEMBLE DE L'EPREUVE

1. Distribution des résultats à l'ensemble de l'épreuve

Du point de vue diagnostique, l'analyse par question ou groupes de questions est sans doute la plus intéressante. Toutefois, il est également utile de se faire une idée d'ensemble de la maîtrise qu'ont les élèves des compétences attendues à l'issue du premier cycle. Dans ce but, un score global a été calculé pour chaque élève sur l'ensemble du test. Pour calculer un tel score, les codes de toutes les questions ont été transformés en scores dichotomiques: 1 pour une réponse correcte et 0 pour une réponse incorrecte. Le score 1 correspond au code 2 et le score 0 correspond à tous les autres codes (1,7,9 et 0). Les scores ont ensuite été additionnés pour obtenir un score total. Comme le test comprend 60 questions (la question 4.11 ayant été exclue), le score total est au maximum de 60 points.



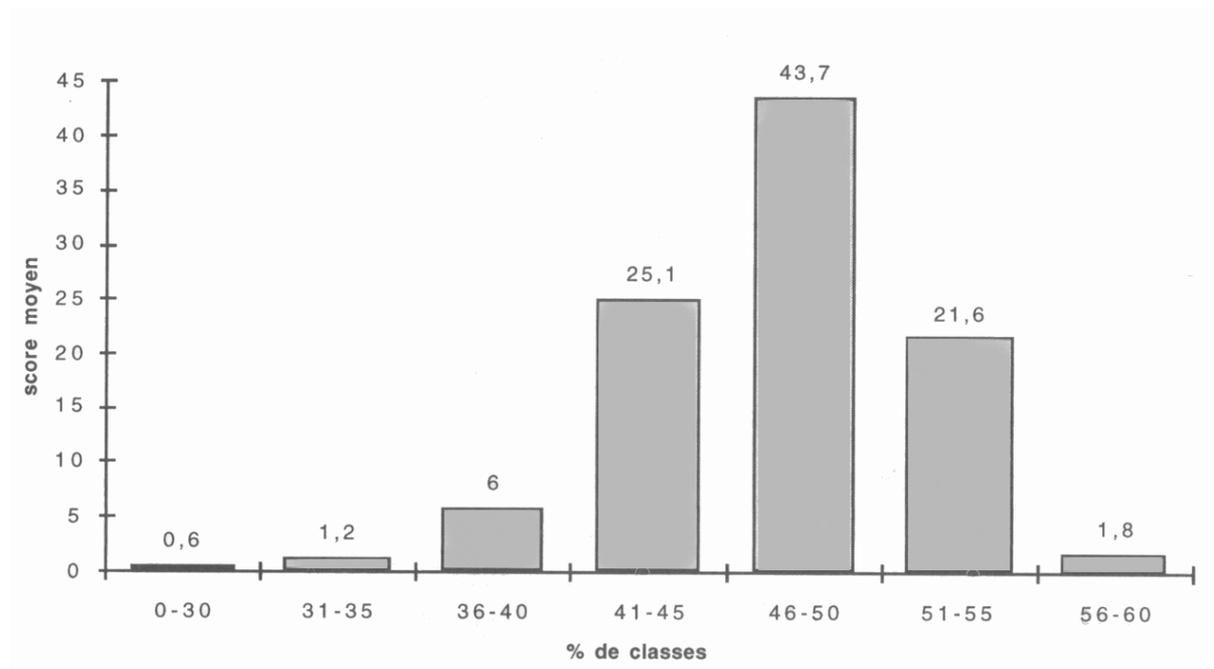
Le graphique ci-dessus présente la distribution de fréquences des scores totaux pour l'ensemble de l'échantillon. Les scores ont été regroupés par intervalle de 5 points. Seul le premier intervalle est plus large puisqu'il s'étend de 0 à 24 points. Cet intervalle ne regroupe en fait que 2,4% des sujets. Il rassemble des scores dont les fréquences sont très faibles (inférieures à 1%).

Le tableau offre une présentation plus détaillée des résultats. La première colonne indique le score total à l'ensemble de l'épreuve. La deuxième colonne donne le pourcentage de sujets de l'échantillon ayant obtenu ce score. La troisième colonne donne le pourcentage cumulé de ces sujets. Par exemple, en face du score 38, le pourcentage cumulé est de 17,8%. Cela signifie que 17,8% des sujets de l'échantillon ont obtenu un score égal ou inférieur à 38. Les pourcentages cumulés permettent de se faire aisément une idée de la position du résultat d'un sujet au sein de la distribution de fréquence des scores.

Score total	Echantillon	Echantillon cumulé
20 et moins	1,1	1,1
21	0,3	1,4
22	0,2	1,6
23	0,3	2,0
24	0,2	2,3
25	0,5	2,8
26	0,4	3,2
27	0,3	3,6
28	0,7	4,3
29	0,8	5,2
30	0,8	6,0
31	0,7	6,7
32	1,2	8,0
33	1,1	9,2
34	1,2	10,5
35	1,5	12,2
36	1,2	13,5
37	1,6	15,3
38	2,3	17,8
39	2,0	20,0
40	2,4	22,6
41	1,7	24,5
42	2,7	27,5
43	3,3	31,1
44	2,7	34,0
45	3,7	38,0
46	3,4	41,8
47	2,8	44,9
48	4,9	48,4
49	4,9	53,8
50	3,9	58,1
51	4,5	63,1
52	4,1	67,5
53	4,2	72,2
54	4,7	77,3
55	4,9	82,6
56	4,3	87,4
57	4,3	92,1
58	3,3	95,7
59	2,7	98,7
60	1,2	100,0

2. Comment situer votre classe par rapport à l'ensemble des classes

Le premier ensemble de résultats fournis concerne les élèves. Tout aussi important pour l'enseignant est de pouvoir situer sa classe. Vous avez, suivant nos indications (cf p. 3), calculé le score moyen de votre classe au test. Le tableau suivant vous permettra de voir où se situe votre classe par rapport aux autres classes.



Score moyen	% de classes dont le score moyen est dans cet intervalle	Votre classe	Votre école
Entre 0 et 30	0,6 %		
Entre 31 et 35	1,2 %		
Entre 36 et 40	6 %		
Entre 41 et 45	25,1 %		
Entre 46 et 50	43,7 %		
Entre 51 et 55	21,6 %		
Entre 56 et 60	1,8 %		

Les classes peuvent obtenir au test un score qui se situe entre 0 et 60; le score moyen est de 46,9 (78 % de réussite). Nous avons découpé l'échelle des scores en intervalles de 5 points et regroupé les scores inférieurs à 30 points. A la lecture du tableau et du graphique, vous pouvez voir combien de classes se situent dans chacun des intervalles de 5 points.

Pour l'ensemble de la population, on voit que les résultats se concentrent dans les trois catégories allant de 41 à 55 (90 % des classes). Très peu de classes obtiennent un score moyen inférieur à 41 ou supérieur à 56.

Pour le tableau suivant, nous avons procédé un peu différemment. A partir de la grille détaillée de résultats (non fournie ici) où l'on peut lire les scores de chacune des 159 classes, ordonnés du plus faible au meilleur, on constitue dix catégories comprenant chacune 10 % des classes. A chacune de ces dix catégories (déciles), correspond une fourchette de scores. A partir du score moyen de votre classe (première colonne), vous pouvez, dans ce tableau, déterminer dans quelle catégorie se situe votre classe.

Exemple

Votre classe a obtenu un score de 47,1 au test. A la lecture du tableau précédent (p. 37) vous pouvez constater que 43,7 % des classes se situent dans cet intervalle de scores. La lecture du tableau ci-dessous vous apprend plus précisément que ce résultat situe votre classe entre la 40^e et la 50^e position.

Si votre classe a obtenu un score moyen...	Elle se situe...
inférieur à 41,8	entre la 1 ^{re} et la 10 ^e position
Entre 41,8 et 42,8	entre la 10 ^e et la 20 ^e position
entre 42,9 et 44,7	entre la 20 ^e et la 30 ^e position
Entre 44,8 et 46,6	entre la 30 ^e et la 40 ^e position
Entre 46,7 et 47,9	entre la 40 ^e et la 50 ^e position
entre 48 et 48,9	entre la 50 ^e et la 60 ^e position
Entre 49 et 49,9	entre la 60 ^e et la 70 ^e position
Entre 50 et 51	entre la 70 ^e et la 80 ^e position
supérieur à 51	entre la 90 ^e et la 100 ^e position

3. Comment situer votre classe par rapport à des classes qui lui « ressemblent » ?

Le score moyen de réussite des élèves, on l'a dit, est de 46,9/60. Votre classe obtient, selon les cas, un score qui est proche de cette moyenne, ou qui lui est inférieur ou supérieur, de peu ou de beaucoup. C'est là une première façon de se situer.

Vous savez cependant, même si cela n'obéit pas à un déterminisme strict, qu'il est plus facile ou plus courant d'obtenir de bons résultats quand la classe compte davantage d'élèves de milieu favorisé, parlant le français à la maison ... Pour chacune des classes, nous avons recueilli des informations via le questionnaire "Enseignant" sur la proportion d'élèves qui

- ont redoublé au moins un an au cours de leur scolarité antérieure; ne parlent pas le français à la maison;
- vivent dans des conditions de vie précaire (parents au chômage, familles très nombreuses, problèmes sociaux, ...);
- sont issus de milieux socio-culturellement favorisés (enseignants, cadres, professions libérales, ...).

Nous avons demandé, en outre, si l'école fait ou non partie d'un projet ZEP, ZAP ou SOUTIEN.

A partir de ces renseignements sur la classe, nous sommes en mesure de fournir des résultats par "types" de classes (classes ZEP ou pas ZEP, classes comptant plus de 20 % d'élèves ne parlant pas le français à la maison, ...). Ces résultats moyens par "sous-groupes" vous permettront de vous situer plus précisément par rapport à des classes qui "ressemblent" à la vôtre. Pour rappel, le score moyen général au test est de 46,9/60.

	Score total	% de ce type de
Classes ZEP, ZAP ou soutien	44,4	12%
Autres classes	47,5	88%
Classes comprenant 25% d'élèves ou plus ayant redoublé au moins une fois	45,2	28%
Autres classes	47,7	72%
Classes comprenant 50% d'élèves ou plus issus de milieu favorisé	49,9	24%
Autres classes	46,1	76%
Classes comprenant 20% d'élèves ou plus ne parlant pas le français à la maison	43,6	15%
Autres classes	47,7	85%
Classes comprenant 25% d'élèves ou issus d'une famille en situation précaire	43,4	33%
Autres classes	48,7	66%

Toutes les différences entre les catégories de classes ainsi constituées sont statistiquement très significatives.

Ces résultats, certes, ne sont qu'indicatifs. A l'intérieur de chacune des catégories constituées (par exemple, plus ou moins 20 % d'élèves ne parlant pas le français), il existe des situations bien différentes (la classe comptant 100 % d'élèves non francophones est plus défavorisée que celle n'en comptant que 21 %). Nous pensons néanmoins qu'ils vous permettront de situer un peu plus précisément les résultats de votre classe. Ainsi, pour prendre des exemples un peu extrêmes, mais éclairants, si une classe ZEP qui, par définition, regroupe beaucoup d'élèves étrangers, défavorisés, ... obtient un résultat se rapprochant de celui des écoles non ZEP, on peut dire qu'il s'agit là, relativement, d'un très bon résultat. A l'inverse, si une classe comptant plus de 50 % d'élèves favorisés obtient un résultat se rapprochant de celui des écoles moins favorisées, on peut considérer qu'il s'agit là, relativement, d'un résultat un peu "inquiétant". C'est à vous qu'il appartient, à la lumière de ce que vous savez de votre classe, de son niveau, de la façon dont l'épreuve a été administrée, d'interpréter vos résultats ainsi "relativisés".

4. Résultats en fonction des caractéristiques des élèves

Les résultats qui viennent d'être présentés peuvent être complétés par une analyse en fonction des caractéristiques des élèves d'après les renseignements recueillis grâce à la fiche "Elève".

	Score total moyen de ce type d'élèves	% de ce type d'élèves
Sexe		
Garçons	46,9	50,1 %
Filles	46,8	49,9 %
Français parlé à la maison		
Non	47,4	10 %
Oui	41,5	90 %
Retard scolaire		
Non	47,9	84 %
Oui	41,1	16 %
A redoublé la		
1 ^{ère} année	39,5	7 %
2 ^e année	39,2	6 %
3 ^e année	44,8	4 %

On n'observe pas de différence significative de réussite au test entre les filles et les garçons, ce qui peut surprendre : les filles obtiennent généralement de meilleurs scores, particulièrement à ce niveau de la scolarité.

La différence entre élèves parlant habituellement le français à la maison ou non est en revanche très significative, ainsi que celle entre les élèves à l'heure (et avancés) et les élèves en retard scolaire. La différence est particulièrement nette pour les élèves qui ont redoublé la 1^{re} ou la 2^e année. Toutefois, même les élèves qui redoublent l'année en cours (3^e) restent plus faibles que la moyenne des élèves à l'heure, alors que l'année est peu avancée. Ces chiffres conduisent, une nouvelle fois, à s'interroger sur l'efficacité du redoublement¹².

¹² Chaque enseignant recevra, d'ici quelques semaines, un document intitulé *De deux ans et demi à 18 ans, Réussir l'école ...*, Ministère de l'Éducation, de la Recherche et de la Formation, où la question de l'efficacité du redoublement est discutée.

5. Niveaux de compétence des élèves

Les 60 items du test¹³ ont été soumis à une analyse statistique (analyse de RASCH)¹⁴ qui permet

- de hiérarchiser les items sur une échelle, du plus simple au plus difficile; -d'exprimer sur la même échelle le niveau de compétences des élèves.

Parmi les items ainsi hiérarchisés, on peut, en regroupant des "blocs" d'items proches, déterminer des "niveaux de compétence" et voir combien d'élèves se situent à chacun des niveaux.

Sur l'échelle, la moyenne de la difficulté des items est par convention fixée à 0 et l'écart type à 1. Les items sont rangés du plus facile (indice = - 2.588) au plus difficile (2.986). L'indice de difficulté exprime aussi la distance entre les items : deux items consécutifs peuvent être pratiquement de même difficulté, ou l'un beaucoup plus difficile que l'autre (c'est par exemple le cas du dernier et de l'avant-dernier item).

Les résultats doivent se comprendre comme suit : un élève qui a une compétence correspondant à la difficulté de tel item (par exemple l'item 2.10.1 : lecture de l'heure - 8 h ou 20 h) a une probabilité de 80 % de réussir cet item et tous les items plus faciles de l'échelle (de 1.2.1 à 3.17), et une probabilité inférieure à 80 % de réussir les items plus difficiles sur l'échelle (de 1.3 à 4.7). Le seuil a été fixé à 80 % parce que c'est habituellement à ce niveau que l'on fixe le seuil de maîtrise.

La hiérarchie des items obtenue est la suivante :

Des items les plus faciles...

		Indice de
Niveau 1 : Élémentaire		
1.2.1.	Prix: 70 F	- 2.588
1.2.2.	Prix: 200 F	- 2.245
1.2.3.	Prix: 560 F	- 2.214
3.9.	Calcul: 60	- 1.917
3.7.	Calcul: 35	- 1.885
3.14.	Calcul: 30	- 1.847
4.2.2.	Pion Laurent, case 10	- 1.509
4.2.1.	Pion Julie, case 9	- 1.473
3.6.	Calcul: 39	- 1.455
3.11.	Calcul: 40	- 1.415
3.8.	Calcul: 31	- 1.275
3.10.	Calcul: 30	- 1.260
3.12.	Calcul: 35	- 1.105
Niveau 2 : Intermédiaire		
2.6.	Heure d'entrée: 9 heures	- 1.016
3.13.	Calcul : 12	- 0.960
3.1.	Argent de Julie: 48 F	- 0.724
1.4.2.	Adresse Laurent	- 0.677
3.17.	Opération: 20 - 5	- 0.597
2.10.1.	Lecture d'heure : 8 h ou 20 h	- 0.590
1.3.	Ouverture piscine: 9 h	- 0.583
4.1.	Laurent va commencer	- 0.500
3.16.	Opération: 20 x 5	- 0.436

¹³ Le dernier item a été exclu de l'analyse.

¹⁴ Pour plus de détails sur cette analyse, voir le glossaire en fin de document.

2.9.1.	Relier Lavabo - 20 cm	- 0.405
3.3.	Colorier les pièces : 1 p. de 20 F et 2 p de 5 F	- 0.296
4.8.	Case 10	- 0.238
1.4.3.	Adresse n° 3 : Piscine	- 0.210
1.4.1	Adresse n° 1 : Julie	- 0.179
Niveau 3 : Compétent		
1.1.1.	Copie du prénom, du nom et de l'adresse de Laurent.	- 0.132
1.1.5.	L'élève a compris la tâche.	- 0.124
3.18.	Opération : 20 + 5	- 0.113
4.9.	Retourner à la case 4	- 0.094
3.2.	Prix lace de Julie: 30 F	- 0.055
2.4.	Mesure de la longueur: 25 m	- 0.041
2.10.3.	Lecture d'heure : 3 h 30 ou 15 h 30	0.056
1.1.2.	Emploi correct des majuscules	0.121
2.9.2.	Relier Bassin: 2m50	0.229
2.10.2.	Lecture d'heure : 10 h 15 ou 22 h 15	0.303
2.9.3.	Relier baignoire: 60 cm	0.324
4.6.	Réponse correcte : non	0.342
4.10.	Case 8 ou en prison	0.507
2.5.	Nombre de mètres: 50 m	0.592
3.4.	Reste de Julie : 18 F	0.594
1.1.3.	Orthographe : copie correcte et complète	0.657
2.3.	Phrases correctes : les deux phrases correctes (3e et 4e) sont entourées	0.776
2.7.	Heure de sortie : 10 h 30	0.860
3.15.	Calcul :15	0.870
1.1.4.	contenu de la carte : Présence d'une formule d'adresse à Laurent et d'une signature	0.992
2.1.	Profondeur de l'eau : 1 mètre	1.043
2.2.	Nombre de longueurs : 4	1.126
4.3.	Le ion de Julie est sur la case 21	1.216
1.1.6.	Nom de la rue où se trouve la piscine correct : "Grand-rue"	1.275
2.11.2.	Durée du film: 2 heures 30	1.347
4.5.	Réponse correcte : 4	1.392
2.11.1.	Durée du film: 2 heures 15	1.537
2.8.	Durée : 1 heure 30	1.615
3.5.	La monnaie de retour: 55 F	1.615
4.4.	Le ion de Laurent est sur la case 4	1.628
1.1.7.	Heure d'ouverture le samedi correcte : 9 heures, 9-13 heures ou de 9 à 13 heures	1.635
2.11.3.	Durée du film: 30 minutes	2.454
4.7.	Justification précise du NON	2.986

...aux items les plus difficiles

D'un point de vue technique, cette analyse est très solide : la hiérarchie des items qui en résulte est en effet **indépendante** de la population d'élèves qui a passé le test. Le même test, administré à des élèves plus jeunes ou plus âgés donnerait bien sûr des pourcentages de réussite plus ou moins élevés, mais la hiérarchie des items devrait rester la même. Quand sur une telle échelle, un item apparaît plus facile qu'un autre, ce n'est dû ni au hasard, ni à la force des élèves; il y a à cela de "bonnes" raisons sur lesquelles il convient de s'interroger.

Au départ de cette échelle robuste, il est donc légitime de tenter de regrouper des items en "blocs" ou "niveaux" de compétence.

Le nombre de niveaux et l'endroit précis où l'on place les barres entre niveaux dépendent certes de la personne ou du groupe de personnes qui effectuent ce travail. **Le groupe de concepteurs de l'épreuve d'évaluation externe, composé de chercheurs et d'inspecteurs, après examen attentif de la hiérarchie des items, s'est mis d'accord pour déterminer quatre niveaux de compétence dans l'épreuve.** Ce choix d'un découpage en quatre niveaux résulte donc du consensus d'un groupe; il a nécessairement un caractère relatif et en partie arbitraire.

Les quatre niveaux de compétences se définissent comme suit¹⁵

- **Niveau 1 : Élémentaire**

Les élèves qui se situent à ce niveau ne maîtrisent vraiment (à 80 %) que les compétences isolées les plus simples (celles dont plusieurs enseignants dans le questionnaire ont dit qu'elles étaient **trop** simples) : les calculs les plus simples de la situation 3, relier les prix à ce à quoi ils correspondent dans la situation 1. Selon les membres du groupe d'inspecteurs et de chercheurs qui a conçu l'épreuve, tous les élèves de fin de 2e devraient maîtriser ces compétences.

- **Niveau 2 : Intermédiaire.**

Les élèves qui se situent à ce niveau maîtrisent les compétences du niveau 1 et des compétences déjà plus complexes. Parmi celles-ci, on retrouve des compétences isolées

- certains calculs ou opérations (20 - 5) et (20 x 5), l'identification des adresses, l'heure d'ouverture de la piscine (QCM), la lecture des heures les plus simples (heure juste);
- certaines questions peu complexes posées en situation : combien Julie a-t-elle d'argent ?, colorier les pièces, qui va commencer le jeu ? ...

Toutes ces questions sont bien ciblées et demandent un traitement de l'information simple (une seule opération à effectuer, pas de compréhension complexe). Selon les membres du groupe

concepteur de l'épreuve, ces compétences sont toutes à la portée d'élèves de fin de 2e et doivent donc être maîtrisées. Aucune notion qui n'aurait été abordée dans le programme ne figure parmi ce groupe.

- **Niveau 3 : Compétent.**

Les élèves qui se situent à ce niveau sont capables de réussir à 80 % des questions qui correspondent

- à des compétences testées isolément telles que : la lecture de l'heure (demi-heure et quart d'heure), des opérations (20 + 5), la lecture isolée des règles du jeu de la situation 4, l'évaluation de la profondeur de la baignoire et du bassin et le calcul 50 - 35;
- à des compétences testées en situation, mais restant d'un ordre de complexité accessible copie de la carte (sans rien oublier, majuscules et orthographe), compréhension globale de la tâche (Sit. 1), heure de sortie de la piscine, prix de la glace de Julie, reste de Julie (18 F), Laurent ne peut pas gagner, longueur de la piscine, ...

Il apparaît souhaitable aux concepteurs de l'épreuve que tous les élèves maîtrisent ce niveau de compétences.

- **Niveau 4 : Avancé.**

Les compétences figurant dans ce groupe d'items sont en effet des compétences complexes dont on peut penser

- soit qu'elles excèdent quelque peu, sur le plan des notions, ce qui est requis en fin de 2^e on retrouve ici tous les items relatifs à la durée, poser et résoudre la soustraction 100 - 45 = 55, l'estimation des profondeurs, et les questions relatives à la longueur;
- soit qu'elles sont évaluées dans un contexte qui n'est pas totalement adéquat (le contexte de jeu simulé de la situation 4 a posé certains problèmes) ou est trop chargé sur le plan cognitif (l'élève a trop de choses à faire à la fois).

Par ailleurs, certains des items de ce groupe ne sont pas pleinement satisfaisants sur le plan technique : ils ont posé des problèmes de compréhension ou de correction.

¹⁵ Aux pages 47 et 48, ces niveaux de compétence seront confrontés avec les réponses que les enseignants ont fournies dans le questionnaire à propos des items du test.

Pour ces différentes raisons, il apparaît légitime aux concepteurs de l'épreuve, de considérer que les élèves de fin de 2e ne doivent pas réussir (à 80 %) ce type d'items. Maîtriser ces compétences telles qu'elles ont été évaluées ici représente en quelque sorte "un plus", mais ne fait pas partie du bagage que le groupe juge indispensable à ce niveau.

Pour chacun des paliers ainsi définis, on peut déterminer combien d'élèves se situent à ce niveau. Rappelons que ces niveaux n'ont pas de valeur absolue; ils sont déterminés par rapport à l'épreuve et aux compétences qui y sont évaluées, sous la forme où elles y sont évaluées.

	% d'élèves qui ont une probabilité de 80 % de réussir les items de ce groupe
Niveau 1 : Elémentaire	11 %
Niveau 2 : Intermédiaire	20 %
Niveau 3 : Compétent	37 %
Niveau 4 : Avancé	32 %

Selon le groupe concepteur de l'épreuve, le niveau 3 peut être considéré comme celui que devrait atteindre les élèves à l'issue du 1^{er} cycle. On voit que 69 % des élèves ont atteint ce niveau de compétences. Parmi ces 69 %, 32 % atteignent, en outre, un niveau de compétences approfondies ou complexes. Vingt pour cent des élèves (ceux du niveau 2) sont en outre sur le chemin de cette maîtrise. Leur probabilité de réussir les items du niveau 3 est en effet inférieure à 80 %, sans être négligeable. Pour les élèves se situant au niveau 1 (11 % des élèves), le chemin vers la maîtrise reste encore long à parcourir; ces élèves ne maîtrisent en effet que le niveau des compétences les plus simples évaluées isolément et sont en difficulté dès que les compétences se complexifient du sont insérées dans un contexte.

Même si le cas de ces 11 % d'élèves est inquiétant, il n'empêche que le bilan, pour les compétences évaluées, s'avère globalement satisfaisant. Pour tenter de cerner le profil des élèves se situant pour cette épreuve à chacun des niveaux, on a croisé l'appartenance au niveau avec les variables "parle le français à la maison " d'une part, "retard scolaire" d'autre part.

	Ne parlent pas le français à la maison	Parlent le français à la maison
Niveau1:Elémentaire	25 %	9 %
Niveau2:Intermédiaire	30 %	20 %
Niveau 3 : Compétent	31 %	37 %
Niveau 4 : Avancé	14 %	34 %
	100 %	100 %

Le tableau fait clairement apparaître la sur-représentation, statistiquement très significative, des élèves ne parlant pas le français à la maison dans les niveaux 1 et 2 (55 % des élèves) contre 29 % dans le groupe des enfants parlant le français à la maison.

	Elèves en retard scolaire	Elèves à l'heure ou avancés
Niveau1: Elémentaire	22 %	9 %
Niveau 2:Intermédiaire	33 %	18 %
Niveau 3 : Compétent	35 %	37 %
Niveau 4 : Avancé	10 %	36 %
	100 %	100 %

La différence de répartition entre niveaux est manifeste, et plus accusée encore que pour la langue parlée à la maison. Cinquante-cinq pour cent des élèves en retard scolaire ne dépassent pas le niveau intermédiaire, contre 27 % des élèves à l'heure. Seuls 45 % des élèves retardés atteignent le se niveau "compétent", quand 73 % des élèves à l'heure y parviennent.

6. Evaluation par les titulaires et les inspecteurs du « niveau » des compétences mises en œuvre dans l'épreuve

Aussi bien les inspecteurs que les titulaires ont été interrogés à propos des compétences évaluées dans l'épreuve. Pour chacune d'entre elles, nous leur avons demandé d'estimer s'il 'agissait:

- d'une compétence "à maîtriser" = d'une compétence que tous les enfants en entrant en 3^e année devraient maîtriser à un degré déterminé.
- d'une compétence "en construction" = d'une compétence que les élèves doivent progressivement construire, mais qui n'est pas encore maîtrisée en fin de 1^{er} cycle.

Les réponses à ces questions prennent une dimension supplémentaire du fait de l'analyse effectuée dans les pages précédentes.

D'une manière générale, pour l'ensemble de l'épreuve, les différences d'appréciation entre titulaires et inspecteurs sont faibles.

	Compétences à maîtriser	Compétences en construction	Omissions
Moyenne des titulaires	54,3 %	43,7 %	2 %
Moyenne des inspecteurs	52,2 %	43,7 %	4 %

Cette moyenne est d'un intérêt relativement limité, car les jugements sont très variables d'un item à l'autre. Elle permet néanmoins de voir que globalement, inspecteurs et titulaires sont "sur la même longueur d'ondes".

On ne constate pas de divergences importantes entre les inspecteurs et les titulaires dans le classement des items; les différences observées concernent plutôt le degré d'exigence. D'une manière générale, les titulaires se montrent plus exigeants (estiment plus souvent qu'il faut maîtriser) pour les compétences peu complexes et évaluées isolément (lire l'heure, les calculs, les opérations), compétences dont on est assuré qu'elles sont régulièrement exercées en classe sous cette forme, tandis que les inspecteurs auraient tendance à l'être davantage pour certaines des compétences du niveau plus avancé (durées, estimation de la profondeur; certaines questions de la situation 4)¹⁶.

¹⁶ Précisons que les inspecteurs, à la différence des titulaires, n'ont pas assisté à la passation du test.

Si l'on confronte, à titre indicatif, les résultats de ces deux classements avec l'analyse, de RASCH, on constate aussi d'importantes similitudes.

	moyen des titulaires estimant que des compétences du niveau doivent être maîtrisées	% d'inspecteurs
Niveau 1	80,2 %	81,4 %
Niveau 2	61,8 %	54,3 %
Niveau 3	55,9 %	55,0 %
Niveau 4	27,3 %	34,8 %

Cette comparaison ne peut toutefois être qu'indicative car l'analyse de RASCH produit une hiérarchie des items, tandis que les inspecteurs et les titulaires ont été interrogés sur des compétences qui regroupaient dans certains cas plusieurs items. Une évaluation par item produirait des classements plus nuancés. De surcroît l'évaluation demandée n'insistait pas suffisamment sur le caractère **contingent** des compétences (ces compétences telles qu'on les avait évaluées dans l'épreuve et non dans l'absolu), ce qui a pu susciter des malentendus. Les quelques divergences d'importance relevées (par exemple items 4.3 et 4.4 - placer les pions de Julie et Laurent après le 2^e coup, items mal réussis et dont une large majorité estiment qu'ils devraient être maîtrisés) tiennent sans doute à cette ambiguïté : une chose est de "maîtriser" ces compétences en situation de jeu réelle, une autre est de le faire dans un contexte simulé, avec les contraintes matérielles que l'on sait.

Malgré ces quelques réserves, on peut constater qu'une majorité de plus de 50% des titulaires et des inspecteurs estiment que les compétences groupées dans les niveaux 1, 2 et 3 devraient être maîtrisées par **tous** les élèves. En revanche, une large majorité considère que les compétences groupées dans le niveau 4 sont « en construction » à la fin du 1^{er} cycle.

Il est intéressant de constater que de manière indépendante, le groupe de concepteurs d'une part, la majorité des titulaires et des inspecteurs d'autre part, placent la "barre" de ce qui est « requis » et de ce qui dépasse le niveau requis, entre le niveau 3 et le niveau 4.

7. Les résultats au test correspondent-ils au "niveau" de la classe apprécié par le titulaire ?

Dans le questionnaire au titulaire, il a été demandé à celui-ci d'apprécier le niveau moyen de sa classe de 3e actuelle sur une échelle à cinq degrés (de « très faible » à « excellent »). Les résultats se répartissent comme suit

Très faible	0,6 %
Faible	11,0 %
Moyen.....	37,0 %
Bon	49,4 %
Excellent	2,0 %
	<hr/>
	100 %

Même si l'estimation demandée ne représente qu'une indication, il est intéressant de la confronter avec les résultats observés au test.

Une *corrélation* a été calculée entre le jugement du titulaire et le score moyen de sa classe. Il ressort de cette analyse que les deux "évaluations" sont concordantes dans une majorité de cas. On constate néanmoins un nombre non négligeable de discordances. Soit le titulaire juge sa classe "faible" ou "moyenne", alors que celle-ci obtient de bons résultats relatifs au test. On parlera dans ce cas de "sous-estimation" (par rapport à ce test, bien entendu). Soit le titulaire juge sa classe "bonne", voire "excellente" et les résultats au test sont relativement moyens ou mauvais. On considérera dans ce second cas que le titulaire "surestime" sa classe. Ces phénomènes (surestimation, sous-estimation) n'ont bien entendu de valeur que par rapport à l'épreuve qui a été passée et il convient de les considérer avec prudence car les sources de ces différences d'estimation sont diverses et multiples. Pour les classes qui se trouveraient dans le cash, cette différence d'appréciation mérite toutefois réflexion.

Mais examinons plus en détail ces résultats

- Parmi les 4 classes obtenant un résultat inférieur à 36/60 (60 % de réussite au test), 2 sont jugées "faibles", 2 "moyennes". Ces deux dernières sont sans doute surestimées par leur titulaire. Les résultats au test sont en dessous de ce que laisse attendre le jugement du maître.

- Parmi les 22 classes obtenant un score entre 36/60 et 42/60 (60 et 70 % de réussite)

- une est jugée "très faible",
- 4 sont jugées "faibles",
- 14 sont jugées "moyennes",
- 3 sont jugées "bonnes".

Il semble légitime de penser que ces trois dernières classes sont "surestimées". Leurs élèves ont tendance à obtenir au test des résultats décevants compte tenu de l'estimation du titulaire.

- Parmi les 55 classes obtenant un score se situant entre 42/60 et 48/60 (70 et 80 % de réussite)

- 8 sont jugées "faibles",
- 24 sont jugées "moyennes",
- 23 sont jugées "bonnes".

Les 8 classes jugées faibles, au vu des résultats du test, semblent "sous-estimées" par le titulaire. Les élèves obtiennent au test des résultats meilleurs que ne le laisserait attendre le jugement du titulaire.

- Parmi les 68 classes obtenant un score supérieur à 48/60 (plus de 80 % de réussite)

¹⁷ Vous pouvez, pour votre classe, opérer la même analyse que ci-dessous, en confrontant le score moyen et votre appréciation du niveau de la classe.

- 2 classes sont jugées "faibles" et sont donc fortement sous-estimées;
- 18 classes sont jugées "moyennes",
- 45 classes sont jugées "bonnes",
- 3 classes sont jugées "excellentes". L'une de ces classes paraît "surestimée" (score de 50/60).

Au total, à côté d'un large ensemble de résultats "concordants", on décèle donc cinq cas assez manifestes de "surestimation" et dix cas de "sous-estimation". Dans environ 10 % des classes, le titulaire se trouve donc confronté, via l'évaluation externe, à une performance de sa classe très "discordante" par rapport à la représentation qu'il se fait du niveau de celle-ci. Ces "distorsions" peuvent tenir à différents facteurs:

- écart important entre la pédagogie pratiquée en classe, en 3^e et au cours du 1^{er} degré, et la nature du test;
- "isolement" du titulaire, qui ne dispose pas de critères de référence pour situer sa classe;
- dépassement des exigences du programme et tendance à sous-estimer sa classe;
- phénomènes particuliers lors de la passation ou de la correction du test (titulaire absent, sévérité plus forte que la moyenne, davantage d'aide que prévu aux élèves, ou le contraire ...);
- ...

Il revient aux titulaires, qui sont les seuls à disposer du faisceau d'informations nécessaires,, pour interpréter ces phénomènes, de pousser plus loin l'analyse pour leur classe. Une fois cette analyse effectuée, la représentation du "niveau" de leur classe devra peut-être se réajuster quelque peu. Ce réajustement quelquefois nécessaire correspond à l'une des missions de l'évaluation externe.

QUATRIÈME PARTIE. RÉSULTATS DU QUESTIONNAIRE AU TITULAIRE

Pour rappel, ce questionnaire concernait trois aspects:

- les caractéristiques de la classe et de l'école,
- l'appréciation du niveau de difficulté de l'épreuve,
- le classement des items de l'épreuve en compétences à maîtriser en fin de 1er cycle et compétences en construction. Ce troisième aspect a déjà été développé.

1. Caractéristiques de l'école et de la classe

La taille moyenne de l'école est de 203 élèves (min = 13, max = 710) : 24 % des écoles comptent moins de 100 élèves, 31 % des écoles ont entre 100 et 200 élèves, 24 % des écoles comptent 200 à 300 élèves, 27 % des écoles plus de 300 élèves. On constate une corrélation positive, mais assez faible, entre la taille de l'école et le score au test : la réussite tend à être meilleure dans les plus grosses écoles. Ceci est sans doute dû au fait que les grosses écoles recrutent, en général, une population socialement plus favorisée.

Les classes où les élèves ont été testés comptent en moyenne 17,2 élèves de 3e (écart type 7,3).

En moyenne, les classes sont cependant un peu plus nombreuses (18,9 élèves), du fait des classes multi-âges.

Taille des classes	% de classes
15 élèves ou moins	25,5 %
entre 16 et 20 élèves	35,3 %
entre 21 et 25 élèves	31,4 %
plus de 25 élèves	7,8 %
	100 %

La corrélation entre la taille de la classe et la réussite au test n'est pas significative.

Les classes de 3e comptent en moyenne

- **17 % de redoublants** (minimum 0 %, maximum 67 %)

36 % des classes comptent moins de 10 % de redoublants,

17 % des classes en comptent plus de 30 %.

La corrélation négative (-0.27) entre la proportion de redoublants dans la classe et la réussite au test est très significative (0.0001);

- 11 % d'élèves ne parlant pas le français à la maison.

Soixante pour cent des classes ne comptent aucun enfant de ce type. A l'opposé, 6 % des classes en comptent plus d'un tiers. La corrélation négative (-0.40) entre la proportion d'élèves ne parlant pas le français à la maison et le score au test est très significative (0.0001);

- 32 % d'élèves de milieu favorisé (minimum 0 %, maximum 100 %).

La corrélation (0.42) entre la proportion d'élèves favorisés dans la classe et le score au test est très significative (0.0001).

- 22 % d'élèves en situation de précarité (minimum 0 %, maximum 100 %).

La corrélation négative (-0.52) entre la proportion d'élèves en situation de précarité et le score au test est très significative (0.0001).

Douze pour cent des écoles sont inscrites dans un projet ZEP, ZAP ou SOUTIEN.

Au cours du 1^{er} cycle, en général

- 7 % des élèves étaient dans le cycle 5-8;
- 55,5 % des élèves ont eu un titulaire en 1^{re} année et un titulaire différent en 2^e année;
- 21 % des élèves ont eu un titulaire qui est "monté" avec sa classe;
- 18 % des élèves ont eu un même titulaire dans une classe multi - âges.

Près d'un élève sur deux a été suivi par un même titulaire au cours du 1^{er} cycle, sous des modalités diverses.

Un peu plus d'un quart des classes sont organisées en classes multi-âges. Toutes les formes de groupement sont observées, le mode le plus courant est une classe regroupant 3^e et 4^e (12,2 % des classes) suivi par le groupement 1-2-3 (6,1 % des classes). La réussite au test ne diffère pas significativement selon qu'il s'agit d'une classe comprenant uniquement des élèves de 3^e ou une classe multi-âges.

2. Appréciation de la difficulté de l'épreuve par situation

SITUATION	1	2	3	4
<input type="checkbox"/> beaucoup trop facile pour des élèves de début de 3e année	-	-	3	0,6
<input type="checkbox"/> un peu trop facile pour des élèves de début de 3e année	5,5	7,3	25	6,1
<input type="checkbox"/> bien adaptée pour des élèves de début de 3e année	69,5	62,8	64,6	43,9
<input type="checkbox"/> un peu trop difficile pour des élèves de début de 3e année	23,2	26,2	4,9	42,7
<input type="checkbox"/> beaucoup trop difficile pour des élèves de début de 3e année	0,6	1,8	0,6	5,6
Omissions	1,2	1,8	1,8	1,2
Total	100 %	100 %	100 %	100 %

Pour les quatre situations, c'est la réponse "bien adaptée pour des élèves de début de 3e" qui recueille le plus de faveurs. Les nuances de jugement selon les situations sont cependant importantes. Les situations 1 et 2 sont jugées bien adaptées par plus de 60° % des titulaires et un peu trop difficiles par environ un quart d'entre eux. Très rares sont ceux qui les trouvent trop faciles.

La situation 3 est également jugée adaptée par 65 % des instituteurs/trices, mais un quart d'entre eux la jugent trop facile, et une minorité trop difficile.

La situation 4 présente quant à elle un profil singulier; les enseignants sont moins nombreux (44 %) à la juger adaptée, et sont presque aussi nombreux (43 %) à la juger trop difficile.

Ces appréciations de la difficulté des situations concordent-elles avec la réussite observée chez les élèves ?

	% de réussite observé
Situation 1	80 %
Situation 2	70 %
Situation 3	86 %
Situation 4	73 %

La situation 3, comme le prévoyaient les enseignants, se révèle la plus facile. En revanche la situation 4 apparaît moins "difficile" qu'escompté; les élèves y obtiennent en tout cas un meilleur résultat que dans la situation 2, dont la difficulté relative apparaît un peu sous-estimée. Rappelons que cette situation fait intervenir le concept de durée qui semble poser problème à ce niveau de la scolarité. Elle s'inscrit en revanche dans un contexte de questionnement peu difficile à gérer, ce qui peut lui donner une apparente simplicité.

BILAN ET PERSPECTIVES

L'analyse des résultats globaux effectuée (analyse de RASCH) est intéressante à plus d'un titre. D'un point de vue technique, cette analyse est très solide : la hiérarchie des items qui en résulte est en effet **indépendante** de la population d'élèves qui a passé le test.

Au départ de cette hiérarchie robuste, le groupe de concepteurs de l'épreuve s'est attaché à définir des "niveaux" de compétence. Ces "niveaux", répétons-le, se déterminent en relation avec les compétences évaluées dans l'épreuve et ont donc un caractère limité et par définition arbitraire. Toute tentative de définition d'un "niveau" requis pour tel degré de la scolarité souffre cependant du même arbitraire.

Où placer la barre ? Quel "niveau" les élèves doivent-ils atteindre à 8,12 ou 14 ans ? La réflexion menée à propos de l'épreuve d'évaluation externe est du même ordre que celle qui doit guider l'élaboration de "socles de compétence".

Dans cette perspective, on pourrait concevoir que des épreuves du type "Evaluation externe", couvrant un large ensemble de compétences, élaborées collectivement par des chercheurs en éducation et des inspecteurs, concourent à préciser progressivement des "socles" ou des "niveaux" de compétences d'une manière plus objective et plus fiable.

On le voit, la perspective adoptée cette année pour l'évaluation externe en 3e année représente un "saut qualitatif". Par rapport à l'évaluation réalisée en 5e, qui restait normative (quel pourcentage de réussite moyen aux épreuves), la présente perspective ouvre modestement de nouveaux horizons prometteurs

- pouvoir déterminer avec suffisamment d'objectivité, quelles proportions d'élèves atteignent des niveaux de compétence donnés;
- contribuer à une définition plus rigoureuse des socles de compétence et des niveaux d'exigence requis.

Le bilan effectué cette année en début de 3e année s'avère dans l'ensemble globalement satisfaisant : 69 % des élèves atteignent le niveau de compétences que le groupe concepteur de l'épreuve considère comme requis et 20 % sont sur la voie de cette maîtrise. Mais pour cette moyenne plutôt satisfaisante, que de disparités ! Dans les classes coexistent en effet des élèves maîtrisant les seules compétences "élémentaires" et des élèves possédant des compétences "avancées" ... Le problème de l'hétérogénéité des compétences au sein des classes se pose plus que jamais.

La confrontation entre les résultats à l'épreuve et le niveau de la classe ouvre d'autres perspectives intéressantes. Si dans la majorité des cas, le jugement du titulaire et le résultat au test s'accordent dans les grandes lignes, les quelques 10 % de classes où se produisent des discordances manifestes appellent une réflexion, principalement chez les titulaires concernés.

Des perspectives didactiques, enfin, s'esquissent déjà. Les résultats de l'épreuve invitent, en effet à se pencher sur plusieurs questions essentielles:

- pourquoi les compétences très bien maîtrisées à l'état isolé le sont-elles moins dans un contexte plus large (résolution de problème, situation de jeu, rédaction de la carte) ?
- pourquoi des compétences exercées très systématiquement au cours du 1^{er} cycle (par exemple : la lecture de l'heure) ne sont-elles pas mieux maîtrisées ?

Le document "Pistes didactiques" qui sera rédigé et envoyé ultérieurement dans les classes abordera ces questions, parmi d'autres, ainsi que celle, cruciale au regard des résultats, de la gestion de l'hétérogénéité de compétences des élèves et de la différenciation des apprentissages.

GLOSSAIRE

Corrélation

Une corrélation est une liaison entre deux variables. Le coefficient de corrélation nous informe de l'importance de cette liaison. Ce coefficient peut varier de -1 à +1. Un coefficient égal à 0 indique une relation nulle entre les variables considérées. Un coefficient positif signifie que les variables tendent à évoluer de concert et dans le même sens. C'est, par exemple, le cas de la corrélation entre l'année scolaire et la vitesse de lecture: plus les élèves avancent dans leur scolarité, plus leur vitesse de lecture augmente. A l'opposé, une corrélation négative signifie que les variables évoluent de concert mais en sens contraire. C'est, par exemple, le cas de la corrélation entre le nombre de jours d'absence et les résultats scolaires: plus les élèves sont absents, moins bons sont leurs résultats.

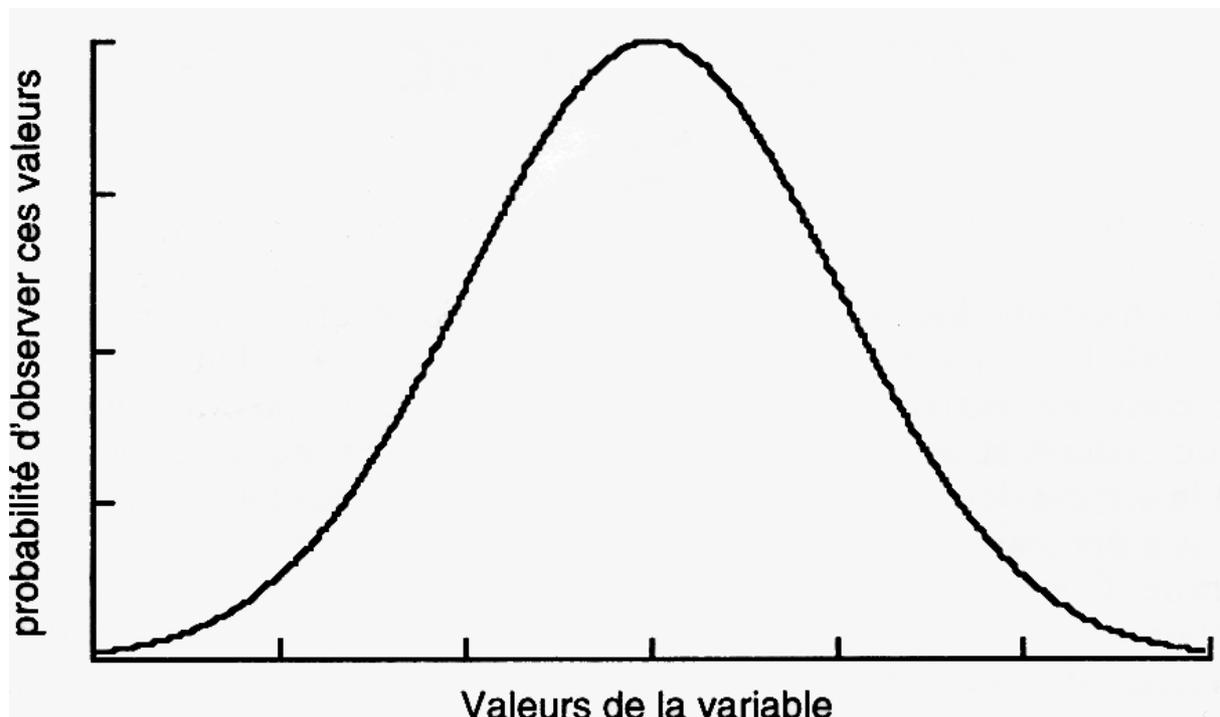
Il faut souligner que l'existence d'une corrélation n'implique pas ipso facto une relation de cause à effet entre les variables considérées. Ainsi, la corrélation positive entre l'année scolaire et la vitesse de lecture n'implique pas que la première variable soit la cause de la seconde. Dans ce cas, la relation entre les deux variables découle des règles de passage d'une année à l'autre (si l'élève lit correctement, il passera vraisemblablement dans l'année supérieure, et réciproquement) et de l'entraînement à la lecture qui augmente au cours des années. Cet exemple montre bien qu'un coefficient de corrélation n'a pas de sens par lui-même. Il doit toujours être l'objet d'une interprétation.

Par ailleurs, une corrélation non nulle peut être due au hasard. Toutefois, plus cette corrélation s'écarte de 0, moins il est probable qu'elle soit le fruit du hasard. Un coefficient de corrélation est considéré comme significatif si la probabilité qu'il résulte du simple jeu du hasard est très faible (habituellement, cette probabilité est inférieure à une chance sur cent). La signification statistique d'un coefficient de corrélation est directement liée à la taille de l'échantillon de sujets utilisé pour le calculer.

Le fait qu'un coefficient de corrélation soit significatif ne signifie pas que la liaison entre les variables est nécessairement très étroite. Il existe des degrés dans l'intensité de cette liaison. Ainsi, un coefficient de 0,90 (ou -0,90) indique une relation beaucoup plus étroite entre les variables qu'un coefficient de 0,30 (ou -0,30), même si ces deux coefficients sont significatifs.

Courbe normale

La distribution normale a été définie de manière précise par Gauss (1777-1855). Par conséquent, la forme de cette distribution est appelée indifféremment « courbe normale » ou « courbe de Gauss ». Le graphique ci-dessous offre une illustration de la distribution normale. L'axe horizontal représente toutes les valeurs possibles que peut prendre une variable. L'axe vertical représente la probabilité d'observer ces différentes valeurs au sein de la population. On peut constater que les valeurs centrales ont une forte probabilité d'être observées. Par contre, plus on s'éloigne du centre, moins les valeurs ont de chance d'être observées. Dans le domaine de l'éducation, il est fréquent d'observer une distribution des résultats des élèves proche de la distribution normale: beaucoup d'élèves obtiennent des résultats moyens et peu d'élèves obtiennent des résultats faibles ou élevés.

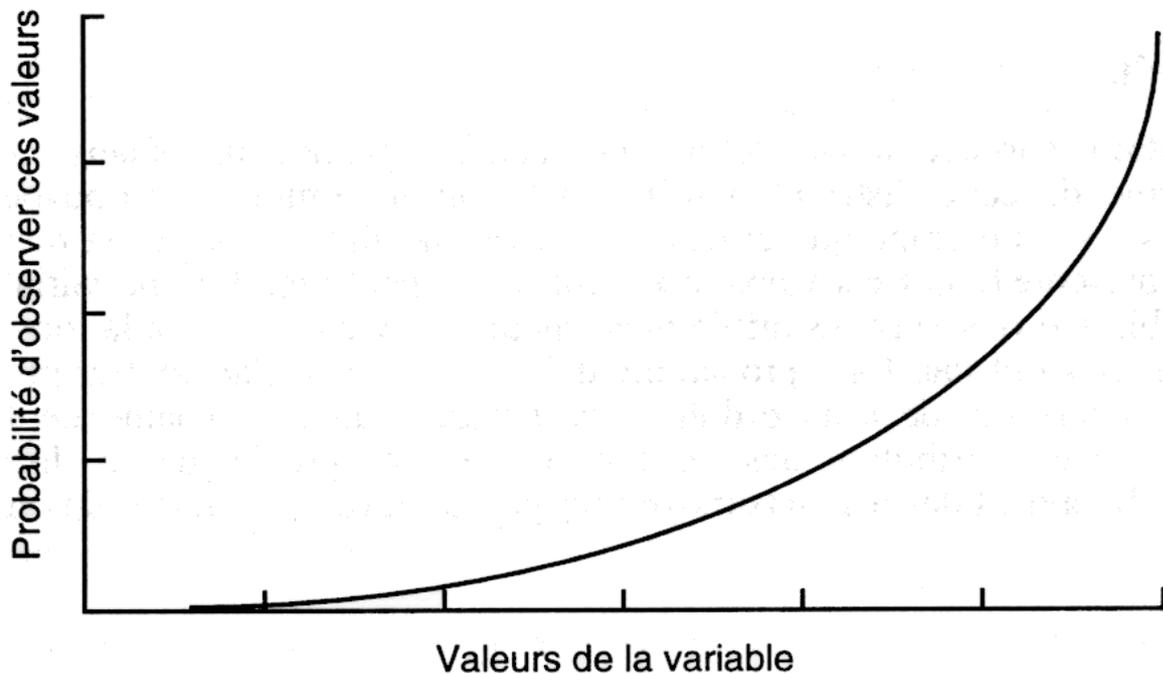


Courbe en J

Souvent, les acteurs du système éducatif considèrent que la distribution normale des résultats des élèves est une fatalité. Pourtant, l'objectif de l'enseignement est, au contraire, d'obtenir une distribution des résultats en forme de « J » majuscule (voir graphique): une majorité d'élèves obtient un résultat proche du maximum et une minorité obtient un résultat plus faible. Une telle distribution s'observe régulièrement à la fin des cycles d'apprentissage, lorsque les élèves ont, dans leur grande majorité, atteint le niveau de maîtrise visé.

La forme de la distribution des résultats est liée aux pratiques d'enseignement. Si l'enseignement est le même pour tous et correspond au niveau moyen des élèves de la classe, la distribution des résultats aura vraisemblablement la forme d'une courbe de Gauss. En effet, la matière enseignée risque d'être peu abordable pour les élèves faibles, de niveau adapté pour les élèves moyens et très facile pour les élèves doués. Les résultats d'un tel enseignement tendent à reproduire la distribution gaussienne des compétences initiales des élèves.

De nombreux pédagogues, en particulier Bloom, ont encouragé les enseignants à différencier leur enseignement en fonction des caractéristiques des élèves. En agissant de la sorte, il est possible d'amener la majorité des élèves à maîtriser les objectifs de l'enseignement. Si c'est le cas, la distribution des résultats des élèves prendra la forme d'une courbe en « J » .



Ecart-type

L'écart-type est une valeur statistique qui nous donne une information à propos de la dispersion des résultats au sein d'une distribution. Plus les résultats sont dispersés, plus l'écart-type est important; et réciproquement.

Analyse de Rasch

La difficulté d'un item est classiquement définie comme la proportion de sujets qui répondent correctement à cet item. Par conséquent, si les sujets testés sont faibles, l'item sera considéré comme difficile. Par contre, si les sujets testés possèdent un niveau de compétence élevé, l'item sera considéré comme facile. Au début des années 60, le mathématicien danois Georg Rasch a proposé une nouvelle façon d'estimer la difficulté des items qui ne soit plus relative à l'échantillon des sujets. L'analyse de Rasch permet de mettre en évidence une hiérarchie des items particulièrement robuste puisque l'ordre des items et la distance qui les sépare restent identiques quel que soit l'échantillon de sujets.

Dans le modèle proposé par Rasch, la difficulté d'un item est définie en terme de probabilité de réussite. Plus le niveau de compétence d'un élève est élevé, plus sa probabilité de réussir un item donné est élevée; et réciproquement. Cette conception de la difficulté est réaliste. En effet, un élève très compétent a rarement 100% de chance de réussir un item donné. De même, un élève peu compétent a souvent une chance supérieure à 0 de réussir ce même item.