

Projet CLEF – Bénin

**Traitement et analyse des résultats de l'examen de juillet 1998
du CEP nouveaux programmes d'études**

Rapport de mission de Léo Laroche

(21-01-99 au 05-02-99)

pour The Mitchell Group Inc

Mars 1999

Table des matieres

Introduction	1
Travaux preparatoires	2
Activités de perfectionnement	4
Organisation et logistique	4
Collecte des données sur les epreuves de juillet 1998	5
Activités de formation a l'utilisation du logiciel EduStat	5
Premiere sensibilisation	6
Formation d'une equipe technique	6
Analyse des résultats obtenus aux examens de juillet 1998	7
Les epreuves de français	7
Expression écrite	7
Comprehension orale	9
Comprehension écrite	11
Vocabulaire	12
Mathématique	13
Education scientifique et technologique	15
Education sociale	16
Education physique et sportive	19
Dessin	20
Fusion des resultats	21
Tenue de rencontres et activites quotidiennes	22
Le jeudi, 21 janvier 1999	22
Le vendredi, 22 janvier 1999	23
Le samedi, 23 janvier 1999	23
Le lundi 25 janvier 1999	23
Le mardi, 26 janvier 1999	24
Le mercredi, 27 janvier 1999	25
Le jeudi, 28 janvier 1999	26
Le vendredi, 29 janvier 1999	27
Le samedi, 30 janvier 1999	28
Le lundi, 1 ^{er} fevrier 1999	29
Le mardi, 2 fevrier 1999	29
Le mercredi 3 fevrier 1999	29
Le jeudi, 4 fevrier 1999	30
Le vendredi 5 fevrier 1999	30
Mesures susceptibles d'ameliorer la qualite des prochains examens du CEP	31
La preparation des epreuves	31
La notation des epreuves	32
Le controle de la qualite	32
Les compilations statistiques	34
La realisation d'analyses complementaires	34
Un sommaire des operations a realiser	34
Annexe « A » Definition de quelques indices statistiques	36
Mesure de tendance centrale	36
Mesure de dispersion des resultats	36
Mesure de discrimination	36
Mesure de coherence interne	37

Annexe « B » Resultats obtenus par les eleves aux differents epreuves	38
Annexe « C » Coefficients de correlation entre les resultats obtenus aux differentes epreuves	41
Annexe « D » L'instrumentation	42
Démarche d'elaboration d'une epreuve	42
Particularite de l instrumentation utilisee pour realiser une evaluation sommative	43
Annexe « E » La definition de domaine / la preparation du tableau de specification	44
Analyser le programme en vue de construire un tableau de specification	44
Les contenus notionnels	45
Les habiletés intellectuelles	45
La ponderation des contenus notionnels et des habiletés intellectuelles	45
Constituer le tableau de specification	46
Annexe « F » La correction des items a reponses construites	47
Determination des criteres de correction	47
Formation a la correction	47
Contrôle de la qualite de la correction	48
Corrections de groupe	48
Double correction	48
Identification des « deviances »	48
Annexe G Le controle de la qualite	49

Introduction

La présente mission au Bénin est la suite d'une demande faite en octobre 1998. Ce rapport touche la mission de **Leo Laroche**, ci-après nommé **le consultant**, qui s'est déroulée du 21 janvier au 5 février 1999. Voici un rappel des termes de référence de cette mission. En collaboration avec les membres de l'équipe du Plan d'Actions « Evaluation des connaissances et orientation des élèves » et l'Assistante Technique du Projet CLEF, le Consultant devra

- 1- Assurer la formation de membres de l'équipe d'évaluation à la dernière version du logiciel « EduStat », logiciel de traitement et d'analyse des données relatives aux examens et tests d'évaluation,
- 2- Confectionner les fichiers descripteurs et y transférer toutes les données collectées par la Direction des Examens et Concours sur les résultats du CEP nouveaux programmes,
- 3- Examiner rigoureusement les épreuves retenues dans chaque champ de formation (analyse d'items),
- 4- Effectuer une analyse statistique et pédagogique des résultats obtenus par les candidats au niveau de chaque épreuve du CEP nouveaux programmes,
- 5- Identifier toutes les insuffisances constatées et proposer des recommandations pertinentes en vue de l'amélioration du contenu et de l'administration des prochaines sessions du CEP nouveaux programmes

Le présent rapport se compose des chapitres suivants

- Travaux préparatoires
- Activités de perfectionnement
- Analyses statistiques
- Tenue de rencontres et activités quotidiennes
- Mesures susceptibles d'améliorer la qualité des prochains examens du CEP

Quelques annexes contiennent des renseignements complémentaires

- 1 Définition de quelques indices statistiques
- 2 Résultats obtenus par les élèves aux différentes
- 3 Coefficients de corrélation entre les résultats obtenus aux différentes épreuves
- 4 L'instrumentation
- 5 La définition de domaine / la préparation du tableau de spécification
- 6 La correction des items à réponses construites
- 7 Le contrôle de la qualité

Travaux préparatoires

Quelques mois avant la venue du consultant au Bénin, il y a eu échange de renseignements afin de mieux préparer les travaux qui devaient prendre place dans le cadre de cette consultation. C'est ainsi qu'un exemplaire des examens du CEP nouveaux programmes utilisés en juillet 1998 a été rendu disponible. Des fichiers de données préparés par la Direction des Examens et Concours ont aussi été expédiés au consultant. Il a été possible d'examiner les documents reçus afin de mieux connaître le format des examens utilisés. Quant aux bases de données, il s'est avéré que l'information disponible était peu utile pour le type de travaux devant être réalisés dans le cadre de la mission. Les fichiers contiennent des renseignements utiles pour des fins de gestion de l'examen du CEP. Cependant, les termes de référence élaborés pour ce séjour se réfèrent davantage à la conduite d'analyses de tous les éléments d'information présents sur les copies remises par les candidats au CEP. Il a donc fallu exploiter directement les copies remplies par les élèves et faire la codification des renseignements disponibles, c'est à partir de ces renseignements qu'il a été possible de réaliser les activités prévues.

Le logiciel EduStat a déjà été utilisé au Bénin dans le cadre d'une évaluation sur échantillons qui s'est déroulée il y a quelques années. La version Dos de ce logiciel avait alors été acquise. Ce logiciel est maintenant disponible dans l'environnement Windows (Windows 95 ou supérieur). Plusieurs améliorations ont été apportées depuis l'utilisation qui en a été faite au Bénin à partir de 1994. Une copie du Guide d'utilisation du logiciel a été remise au Pilote (ce Guide est d'ailleurs directement accessible au moment de l'utilisation du logiciel). Cependant, le consultant a préparé une série de « films » illustrant le fonctionnement des différentes options du logiciel. Plusieurs modules sont disponibles et ont servi d'introduction aux différentes activités de formation, de plus, les participants ont eu la possibilité de revoir à leur convenance ces « films » fonctionnant de façon autonome sur un micro-ordinateur. Voici la liste de ces modules de présentation du logiciel EduStat.

- 1 Première visite descripteur, exploitation de données, commandes (durée 4 minutes)
- 2 Création d'un fichier descripteur (durée 9 minutes)
- 3 Entrée de données avec EduData (durée 2 minutes)
- 4 Importation de données, exploitation de la base, distribution de fréquences (durée 3 minutes)
- 5 Fréquences, analyse d'items et résultats moyens – format de rapports texte (durée 4 minutes)
- 6 Fréquences, analyse d'items et résultats moyens – format de rapports RTF (durée 5 minutes)
- 7 Ouverture d'un rapport en format RTF (durée 2 minutes)
- 8 Fréquences choix d'enregistrements et variable de distribution (durée 2 minutes)
- 9 Tableau de résultats moyens de fréquences (durée 4 minutes)
- 10 Analyse d'opinion (durée 2 minutes)

11 Visite du module d'échantillonnage (duree 5 minutes)

Ces fichiers auto-executables pourront être installes sur des ordinateurs du projet CLEF Il sera alors possible de les visionner de façon periodique

Activités de perfectionnement

Cette partie du rapport traite des activités tenues au cours de l'atelier prévu dans le cadre de cette mission. Plus de vingt personnes ont participé activement aux exercices proposés. Le consultant a proposé de s'inspirer des pratiques courantes afin d'illustrer les concepts liés à l'évaluation pédagogique. C'est ainsi, qu'il a été jugé important de partir de copies remplies par les élèves lors de la session de l'examen de juillet 1998 pour faire l'analyse du fonctionnement de la procédure en usage pour l'attribution du Certificat d'études primaires. Cette façon de faire a permis de situer les aspects importants de l'évaluation pour fins de sanction dans un contexte connu des participants.

Après avoir fourni quelques informations sur l'organisation de ces activités et des aspects liés à la logistique, nous examinerons la méthode utilisée pour obtenir l'information requise sur les épreuves de juillet 1998. Par la suite, nous pourrions aborder les activités de perfectionnement à l'utilisation du logiciel EduStat, ces activités se sont tenues en deux moments. Dans un premier temps, l'ensemble des participants a pu s'initier à la micro-informatique à partir d'EduStat, par la suite, une équipe technique a poussé son perfectionnement à l'utilisation de ce logiciel.

Organisation et logistique

Au début de la deuxième semaine de cette mission, quatre micro-ordinateurs ont été rendus disponibles pour la tenue des activités de formation sur le logiciel EduStat. Trois micros ont pu être installés, le dernier demeurant non branché à cause de la non disponibilité de modulateur. Après quelques heures de fonctionnement, le modulateur utilisé par un poste de travail cesse de fonctionner. Seulement deux appareils demeurent alors disponibles. Le consultant réussit à brancher un troisième ordinateur sur l'unique modulateur disponible. Il faudra se contenter de trois appareils pour la suite des opérations. L'imprimante prévue pour produire des rapports n'est pas disponible.

Par ailleurs, il s'avère que les appareils fournis pour cet atelier ne permettent pas l'exploitation des « films » de présentation du logiciel EduStat, la capacité graphique des machines est insuffisante. Les films sont installés sur un ordinateur utilisé par le personnel du projet CLEF. Les participants à l'atelier sont invités à visionner les films. Cependant, il n'y a pas de réelle intégration de ces films à la formation donnée.

Comme il y a peu d'appareils et que le local est plutôt petit, il faut constituer deux groupes pour faire les présentations sur le logiciel, diminuant ainsi le temps prévu pour chacun à des exercices pratiques.

Collecte des données sur les épreuves de juillet 1998

Au tout début de la mission, il a été décidé de se servir des données recueillies en juillet dernier à l'occasion de l'examen du CEP, nouveaux programmes. On souhaitait couvrir l'ensemble des champs de formation, c'est-à-dire

- Le français ,
- La mathématique ,
- L'éducation scientifique et technologique ,
- L'éducation sociale ,
- L'éducation physique et sportive ,
- L'éducation artistique

Le personnel de la Direction des Examens et Concours a vu à mettre à la disposition des participants un échantillon d'une centaine de copies pour les différentes épreuves retenues. Les mêmes élèves ont été sélectionnés afin de garantir la comparabilité des échantillons et afin de pouvoir établir des liens entre les résultats obtenus aux différentes épreuves.

Pour chaque épreuve choisie, les opérations suivantes ont été réalisées

- 1 Choix de l'échantillon de copies
- 2 Examen de l'épreuve et du guide de correction
- 3 Fabrication d'un formulaire de codification
- 4 Notation et codification de l'échantillon de copies
- 5 Fabrication du fichier descripteur compatible au logiciel EduStat
- 6 Entrée des données et création d'une base de données EduStat
- 7 Production de tableaux statistiques à l'aide du logiciel EduStat

Neuf épreuves ont été soumises à cette démarche de recorection et de préparation d'une base de données. Ces opérations se sont déroulées au cours de la deuxième semaine de la tenue de cette mission.

Activités de formation à l'utilisation du logiciel EduStat

Les termes de référence indiquent une formation à l'utilisation du logiciel EduStat. Dès la première activité de formation, il s'est avéré que pour la majorité des participants, il s'agissait d'un premier contact avec l'ordinateur. Aussi, le consultant a proposé deux séries d'activités de formation à l'utilisation de ce

logiciel Dans un premier temps, il y a eu sensibilisation generale destinee a l'ensemble des participants Par la suite, des activites ont ete proposees a un equipe plus reduite

Premiere sensibilisation

Les premiers jours de la formation ont porte sur la manipulation des differents peripheriques utilises pour communiquer de l'information a l'ordinateur Il y a eu des exercices pratiques de manipulation de la souris, certains eprouvant des difficultes a realiser les operations souhaitees Par la suite, il y a eu la presentation des differentes fenetres prevues dans un logiciel comme EduStat C'est ainsi que les participants ont appris a selectionner un ou plusieurs elements a lancer l'execution d'une operation, a quitter une fenetre Ils ont pu apprendre a demarrer le logiciel, a rendre active une base de donnees, a explorer les variables comprises dans la base active, a realiser certaines compilations Ils ont examine a l'ecran les rapports produits

Les differentes activites se sont tenues dans la bonne humeur, les participants s'entraidant Avant le debut et apres la fin des activites, plusieurs faisaient des exercices supplementaires afin de mieux maıtriser ce qui avait ete vu

Ces activites de formation, męme si elles ont porte sur les prealables a l'utilisation de l'ordinateur, ont suscite beaucoup d'interę de la part des participants On peut cependant se demander si le cadre d'une telle mission se prętait bien a la tenue d'activites de sensibilisation a la micro-informatique

Formation d'une equipe technique

Une equipe limitee a eu l'occasion d'approfondir ses connaissances sur le fonctionnement du logiciel EduStat Ces activites se sont tenues au cours de la troisieme semaine de la mission Les participants ont prepares plusieurs fichiers descripteurs Il faut mettre en application les connaissances acquises si l'on souhaite ne pas « perdre » les habiletés acquises au prix de nombreux efforts

Analyse des résultats obtenus aux examens de juillet 1998

Cette section du rapport concerne l'analyse statistique et pédagogique réalisée à partir des résultats rendus disponibles par la codification et la recorection de copies d'élèves de juillet 1998

Comme il est indiqué ailleurs, trois équipes ont examiné les tableaux statistiques sur l'analyse d'items confectionnés à l'aide du logiciel EduStat. Un document rédigé par une équipe réduite met en lumière l'analyse faite par les participants à cet atelier. Nous ne reprendrons pas ici les conclusions tirées de ces travaux. Nous pourrions formuler des commentaires généraux sur chaque épreuve à la suite de l'examen du matériel disponible. Le consultant a dû expliquer à plusieurs occasions le sens des indices statistiques apparaissant sur les tableaux. Comme aide-mémoire, un court texte comportant des définitions de ces indices a été préparé, l'annexe « A » contient ce document qui a été remis aux participants.

Les épreuves de français

Quatre épreuves de français ont été examinées : 1) l'expression écrite, 2) la compréhension orale, 3) la lecture et 4) l'orthographe. Nous ne disposons pas de tableaux de spécification propres à chaque épreuve, ceci a rendu plus difficile l'examen de la couverture des épreuves par rapport à l'instrumentation utilisée.

Expression écrite

L'épreuve d'expression écrite invitait les élèves à rédiger une lettre sur un thème. La grille de correction utilisée comprend six critères, certains se subdivisant en unités plus fines. Rapidement, les participants ont noté que les correcteurs n'étaient pas tenus de reporter leur correction de façon détaillée, une note d'ensemble a été inscrite sur les copies.

Voici donc les tableaux statistiques prepares pour cette epreuve

Expression ecrite / Examen du CEP – Juillet 1998

Analyse classique d'items

Items a reponses continues

Items	N=	Minimum	Maximum	Moyenne	Rés/100	Corr.
NO_I	104	0,000	1 000	0 863	86 298	0,225
NO_II_1	104	1 000	3 000	2 635	87 821	0 445
NO_II_2	104	0,500	2 000	1 082	54 087	0 727
NO_II_3	104	0,500	2 000	0 803	40 144	0,654
NO_III	104	1,000	4 000	1 971	49 279	0,682
NO_IV_1	104	0 500	3,000	1 221	40 705	0,751
NO_IV_2	104	0 500	2 000	0 822	41 106	0,536
NO_V	104	0 000	1 000	0 683	68 269	0 415
NO_VI	104	0 500	2 000	1 346	67 308	0 445

Coefficient Alpha de Cronbach 0,657

Distribution de frequences des resultats

Classes de Resultats	Nombre d'eleves	%
22,000 <29,000	1	0,962
29,000 <36,000	2	1,923
36,000 <43,000	13	12,500
43,000 <50,000	19	18,269
50,000 <57,000	15	14,423
57,000 <64,000	19	18,269
64,000 <71,000	18	17,308
71,000 <78,000	12	11,538
78,000 <85,000	3	2,885
85,000 <90,000	2	1,923
Total	104	100,000

Statistiques descriptives

Etendue	22 500-	90 000
Moyenne	57 127	
Ecart type	13 359	
Erreur type	1 316	
Mediane	57 5	

Voici certains constats que l'on peut faire a la suite de l'examen de ces statistiques ainsi que de l'instrumentation utilisee

- Le minimum prevu pour certains criteres a ete fixe a une valeur superieur a 0, cette façon de faire diminue l'etendue possible des resultats obtenus par les eleves

- Les coefficients de corrélation items / résultats d'ensemble sont assez élevés indiquant que ces résultats respectent le modèle habituellement attendu, c'est-à-dire que ce sont les élèves les plus performants qui obtiennent de meilleures notes à chaque critère
- La moyenne générale obtenue par l'ensemble de cet échantillon de 104 élèves n'est pas tellement élevée (57,1 sur cent), cependant, les élèves sont assez regroupés autour de cette note moyenne, l'écart type étant plutôt faible
- La consistance interne des résultats obtenus par les élèves aux différents critères prévus dans la grille est plutôt faible. Les critères retenus ne mesurent probablement pas une même dimension de la langue

Comprehension orale

Cette épreuve d'une durée de 30 minutes était constituée d'un texte lu par l'examinateur ainsi qu'un questionnaire comportant 5 questions

Voici les résultats obtenus à l'analyse d'items réalisée à partir des données de 104 candidats

Comprehension orale / Examen du CEP -- Juillet 1998

Analyse classique d'items

Items à choix multiple

Items	N°	%	Pt Bis	Résultats moyens	
				Succès	Échec
Q_1	104	96 154	0 350	77 170	45 875
Q_2A	104	98 077	0 179	76 397	54 000
Q_2B	104	98 077	0 382	76 887	29 000
Q_2C	104	98 077	0 281	76 642	41 500
Q_2D	104	99 038	0 086	75 820	91 000
Q_2E	104	99 038	0 166	76 248	47 000
Q_3	104	89 423	0 478	78 796	52 045
Q_4A	104	97 115	0 170	76 470	59 000
Q_4B	104	76 923	0 337	79 144	65 375
Q_4C	104	56 731	0 170	78 525	72 611
Q_4D	104	67 308	0 463	81 514	64 544
Q_4E	104	56 731	0 168	78 492	72 656
Q_5A1	104	58 654	0 782	87 262	59 942
Q_5A2	104	49 038	0 731	88 784	63 632
Q_5B1	104	53 846	0 726	87 518	62 490
Q_5B2	104	45 192	0 723	89 649	64 684

Compilations réalisées en janvier 1999

Coefficient Alpha de Cronbach 0,769

Distribution de frequences des resultats

Classes de Resultats	Nombre d'eleves	%	
26 0 < 34 0	2	1 9	1 9
34 0 < 42 0	0	0 0	1 9
42 0 < 50 0	3	2 9	4 8
50 0 < 58 0	11	10 6	15 4
58 0 < 66 0	21	20 2	35 6
66 0 < 74 0	10	9 6	45 2
74 0 < 82 0	13	12 5	57 7
82 0 < 90 0	11	10 6	68 3
90 0 < 98 0	25	24 0	92 3
98 0 <100 0	8	7 7	100 0
Total	104	100 0	100 0

Statistiques descriptives

Etendue	26 0-	100 0
Moyenne	76 0	
Ecart type	17 2	
Erreur type	1 7	
Mediane	80 0	

Voici quelques commentaires que l'on peut faire a la suite de l'examen de ces tableaux statistiques

- Les deux premieres questions posees aux eleves sont beaucoup trop faciles. L'information recueillie ne permet pas de discriminer les candidats qui maîtrisent et ceux qui maîtrisent moins les habiletés évaluées.
- On peut se demander si certaines questions posees aux candidats mesurent véritablement leurs habiletés a comprendre l'information contenue dans le texte lu, il faut veiller a poser des questions qui permettent d'évaluer les habiletés a l'écoute plutôt que l'appel a des connaissances générales acquises par ailleurs.

Comprehension ecrite

L'eleve disposait d'un texte a lire ainsi que d'un questionnaire comprenant 5 questions. Voici les resultats de l'analyse d'items realisee avec un echantillon de 103 candidats, il est a noter qu'il s'agit ici d'un nouveau rapport produit apres une nouvelle saisie des donnees

Lecture / Examen du CEP -- Juillet 1998

Analyse classique d'items

Items a choix multiple

Items	N=	%	Pt.Bis.	Résultats moyens	
				Succes	Échec
NO_1	103	92,233	0 404	81 789	53 750
NO_4	103	72 816	0 551	85 867	62 857

Lecture / Examen du CEP -- Juillet 1998

Analyse classique d'items

Items a reponses continues

Items	N=	Minimum	Maximum	Moyenne	Res/100	Corr
NO_2A	103	0 000	2 500	1 893	75 728	0 244
NO_2B	103	0 000	2 500	1 456	58 252	0 480
NO_3	103	0 000	5 000	4 320	86 408	0 510
NO_5	103	0 000	5 000	3 981	79 612	0 573

Coefficient Alpha de Cronbach 0,258

Distribution de frequences des resultats

Classes de Resultats	Nombre d'eleves	%
40,000 < 46,000	8	7,767
46,000 < 52,000	6	5,825
52,000 < 58,000	0	0,000
58,000 < 64,000	8	7,767
64,000 < 70,000	0	0,000
70,000 < 76,000	10	9,709
76,000 < 82,000	29	28,155
82,000 < 88,000	0	0,000
88,000 < 94,000	12	11,650
94,000 <100,000	0	0,000
100,000 <100,000	30	29,126
Total	103	100,000

Statistiques descriptives

Etendue	40 000-	100 000
Moyenne	79 612	
Ecart type	18 587	
Erreur type	1 840	
Mediane	80 000	

Voici quelques commentaires tenant compte des résultats obtenus à cette épreuve

- Il y a eu erreur au moment de la saisie des données. Un contrôle fait sur quelques copies d'élèves indique que certains élèves ont obtenu des notes supérieures à 0 par rapport à la question 5. Ce fait met en évidence l'importance de prévoir des mécanismes de contrôle de la qualité tout au long du processus.
- La moyenne obtenue à cette épreuve est assez élevée, s'établissant à près de 80%.
- La valeur du coefficient Alpha est très faible, indiquant un manque de consistance interne de l'épreuve.

Vocabulaire

Cette épreuve comprenait un texte comportant quinze mots incorrectement orthographiés, 9 mots concernent l'orthographe grammaticale alors que les 6 autres visent l'orthographe d'usage.

Voici les statistiques complètes à partir de l'échantillon de 106 copies recorrigées

Composition d'orthographe / Examen du CEP -- Juillet 1998

Analyse classique d'items

Items à réponses continues

Items	N°	Minimum	Maximum	Moyenne	Rés/100	Corr.
NO_1	106	0,000	0 500	0,429	85 849	0 380
NO_2	106	0,000	0 500	0 439	87 736	0,249
NO_3	106	0,000	0 500	0,439	87 736	0,382
NO_4	106	0,000	0,750	0 590	78 616	0,556
NO_5	106	0,000	0 750	0,481	64 151	0,499
NO_6	106	0,000	0 750	0,134	17 925	0,475
NO_7	106	0,000	0 750	0,361	48 113	0,555
NO_8	106	0,000	0,750	0,410	54 717	0,368
NO_9	106	0,000	0 500	0,420	83 962	0 289
NO_10	106	0 000	0 500	0 443	88 679	0 420
NO_11	106	0 000	0 750	0 434	57 862	0 458
NO_12	106	0 000	0 750	0 580	77 358	0 557
NO_13	106	0 000	0 750	0 500	66 667	0 480
NO_14	106	0 000	0 750	0 425	56 604	0 506
NO_15	106	0 000	0 750	0 531	70 755	0 521

Coefficient Alpha de Cronbach 0,750

Distribution de frequences des resultats

Classes de Resultats	Nombre d'eleves	%
- < 5,000	1	0,943
5,000 < 15,000	3	2,830
15,000 < 25,000	1	0,943
25,000 < 35,000	0	0,000
35,000 < 45,000	3	2,830
45,000 < 55,000	15	14,151
55,000 < 65,000	16	15,094
65,000 < 75,000	25	23,585
75,000 < 85,000	19	17,925
85,000 < 95,000	18	16,981
95,000 < 100,000	5	4,717
Total	106	100,000

Statistiques descriptives

Etendue	5 000-	100 000
Moyenne	67 803	
Ecart type	19 115	
Erreur type	1 874	
Mediane	70 000	

Voici quelques commentaires formulés à la suite de l'examen de ces tableaux statistiques

- L'orthographe d'usage (items 1, 2, 3, 8, 9 et 10) est mieux réussie que l'orthographe grammaticale
- Les indices de difficulté permettent de bien discriminer les candidats
- Les coefficients de corrélation possèdent des valeurs acceptables

Mathématique

Cette épreuve comporte trois problèmes soumis aux candidats. Chaque problème vise un domaine différent à évaluer. Le problème numéro 1 concerne l'arithmétique, le deuxième problème touche la mesure. Enfin, le dernier vise la géométrie. Il a été impossible de rattacher les items aux habiletés objets d'évaluation.

Voici les tableaux statistiques produits a partir de l'échantillon de 106 candidats

Mathematique / Examen du CEP – Juillet 1998

Analyse classique d'items

Items a reponses continues

Items	N#	Minimum	Maximum	Moyenne	Rés/100	Corr.
Q_1A	106	0 000	2 000	1 406	70 283	0 479
Q_1B	106	0 000	1 000	0 302	30 189	0 496
Q_1C	106	0 000	3 000	0 335	11 164	0,416
Q_1D	106	0 000	1 000	0 094	9 434	0 435
Q_2A	106	0 000	2 000	0 934	46 698	0 550
Q_2B	106	0 000	2 000	0 278	13 915	0 221
Q_2C	106	0 000	2 000	0 085	4 245	0 285
Q_3A	106	0 000	1 000	0 335	33 491	0,489
Q_3B	106	0,000	3,000	1 189	39 623	0,575
Q_3C	106	0 000	1,000	0 585	58 491	0,270
Q_3D	106	0 000	2 000	0 849	42 453	0 667

Coefficient Alpha de Cronbach 0,632

Distribution de frequences des resultats

Classes de Resultats	Nombre d'eleves	%
0,000 < 8,000	7	6,604
8,000 <16,000	16	15,094
16,000 <24,000	19	17,925
24,000 <32,000	13	12,264
32,000 <40,000	11	10,377
40,000 <48,000	14	13,208
48,000 <56,000	15	14,151
56,000 <64,000	6	5,660
64,000 <72,000	4	3,774
72,000 <77,500	1	0,943
Total	106	100,000

Statistiques descriptives

Etendue	0 000-	77 500
Moyenne	32 519	
Ecart type	18 331	
Erreur type	1 789	
Mediane	28 889	

Nous pouvons faire les observations suivantes

- Cet examen a ete tres peu reussi par les candidats (resultat moyen de 32,5) En realisant la recorection, nous avons constate que des notes supplementaires ont ete attribuees aux eleves On peut se demander sur quelle base ces notes supplementaires ont ete attribuees
- Les coefficients de correlation sont acceptables, ceci indique qu'en general, les items discriminent bien les eleves maîtrisant ce qui est evalue de ceux qui ont une moins bonne maîtrise

Éducation scientifique et technologique

Cette épreuve comprend plusieurs items touchant plusieurs aspects de ce domaine. Ceci augmente la qualité de l'instrument, l'élève devant réaliser un éventail plus complet d'objets d'interrogation.

Voici les tableaux statistiques réalisés à partir d'un échantillon de 84 candidats.

Éducation scientifique et technologique / Examen du CEP - Juillet 1998*Analyse classique d'items*

Items à réponses continues

Items	N=	Minimum	Maximum	Moyenne	Res/100	Corr.
Q_1A	84	00	20	17	86.9	0.5
Q_1B	84	00	20	17	82.7	0.3
Q_2A	84	00	05	02	45.2	0.3
Q_2B	84	00	15	05	35.3	0.4
Q_3A	84	00	15	12	81.7	0.4
Q_3B	84	00	15	13	84.5	0.6
Q_3C	84	00	10	05	53.0	0.5
Q_3D	84	00	20	13	66.1	0.4
Q_4A	84	00	20	17	83.9	0.6
Q_4B	84	00	05	02	45.2	0.4
Q_4C	84	00	05	03	54.8	0.5
Q_4D	84	0,0	20	07	35.7	0.5
Q_5A	84	00	10	08	80.4	0.4
Q_5B	84	00	20	14	72.0	0.3
Q_6A1	84	00	10	09	85.7	0.1
Q_6A2	84	00	15	07	44.0	0.3
Q_6B1	84	00	10	08	76.2	0.3
Q_6B2	84	00	15	06	36.9	0.4
Q_7A	84	00	20	18	92.3	0.0
Q_7B	84	00	30	20	66.1	0.4

Coefficient Alpha de Cronbach 0,722

Distribution de frequences des resultats

Classes de Resultats	Nombre d'eleves	%
18 0 < 27 0	1	1 2
27 0 < 36 0	2	2 4
36 0 < 45 0	1	1 2
45 0 < 54 0	11	13 1
54 0 < 63 0	10	11 9
63 0 < 72 0	21	25 0
72 0 < 81 0	17	20 2
81 0 < 90 0	13	15 5
90 0 < 99 0	7	8 3
99 0 < 100 0	1	1 2
Total	84	100 0

Statistiques descriptives

Etendue	19 0-	100 0
Moyenne	69 6	
Ecart type	16 1	
Erreur type	1 8	
Mediane	70 9	

L'examen des resultats obtenus permet de faire ces commentaires

- Les items possèdent des indices de difficulté varies, il s'agit d'une pratique courante de varier la difficulté des items
- Mis a part deux items (Q_6A1 et Q_7A), les coefficients de corrélation sont acceptables
- La dispersion des resultats moyens obtenus a l'épreuve est acceptable (ecart type un peu supérieur a 15)

Éducation sociale

Cette épreuve comporte six questions. Les statistiques obtenues a partir de l'échantillon de 106 élèves sont présentées dans les tableaux suivants

Education sociale / Examen du CEP – Juillet 1998
Analyse classique d'items
 Items a choix multiple

Items	N=	%	Pt.Bis.	Résultats moyens	
				Succes	Échec
Q_4A	106	6,604	0,263	42 437	61 735
Q_5A	106	89 623	0 215	61 796	48 930

Compilations réalisées en janvier 1999

Education sociale / Examen du CEP – Juillet 1998*Analyse classique d'items*

Items a reponses continues

Items	N=	Minimum	Maximum	Moyenne	Rés/100	Corr.
Q_1A	106	0 000	1 500	1 358	90 566	0 580
Q_1B	106	0 000	1,500	0 882	58 805	0 579
Q_2A	106	0 000	1,500	1 165	77 673	0 570
Q_2B	106	0,000	1 500	1 401	93 396	0 348
Q_3A	106	0 000	1 000	0 783	78 302	0 597
Q_3B	106	0 000	2 000	0 670	33 491	0 538
Q_4B	106	0 000	1 000	0 538	53 774	0 376
Q_4C	106	0 000	1,000	0 420	41 981	0 521
Q_5B	106	0 000	1 000	0 208	20 755	0 306
Q_6A	106	0 000	1 000	0 759	75 943	0 379
Q_6B1	106	0 000	1 000	0 604	60 377	0 601
Q_6B2	106	0 000	1 000	0 528	52 830	0 630

Coefficient Alpha de Cronbach 0,701

Distribution de frequences des resultats

Classes de Resultats	Nombre d'eleves	%
0,000 <10,000	1	0,943
10,000 <20,000	1	0,943
20,000 <30,000	6	5,660
30,000 <40,000	7	6,604
40,000 <50,000	10	9,434
50,000 <60,000	20	18,868
60,000 <70,000	18	16,981
70,000 <80,000	32	30,189
80,000 <90,000	10	9,434
90,000 <94,118	1	0,943
Total	106	100,000

Statistiques descriptives

Etendue	0 000-	94 118
Moyenne	60 461	
Ecart type	18 252	
Erreur type	1 781	
Mediane	64 706	

Nous presentons quelques commentaires sur cette epreuve

- Il est preferable de prevoir des directives precises destinees aux eleves. Parfois, l'eleve avait le choix de fournir une liste de deux ou de trois elements
- Les coefficients de correlation sont acceptables
- Les resultats moyens se regroupent assez bien autour de la moyenne qui s'etablit a 60,5

Éducation physique et sportive

Cette épreuve est constituée d'un ensemble de grille d'observation permettant de recueillir de l'information sur les aspects suivants l'endurance, la vitesse et la gymnastique

Voici les statistiques produites à partir de 90 observations

Éducation physique et sportive / Examen du CEP – Juillet 1998**Analyse classique d'items****Items à réponses continues**

Items	N=	Minimum	Maximum	Moyenne	Rés/100	Corr
END_1	90	0 0	8 0	7 7	96 3	0 8
END_2	90	0 0	12 0	6 9	57 5	0 8
VIT_1	90	0 0	5 0	4 3	85 3	0 6
VIT_2	90	0 0	3 0	2 8	93 7	0 7
VIT_3	90	0 0	12 0	6 9	57 4	0 7
GYM_1A	90	0,0	4 0	3 1	77 9	0 7
GYM_1B	90	0 0	4 0	3 0	74 6	0 7
GYM_1C	90	0 0	4 0	3 1	78 3	0 7
GYM_1D	90	0 0	4 0	3 1	76 5	0 7
GYM_2	90	0 0	4 0	2 9	72 5	0 6

Coefficient Alpha de Cronbach 0,870

Distribution de fréquences des résultats

Classes de Résultats	Nombre d'élèves	%
- <45,0	3	3,3
45,0 <51,0	1	1,1
51,0 <57,0	2	2,2
57,0 <63,0	4	4,4
63,0 <69,0	10	11,1
69,0 <75,0	17	18,9
75,0 <81,0	31	34,4
81,0 <87,0	11	12,2
87,0 <93,0	8	8,9
93,0 <97,5	3	3,3
Total	90	100,0

Statistiques descriptives

Etendue	45 8-	97 5
Moyenne	76 0	
Ecart type	9 3	
Erreur type	1 0	
Mediane	76 7	

Nous presentons quelques commentaires sur cette epreuves

- Les valeurs des coefficients de correlation sont elevees, indiquant que les eleves ayant une bonne performance generale reussissent bien a chaque element mesure
- Le coefficient Alpha est eleve, ceci indique une bonne consistance interne de l'epreuve
- L'ecart type est faible, cela signifie que les resultats moyens obtenus par les eleves se regroupent bien autour de la moyenne, c'est-a-dire 76,0

Dessin

L'eleve etait invite a dessiner le sujet de son choix a partir d'une visite faite anterieurement. La production des eleves fut notee a partir de ces criteres : l'utilisation de l'espace, la conformite de la production, le respect des proportions et l'utilisation des couleurs. Voici les statistiques produites a partir des 105 dessins

Dessin / Examen du CEP -- Juillet 1998

Analyse classique d'items

Items a reponses continues

Items	N=	Minimum	Maximum	Moyenne	Rés/100	Corr.
ELEMENT_1	105	0 500	1 000	0 833	83 333	0 340
ELEMENT_2	105	1 000	2 000	1 219	60 952	0 661
ELEMENT_3	105	0,500	1 000	0 724	72 381	0 497
ELEMENT_4	105	0 500	1 000	0 662	66 190	0 334

Coefficient Alpha de Cronbach -0,120

Distribution de frequences des resultats

Classes de Resultats	Nombre d'eleves	%	
50,000 < 55,000	9	8,571	
55,000 < 60,000	0	0,000	
60,000 < 65,000	33	31,429	
65,000 < 70,000	0	0,000	
70,000 < 75,000	36	34,286	
75,000 < 80,000	0	0,000	
80 000 < 85,000	17	16,190	
85,000 < 90,000	0	0,000	
90,000 < 95,000	9	8,571	
95,000 <100,000	0	0,000	
100,000 <100,000	1	0,952	
Total	105	100,000	100,000

Statistiques descriptives

Etendue	50 000-	100 000
Moyenne	68 762	
Ecart type	11 101	
Erreur type	1 089	
Mediane	70 000	

L'examen des statistiques disponibles permet de faire ces observations

- Les quatre criteres utilises pour faire l'appréciation des production semblent être independants les uns par rapport aux autres La valeur du coefficient Alpha est nulle
- L'eleve obtenait au depart 50 pourcent des notes L'etendue des notes possibles se trouvent par le fait même tres reduite

Fusion des resultats

Les resultats moyens obtenus par les eleves aux differentes epreuves ont ete fusionnes dans une base de donnees L'annexe « B » etablit la liste des resultats obtenus par chaque eleve aux differentes epreuves Des coefficients de correlation ont ete calcules pour examiner les liens qui s'etablissent entre les resultats obtenus par chaque eleve, l'annexe « C » presente ces coefficients Il semble s'etablir des liens assez etroits entre les differents resultats obtenus par les mêmes eleves il y a toutefois deux exceptions le dessin et l'education physique et sportive Dans ces deux cas, les eleves ayant de bons resultats ne sont pas necessairement pas ceux qui reussissent bien dans les autres domaines

Tenue de rencontres et activités quotidiennes

Voici un rappel des rencontres tenues dans le cadre de cette mission ainsi que les activités réalisées pendant le séjour

Le jeudi, 21 janvier 1999

Une première rencontre de planification de la mission a lieu aux bureaux de Cotonou du Projet CLEF. Des responsables du CLEF et des responsables nationaux de différents dossiers reliés aux volets pédagogiques sont présents à cette réunion. Il y a tout d'abord rappel des termes de références de cette mission.

Au cours de l'après-midi, il y a rencontre à Porto-Novo avec une vingtaine de personnes qui ont été invitées à participer à l'une ou l'autre activité prévue au cours de cette mission. Dans un premier temps, Michel Toudonou, le pilote du Plan d'actions sur l'évaluation, fait une proposition en ce qui concerne la suite des travaux. Le calendrier proposé est découpé en trois périodes différentes.

- Réalisation de travaux préparatoires (durée 2 jours)
- Formation à EduStat (durée 3 jours)
- Analyse statistique et pédagogique (durée 6 jours)

Après échanges de points de vue, il s'est avéré que l'on devait assurer une plus grande intégration des différents aspects du programme d'activités. Si l'on souhaite une véritable analyse de données recueillies à l'aide des examens du CEP de juillet 1998, il faut disposer de données exploitables par un logiciel de traitement statistique. Une autre planification des activités est alors proposée. Il s'agit de choisir un échantillon de copies d'élèves, d'en faire la codification et la notation afin d'être en mesure d'entrer ces renseignements sur un support informatique compatible au logiciel qui sera utilisé. Par la suite, il sera possible de faire l'exploitation des bases de données disponibles. La formation vise donc l'ensemble du processus de la codification des données à l'exploitation des données et aux analyses statistiques et pédagogiques.

Pour réaliser ces activités, il faut disposer rapidement des copies remplies par les élèves en juillet 1998. Un échantillon doit en être prélevé. Il est convenu de retenir les mêmes élèves pour les différents examens qui seront analysés, il y aura possibilité d'établir des liens entre les différents examens.

Les appareils informatiques nécessaires à la réalisation des activités ne sont pas encore disponibles. Il faudra les avoir le plus rapidement possible, si l'on souhaite respecter le calendrier arrêté. Une rencontre à la Direction des Examens et Concours permet de s'entendre sur les modalités d'obtention des copies nécessaires. Le directeur de la DEC assigne quatre personnes de sa direction aux travaux en cours.

Le vendredi, 22 janvier 1999

Au debut de la journee, il y a assemblage des differents documents necessaires a la poursuite des travaux les epreuves, les cles de correction, les guides d'administration, les tableau de specification, un echantillon de copies d'eleves. Le materiel relie a l'examen de mathematique etant disponible, le groupe travaille cette discipline. Dans un premier temps, il y a preparation d'un formulaire de codification qui permettra de recueillir les differents elements d'information fournis par les eleves. Le formulaire est multiplie par la suite. Deux autres formulaires sont par la suite prepares : la comprehension orale et l'education sociale.

Les responsables de la DEC choisissent un echantillon de copies d'eleves en comprehension orale. Chaque participant fait la correction d'une dizaine de copies. Apres examen du temps disponible, on decide de retenir une centaine de copies par discipline. Il est necessaire de revenir sur les objectifs vises par cette operation de codification et de correction des copies retenues. Il ne s'agit pas toujours de porter un jugement sur le travail fourni par les eleves, il faut le plus souvent inscrire sur le formulaire les reponses fournies par les eleves, le traitement statistique permettra d'apprécier la performance des eleves.

Le consultant recueille les copies corrigees ainsi que les formulaires de codification remplis afin d'en faire l'entree a l'ordinateur. Lors de la prochaine rencontre, il sera possible d'utiliser ces donnees pour illustrer l'ensemble de la demarche.

Le samedi, 23 janvier 1999

Le consultant prepare le fichier descripteur pour l'examen de comprehension orale. Il lui est alors possible d'entrer les donnees a l'aide du logiciel EduData. De premieres compilations statistiques sont realisees avec les donnees de cette discipline.

Par la suite, le consultant commence la redaction du rapport de la mission.

Le lundi, 25 janvier 1999

Les rapports produits a partir de la base de donnees sur la comprehension orale ont ete imprimes. Le consultant a, par la suite, passe en revue les differentes statistiques disponibles. Chaque question comprise dans cette epreuve a pu être analysee. Cet examen a fourni l'occasion de presenter certaines notions de statistiques : mesures de tendance centrale, indices de dispersion, distribution de frequences, resultats moyens, indice de difficulte, indice de discrimination, correlation, consistance interne. Il faudra revenir sur ces notions statistiques a l'occasion de l'analyse des donnees des autres disciplines.

L'examen des statistiques relatives a l'epreuve de comprehension orale permet de faire les commentaires suivants :

- Le resultat global obtenu par les eleves est assez eleve (le resultat moyen s'etablit a 76,0 sur 100)
- Plusieurs items ont un taux de succes quasi-parfait, diminuant ainsi leur discrimination

- La consistance interne de l'épreuve est moyenne

Après ces échanges d'information sur l'épreuve de compréhension orale, il y a correction et codification d'un échantillon de copies en mathématique. Il y a aussi début d'installation des ordinateurs prévus pour la tenue des activités de formation.

Deux techniciens en informatique réalisent l'entrée des données de l'épreuve de mathématique. Il y a correction et codification d'un échantillon de copies en éducation sociale. En fin d'après-midi, l'équipe prépare un projet de formulaire de codification de l'épreuve d'éducation scientifique et technologique.

En soirée, le consultant prépare le formulaire de codification de l'épreuve d'éducation scientifique et technologique. Il y a aussi finalisation du formulaire qui sera utilisé pour codifier les données de l'épreuve d'éducation physique et sportive.

Le mardi, 26 janvier 1999

Il y a production du formulaire de codification de l'épreuve d'éducation scientifique et technologique. Le consultant examine les fichiers de mathématique préparés par l'équipe technique. Les fichiers sont enregistrés sur les différents ordinateurs qui seront utilisés pour les activités de formation.

Il est convenu de tenir une première activité de formation. Le groupe est fractionné en deux équipes. Un premier groupe participe à une activité de formation. Cette première activité de formation aborde les aspects suivants :

- Présentation générale de l'ordinateur et de ses périphériques
- Démarrage du logiciel EduStat
- Présentation générale des différents menus de ce logiciel
- Examen de l'aide en contexte
- Ouverture de la base de données de mathématique
- Examen de la structure de cette base de données (le fichier descripteur)

Pendant que le premier groupe participe à l'activité de perfectionnement, les autres débutent la correction de l'épreuve d'éducation scientifique et technologique.

Dans un deuxième temps, il y a permutation des équipes pour permettre à tous les participants d'avoir une présentation de la première activité de formation.

Le consultant installe sur un micro-ordinateur les différents « films » disponibles (voir la section *Travaux préliminaires*, ci-dessus). Les participants pourront les consulter à leur convenance.

Au debut de l'apres-midi, le consultant invite les participants a exprimer leurs commentaires par rapport aux activites tenues depuis l'arrivee du consultant. Les echanges touchent les quatre points suivants

Realisations Il y a eu preparation de formulaires de codification d'un certain nombre d'epreuves. Par la suite, un echantillon de copies a ete corrige. Les participants ont eu l'occasion de se voir presenter un certain nombre de concepts statistiques. A titre d'illustration, des tableaux statistiques sur la composition orale ont ete examines. Il y a eu premiers contacts avec la micro-informatique.

Formation Les participants ont pu être initiés a plusieurs aspects relies au traitement des donnees recueillies a l'occasion de l'administration d'une epreuve. On voit mieux l'importance de bien preparer les activites relies a la correction des epreuves et d'etablir des mecanismes de contrôle de la qualite. Tous les participants se sont montres interesses a recevoir une initiation a la micro-informatique.

Benefices Les participants croient que les connaissances acquises permettront de mieux preparer le materiel necessaire a la tenue de l'examen du CEP. Il faudra accorder une grande attention a la correction.

Attentes Les participants souhaitent parfaire leurs connaissances en analyses des donnees statistiques. On veut approfondir ses connaissances en micro-informatique. Enfin, on croit que l'on devrait faire la codification d'un autre ensemble d'epreuves.

Une deuxieme activite de formation se deroule par la suite. Les points suivants sont abordes

- Les participants sont invites a reviser les notions vues lors de la derniere activite
- Il y a ensuite importation des donnees sur l'echantillon de copies corrigees en mathematique
- Il y a exploration des donnees comprises dans la base de donnees a l'aide de la commande **Explorer les donnees**

Après le déroulement de l'activite de formation, l'equipe technique debute l'entree des donnees du formulaire de codification d'education sociale. En soiree, le consultant fait l'entree des donnees du formulaire d'education scientifique et technologique. Le consultant realise des analyses d'items a partir des bases de donnees de mathematique et d'education scientifique et technologique.

Le mercredi, 27 janvier 1999

Dans un premier temps, le consultant fait un sommaire des notions informatiques presentees la veille. Certains expriment l'idee de disposer d'un manuel complet d'utilisation de l'ordinateur. Le consultant indique que le fonctionnement de l'ordinateur s'acquiert de façon intuitive. Il faut être en contact frequemment avec l'informatique pour developper les habiletés nécessaires a une utilisation efficiente d'un logiciel. Il y a des modalites de fonctionnement semblables d'un logiciel a l'autre. Enfin, des questions sont posees en ce qui concerne les virus informatiques.

Par la suite, le groupe est fractionné en deux équipes pour une session de formation à la micro-informatique. Après avoir révisé les notions déjà vues (lancement du programme et ouverture d'une base de données), il y a exploration du fichier ouvert. Cet exercice permet d'examiner les points suivants

- Sélection d'éléments sur une liste un ou plusieurs éléments
- Préparation de distribution de fréquences des valeurs d'une variable
- Préparation d'un tableau de fréquences à partir des valeurs de deux variables
- Visualisation des valeurs contenues dans une base de données

Par la suite, il y a correction de l'examen de dessin. Les données recueillies à partir de cette épreuve sont, par la suite, entrées à l'ordinateur.

Les membres du groupe examinent la grille de correction de l'épreuve d'expression écrite. Il s'agit de préparer un formulaire de codification qui sera utilisé pour corriger un échantillon de copies d'élèves. Tout l'après-midi est consacré à discuter des limites de la grille disponible.

Le jeudi, 28 janvier 1999

Pendant que les participants à l'atelier préparent la codification des épreuves de lecture et d'orthographe, il y a évaluation à mi-parcours de la présente mission. Le consultant rappelle les activités déjà réalisées et celles en cours de réalisation, il y a indication quant aux prochaines activités. Presque toutes les épreuves du CEP de juillet 1998 (nouveaux programmes) auront été examinées. Nous disposons de données permettant de nous prononcer sur les caractéristiques métriques de ces épreuves. La dernière semaine de la mission sera consacrée d'une part à l'analyse des données disponibles sur les épreuves recorrigées et d'autre part à la formation d'une équipe réduite à l'utilisation du logiciel EduStat.

Au cours de l'après-midi s'engage une discussion très animée sur l'épreuve d'expression écrite. On constate que les copies disponibles ne contiennent qu'une note globale sur 20, il est alors impossible de connaître la répartition des notes par rapport à la grille de correction utilisée en juillet 1998. On s'entend pour recorriger un nombre limité de copies, au préalable, il semble nécessaire de préciser certains éléments de la grille de correction. Pendant qu'un groupe va mettre en application les notions apprises à l'ordinateur, les autres participants débutent la correction de l'épreuve d'expression écrite.

Le vendredi, 29 janvier 1999

Au debut de la matinee, le consultant presente les differents tableaux statistiques qui seront utilisees par les participants pour realiser les analyses. Il y a tout d'abord les statistiques relatives a l'analyse d'items. Des indices statistiques sont rendus disponibles par rapport a chaque element d'information, chaque item faisant partie de l'epreuve. Deux indices sont ordinairement utilises pour se prononcer sur la qualite metrique des items : la difficulte et la discrimination. Ce type d'analyse permet de se prononcer sur la qualite docimologique des differentes composantes d'une epreuve. Si toutes les conditions sont respectees, il sera possible d'utiliser les scores d'ensemble qui servent habituellement d'indicateurs pour l'evaluation sommative realisee dans le cadre du CEP.

Des echanges tres animes entre les participants ont lieu sur le sens a accorder au fait que plusieurs eleves devraient pouvoir obtenir des resultats « parfaits » aux epreuves du CEP. On veut mettre en place une pedagogie de la reussite.

Le consultant presente d'autres tableaux statistiques mettant en evidence des distributions de frequences par rapport aux resultats d'ensemble obtenus a une epreuve. Ceci donne lieu a indiquer le sens de certaines statistiques : la moyenne, la mediane, l'ecart type, l'etendue.

Pendant qu'un groupe debute la correction des epreuves de lecture et d'orthographe, les autres participants vont dans le local d'informatique. L'equipe technique realise la saisie des donnees pour les dernieres epreuves corrigees. Le consultant prepare les tableaux statistiques pour chaque epreuve examinee. Ces tableaux sont reproduits et distribues aux participants. A partir du lundi suivant, il y aura analyse de ces resultats.

Au cours de l'apres-midi, il y a tenue d'une autre activite de formation a l'utilisation d'EduStat. Les points suivants sont abordes :

- La production de statistiques avec l'utilisation d'une variable de distribution
- Le calcul de resultats moyens

Le samedi, 30 janvier 1999

Le consultant réalise des travaux au cours de cette journée. Neuf épreuves ont été recorrigées. Nous disposons d'un ensemble important de renseignements sur chaque épreuve retenue. Voici les nombres d'éléments d'information actuellement disponibles, nous ajoutons le nombre de copies retenues pour chaque épreuve.

- Français

Expression écrite	9 éléments	104 copies
Compréhension orale	16 éléments	103 copies
Compréhension écrite	6 éléments	102 copies
Vocabulaire	15 éléments	104 copies

- Mathématique
- Education scientifique et technologique
- Education sociale
- Education physique et sportive
- Dessin

	14 éléments	105 copies
	20 éléments	84 copies
	14 éléments	103 copies
	13 éléments	90 copies
	4 éléments	105 copies

Comme les mêmes élèves ont été retenus pour chaque épreuve – la ou c'était possible – il y a eu fusion de l'information par rapport aux résultats globaux obtenus à chaque épreuve après avoir considéré les réponses fournies aux différents éléments d'information disponibles. Un tableau rend compte de ces résultats obtenus par chaque élève à chaque épreuve retenue (voir l'Annexe « B »). Des corrélations ont été calculées afin de constater les liens entre la performance obtenue à une discipline par rapport à celle obtenue aux autres. Ces résultats se retrouvent à l'annexe « C ». L'examen des coefficients de corrélation indique qu'il s'établit des liens assez étroits entre les résultats obtenus par les mêmes élèves aux différentes épreuves, il y a toutefois deux exceptions que l'on doit signaler : l'épreuve d'éducation physique et sportive et l'épreuve de dessin.

Le consultant a aussi participé à une fête organisée dans une école de la région pour souligner la mise à la retraite de huit enseignants et de deux enseignantes.

Le lundi, 1^{er} février 1999

Le groupe est fractionné en quatre équipes

- Analyse des épreuves de français et d'éducation sociale
- Analyse des épreuves d'éducation scientifique et technologique et d'éducation sociale
- Analyse des épreuves de mathématique et d'éducation physique et sportive
- Équipe technique en informatique

Le consultant rend visite à chaque équipe responsable de l'analyse des données par rapport à l'une ou l'autre épreuve

En parallèle, il y a formation de l'équipe technique à la constitution de fichiers descripteurs compatibles au logiciel EduStat. Dans un premier temps, le film numéro 2 sur le fichier descripteur est visionné. Par la suite, les membres de l'équipe technique fabriquent le fichier descripteur se rapportant à l'épreuve de dessin. Après avoir préparé le descripteur, il y a utilisation d'EduData pour faire l'entrée des données d'un lot de 5 copies. Par la suite, il y a importation de ce fichier et interrogation de la base de données (option Explorer les données). Comme dernier exercice portant sur cette base, le film 2 est visionné une deuxième fois.

Au cours de l'après-midi, l'équipe technique débute la préparation du fichier descripteur pour l'épreuve de mathématique. Le travail sera complet le lendemain.

Le mardi, 2 février 1999

Le consultant tient des rencontres avec les équipes chargées de faire l'analyse des résultats obtenus aux épreuves. De son côté, l'équipe technique poursuit la définition du fichier descripteur déjà débuté. Par la suite, par équipe de deux, les participants définissent chacun un fichier descripteur différent.

Le mercredi, 3 février 1999

Il y a plénière pour partager les résultats de l'analyse statistique et pédagogique réalisée en équipes au cours des deux dernières journées.

Le jeudi, 4 février 1999

Une équipe réduite voit à mettre en forme le rapport rendant compte des analyses statistiques et pédagogiques réalisées à partir des données sur les épreuves du CEP de juillet 1998. L'équipe technique passe en revue les questions soulevées à l'utilisation d'EduStat. Il y a nouvelle entrée des données de lecture puisqu'une erreur avait été identifiée au moment de l'analyse des résultats.

Le vendredi, 5 février 1999

Il y a évaluation finale de cette mission de consultation sur les épreuves du CEP et sur l'utilisation d'EduStat. Les participants à cette rencontre souhaitent que des suites soient données à cet atelier. On croit que l'on devrait organiser une activité pour la préparation des examens du prochain CEP (nouveaux programmes), une telle activité devrait se tenir en avril prochain.

Mesures susceptibles d'améliorer la qualité des prochains examens du CEP

Les paragraphes suivants présentent quelques commentaires et recommandations susceptibles d'améliorer l'examen du CEP. Nous aborderons tour à tour les aspects suivants :

- La préparation des épreuves
- La notation des épreuves
- Le contrôle de la qualité
- Les compilations statistiques
- La réalisation d'analyses complémentaires
- La proposition d'une démarche d'opérations à réaliser

La préparation des épreuves

La procédure en usage pour choisir l'examen utilisé pour la certification de l'enseignement primaire permet au personnel scolaire de soumettre des épreuves. Par la suite, il y a sélection par différents responsables scolaires sous la responsabilité de la Direction des Examens et Concours.

Une telle procédure permet à plusieurs personnes de participer à la démarche de rédaction d'épreuves. Cependant, un tel mécanisme ne permet pas habituellement de prévoir une rétroaction destinée aux concepteurs de ces épreuves. D'un autre côté, il est pratiquement assuré que la plupart des épreuves ainsi élaborées ne respectent pas tous les critères en garantissant la rigueur et la haute qualité. Par exemple, est-on assuré qu'un tableau de spécification a été préparé afin de respecter le domaine à évaluer par chaque épreuve en développement ?

Plutôt que de lancer une telle invitation au milieu scolaire pour la rédaction d'épreuves, il serait sans doute plus approprié de confier le travail à une commission nationale formée de spécialistes de contenus ayant suivi une formation sur les points reliés au développement d'une épreuve. Il n'est pas nécessaire d'attendre la fin de l'année scolaire pour mettre en place une telle structure, le travail pourrait se réaliser au cours du premier trimestre de l'année scolaire afin de disposer du temps nécessaire pour accomplir les autres activités reliées à la tenue de l'examen de fin d'année pour le CEP. Si on décidait de procéder ainsi, il faudra toutefois prévoir une période de transition. Il serait donc possible de recueillir des épreuves à partir de productions réalisées par les enseignants. On pourrait alors fournir au personnel de la DEC des outils (grilles d'analyse, tableaux de spécification) leur permettant de juger de façon plus éclairée les épreuves disponibles.

L'annexe « D » fournit un sommaire de points reliés à l'instrumentation utilisée pour réaliser l'évaluation des apprentissages. Par ailleurs, l'annexe « E » traite de l'élaboration des tableaux de spécification.

La notation des épreuves

Des la phase d'élaboration des épreuves, il est nécessaire de prévoir le guide de correction. Pour les questions à choix multiple, il s'agit de prévoir la cle de correction, dans le cas de reponses courtes ou elaborees fournies par les eleves, il est important de prévoir la ou les reponses admissibles qui respectent les criteres. Une preparation attentive de cette partie de l'instrumentation diminue les risques de subjectivite au moment de la notation, les resultats obtenus par les eleves refletem alors mieux leur performance veritable.

L'organisation de la session de notation s'avere être un aspect important du mecanisme d'evaluation. Il y a tout d'abord le choix des correcteurs susceptibles de respecter le plus fidelement possible les consignes de correction et de realiser la notation de façon consistante. Une procedure de selection des correcteurs peut s'averer utile pour assister dans la selection des meilleurs candidats a la correction de l'examen de fin d'annee scolaire. Un test pourrait même être prepare pour être en mesure de porter un jugement objectif sur les qualites professionnelles attendues des correcteurs.

Avant de debuter la correction elle-même, il serait necessaire de soigner la formation du personnel a cette importante operation. C'est ainsi que le personnel assigne a la correction d'une epreuve devrait recevoir une information complete sur les caracteristiques de l'épreuve objet de la notation ainsi que sur les criteres de notation. L'examen de « copies-types » peut s'averer être une bonne strategie. Des exercices de correction de la même copie par l'ensemble des correcteurs peuvent permettre de mieux verifier la comprehension univoque des criteres de correction et detecter les correcteurs ayant besoin d'un complement de formation.

L'annexe F contient des renseignements complementaires sur la correction des items a reponses construites.

Le contrôle de la qualité

Il est necessaire de prévoir des mecanismes de contrôle tout au long de la procedure en usage pour delivrer le certificat d'etudes primaires. Ces contrôles devraient toucher tant le materiel utilisee (par exemple, les directives transmises au milieu scolaire, les formulaires en usage, les épreuves soumises aux eleves) que les mecanismes mis en place pour la realisation des operations (par exemple, la selection des épreuves, le choix des correcteurs, la procedure de notation, la transmission des resultats).

Les contrôles de qualite devraient tout d'abord toucher l'instrumentation elle-même. Il est important que les épreuves elaborees respectent des standards eleves de qualite et qu'elles soient preparees par rapport a des tableaux de specification respectant les domaines d'evaluation. On doit toujours avoir a l'esprit qu'il s'agit d'un examen de sanction qui atteste des connaissances et des competences des sortants de l'enseignement primaire. Voila pourquoi il apparaît necessaire, comme nous l'avons note ailleurs, de confier a une equipe nationale le soin de realiser cette preparation de l'instrumentation, cela peut se realiser en comptant sur des ressources provenant du milieu scolaire (inspecteurs, enseignants, par exemple). S'il est possible, le materiel elabore devrait être experimente.

Il serait approprié de vérifier le respect des procédures prévues au moment de la passation des épreuves par les candidats. Des visites non annoncées dans un échantillon d'écoles pourraient permettre de recueillir des renseignements pour apprécier la conformité aux procédures arrêtées. Il faudrait toutefois prévoir une grille d'observation afin de recueillir, auprès de l'échantillon retenu, l'information jugée significative.

Comme il est nécessaire de réaliser la correction d'un certain nombre d'épreuves, il est important de mettre en place des mécanismes permettant d'apprécier la qualité de la notation. Nous avons signalé plus haut le mécanisme de « notation de groupe », c'est-à-dire le fait de corriger une même copie. Il est possible d'examiner les résultats obtenus à un tel exercice pour être en mesure de mieux encadrer les correcteurs. Il est en effet souhaitable d'identifier le plus rapidement possible les correcteurs ou groupes de correcteurs dont la notation s'écarte de façon significative des critères fixes. Les correctifs qui s'imposent peuvent alors être apportés afin de garantir la qualité et la constance de la correction de toutes les copies d'élèves.

Les autorités du système scolaire peuvent souhaiter avoir des données permettant de juger de la « marge d'erreur » pouvant être due au mécanisme mis en place pour réaliser la notation. Un échantillon national de copies peut être constitué, un jury national se verrait alors confié la tâche de réaliser une deuxième correction de ces copies. Des analyses statistiques permettraient alors de quantifier la fidélité que l'on peut attribuer à l'opération, il s'agirait de comparer les deux séries de jugements pour tirer ces conclusions.

Après avoir réalisé la correction des copies d'élèves, différentes opérations doivent être exécutées dans le but de permettre la consignation des résultats et leur transmission à qui de droit, les autorités seront alors en mesure de poser les gestes administratifs liés à la délivrance du CEP. Encore ici, il faut prévoir des mécanismes de contrôle pour minimiser les risques d'erreurs. Pour cela, il importe d'examiner les procédures en usage et de s'interroger sur la pertinence d'introduire des technologies susceptibles de mieux garantir l'intégrité des données d'évaluation touchant chaque candidat.

L'annexe G établit un sommaire des opérations de contrôle pouvant être intégrées à une activité d'évaluation.

Les compilations statistiques

Le consultant n'a pas eu l'occasion d'examiner les résultats obtenus par les candidats à l'examen du CEP au cours des dernières années. Il est possible que les taux de succès à l'examen de fin d'année fluctuent de façon significative d'une année à l'autre. Si tel est la situation, cela est souvent attribuable aux caractéristiques des épreuves alors utilisées. Il peut y avoir risque de manquer d'équité envers certains groupes d'élèves. Le Certificat d'études primaires devraient attester de compétences équivalentes d'une année scolaire à l'autre.

Idealement, il faudrait réaliser des analyses statistiques préalables au calcul des résultats officiels afin de juger des qualités métriques des différentes épreuves. Les résultats attribués aux élèves ne devraient pas être influencés par des erreurs possibles contenues dans les épreuves ou introduites au moment de la notation.

S'il n'est pas possible de réaliser à court terme des analyses d'items pour juger de la qualité de la mesure, on pourrait à tout le moins s'assurer que les taux de succès (et ainsi d'échec) soient comparables d'une année à l'autre.

La réalisation d'analyses complémentaires

À l'occasion de la tenue de l'examen du CEP, plusieurs données sont recueillies (par exemple, renseignements sur les élèves, réponses des élèves aux épreuves, résultats de la correction). Ces renseignements sont exploités pour la réalisation des opérations requises à la certification des études primaires. Il pourrait y avoir avantage d'exploiter plus largement les données recueillies. Il n'est pas nécessaire de faire un tel dépouillement pour l'ensemble des élèves inscrits à l'examen de fin d'année, un échantillon représentatif peut permettre de disposer d'un ensemble de renseignements susceptibles de fournir des résultats utiles aux différents milieux scolaires.

Un sommaire des opérations à réaliser

Nous proposons ci-dessous une démarche pouvant être retenue pour réaliser les opérations nécessaires à la tenue de l'examen du CEP. En plus de cette liste d'opérations à accomplir, des annexes présentent des compléments d'information sur l'un ou l'autre aspects de la démarche.

Voici donc un sommaire des étapes à respecter dans la mise en place des opérations liées à l'examen du CEP

- 1 Elaboration de l'instrumentation (en annexes, on retrouve des compléments d'information par rapport aux étapes à respecter dans la mise en place d'un mécanisme d'évaluation ainsi que des renseignements sur la préparation du tableau de spécification)
 - Epreuves (en référence aux tableaux de spécification)
 - Guide d'administration
 - Consignes de correction
- 2 Administration des épreuves
 - Distribution du matériel
 - Contrôle de qualité
 - Récupération du matériel
- 3 Correction des épreuves (une annexe fournit des renseignements sur la correction des copies, une autre annexe établit un sommaire des activités de contrôle de la qualité)
 - Formation des correcteurs
 - Contrôle de qualité
 - Respect des consignes de correction
- 4 Compilation des résultats
 - Entrée des résultats sur informatique
 - Contrôle de qualité
- 5 Correction détaillée d'un échantillon de copies
 - Préparation de formulaires de codification
 - Notation et codification
 - Préparation de fichiers descripteur EduStat
 - Entrée des données
 - Compilations statistiques
- 6 Analyse des résultats
 - Analyse des résultats globaux obtenus par l'ensemble des candidats
 - Analyse des données recueillies à l'occasion de la correction détaillée
- 7 Formulation de recommandations

Annexe « A » Définition de quelques indices statistiques

Mesure de tendance centrale

- **Moyenne** Statistique qui s'obtient par la somme des scores divisée par le nombre d'observations, il s'agit d'une mesure de tendance centrale
- **Mediane** Le point sur l'échelle des valeurs au-dessous duquel tombe 50%, c'est-à-dire la moitié des scores, il s'agit d'une mesure de tendance centrale
- **Mode** Dans une distribution de scores, le mode se définit comme étant le score de fréquence maximum, il s'agit d'une mesure de tendance centrale

Mesure de dispersion des résultats

- **Ecart type** C'est la racine carrée de la moyenne arithmétique des carrés des écarts à la moyenne, il s'agit donc d'une mesure de dispersion. Plus la valeur de l'écart type est élevée, plus il y a dispersion des résultats autour de la moyenne
- **Etendue** Mesure de variabilité ou de dispersion d'une distribution qui provient de la différence entre les deux scores extrêmes, le plus fort et le plus faible

Mesure de discrimination

- **Correlation bisériale** La corrélation bisériale indique le lien qui existe entre la réponse fournie à un item et le résultat obtenu à l'épreuve. Pour réaliser ce calcul, il y a constitution, pour chaque item, de deux groupes d'individus : ceux qui ont réussi l'item et ceux qui l'ont échoué. Le programme compare alors les résultats obtenus à l'épreuve par ces deux groupes de personnes. Un taux de corrélation bisériale positivement élevé indique habituellement que ce sont les personnes les plus performantes qui ont réussi l'item et les moins performantes qui l'ont échoué. Il faut néanmoins considérer la difficulté de l'item : si celui-ci a été réussi ou échoué par la presque totalité des individus, le taux de corrélation peut alors prendre des valeurs extrêmes qu'il est difficile d'interpréter. On peut choisir l'un ou l'autre coefficient suivant la corrélation bisériale ou la corrélation du point biserial. Les formules utilisées pour le calcul de la corrélation du point-biserial et de la corrélation bisériale diffèrent. Aussi, les coefficients obtenus ne sont pas les mêmes pour le même item d'une même épreuve. Ordinairement, la valeur limite acceptable pour un coefficient biserial est de 0,30 alors que celle pour un coefficient du point-biserial est de 0,20.

- **Correlation de Pearson** Nous pouvons affirmer qu'il existe une relation entre deux ou plusieurs variables si leurs variations sont reliées l'une à l'autre. Le coefficient de corrélation indique donc le degré d'association existant entre deux variables. Une façon d'interpréter un coefficient de corrélation calculé à partir de deux séries de données est d'élever le coefficient obtenu au carré, le résultat de cette opération indiquera la proportion de la variation dans une variable qui s'explique par la présence de l'autre variable (on parle alors de coefficient de détermination). Ainsi, si on a obtenu une corrélation de 0,8 entre deux séries de données, on peut croire que 64% ($0,8 * 0,8$) de la variation dans une série de données s'explique par la présence de l'autre série.

Mesure de cohérence interne

- **Coefficient Alpha de Cronbach** La qualité qu'a un instrument de mesurer avec la même exactitude à chaque administration réalisée dans des conditions semblables se nomme la «fidélité». Il existe plusieurs méthodes pour apprécier la fidélité d'un instrument. C'est ainsi que le calcul d'un coefficient de consistance interne est souvent retenu. Ce coefficient a pour but de refléter le degré d'association ou de corrélation qui existe entre les variables faisant partie de l'instrument. Plusieurs coefficients peuvent être produits pour rendre compte de la consistance interne. La méthode Kuder-Richardson (KR-20) et la méthode Alpha de Cronbach sont les plus courantes. Au moment de l'analyse d'items, EduStat produit le coefficient Alpha qui convient aux épreuves contenant des items à choix multiple ou sur une échelle continue.

Annexe « B » Résultats obtenus par les élèves aux différents épreuves

#	EP1	EP2	EP3	EP4	EP5	EP6	EP7	EP8	EP9
0034	-	-	-	-	-	-	-	81,7	-
0036	75,0	85,0	100,0	78,9	41,2	73,3	76,5	85,8	70,0
0037	30,0	50,0	85,7	5,0	7,5	30,0	23,3	-	90,0
0039	-	-	-	-	-	70,0	-	-	-
0047	37,5	26,0	50,0	10,0	15,0	68,3	57,7	78,3	90,0
0052	58,8	45,0	75,0	85,0	12,5	89,1	46,4	81,7	80,0
0053	75,0	100,0	100,0	92,5	22,5	72,9	76,5	78,3	70,0
0086	65,0	100,0	100,0	92,5	67,6	80,0	94,1	79,2	90,0
0087	57,5	85,0	100,0	50,0	35,0	81,4	58,6	97,5	70,0
0130	71,3	82,5	87,5	65,0	10,0	70,0	67,9	58,0	60,0
0131	42,5	77,5	37,5	77,5	47,5	68,3	53,8	92,5	60,0
0132	67,5	70,0	75,0	62,5	45,0	86,2	75,0	-	70,0
0136	65,0	85,0	71,4	85,0	50,0	86,7	-	85,8	50,0
0166	57,5	85,0	100,0	80,0	52,5	96,7	61,8	73,3	60,0
0173	42,5	60,0	71,4	5,0	17,5	46,7	23,5	79,2	100,0
0174	50,0	-	75,0	65,0	30,0	51,7	50,0	78,3	60,0
0176	-	85,0	-	-	-	-	-	-	-
0177	62,5	65,0	87,5	80,0	45,0	61,7	50,0	69,2	60,0
0186	47,5	65,0	70,0	77,5	67,5	-	78,1	83,3	80,0
0189	45,0	65,0	70,0	65,0	47,5	-	52,9	83,3	80,0
0193	-	92,5	-	-	-	-	-	-	-
0195	41,3	-	60,0	65,0	20,0	-	76,5	89,2	50,0
0198	41,3	95,0	87,5	47,5	25,0	-	73,3	70,0	90,0
0201	50,0	80,0	75,0	70,0	40,0	-	38,2	70,8	80,0
0236	72,5	92,5	100,0	100,0	57,5	100,0	82,4	61,7	70,0
0237	60,0	70,0	62,5	85,0	52,5	53,6	75,0	75,0	90,0
0238	55,0	67,5	75,0	51,4	15,0	63,3	52,9	78,3	70,0
0239	67,5	100,0	100,0	-	50,0	51,7	58,8	72,7	80,0
0240	77,5	72,5	60,0	82,5	52,5	78,6	67,6	-	70,0
0286	70,0	90,0	75,0	91,9	27,5	98,2	70,6	55,8	60,0
0301	87,5	80,0	80,0	100,0	50,0	90,0	76,5	73,3	60,0
0320	90,0	95,0	100,0	92,5	55,0	98,2	87,5	-	60,0
0328	53,8	75,0	75,0	77,5	35,0	73,3	76,5	94,2	60,0
0335	-	87,5	75,0	57,5	25,0	80,0	61,8	-	60,0
0341	77,5	91,0	75,0	87,5	58,3	86,4	73,5	71,7	70,0
0364	80,0	87,5	100,0	75,7	60,0	86,4	81,3	60,4	90,0
0365	45,0	47,0	50,0	78,4	15,0	19,0	38,2	-	60,0
0374	65,0	80,0	100,0	72,5	45,0	83,3	82,4	75,8	70,0
0375	63,7	60,0	75,0	70,0	50,0	71,7	76,5	75,0	70,0
0386	67,5	95,0	100,0	62,2	20,0	82,8	87,5	72,5	60,0
0408	50,0	50,0	100,0	59,5	17,5	78,3	76,5	-	70,0
0422	72,5	95,0	75,0	65,0	10,0	76,3	88,2	63,3	60,0
0423	52,5	95,0	75,0	85,0	37,5	82,1	79,4	-	70,0

0435	60,0	70,0	75,0	77,5	27,5	86,7	76,5	81,7	80,0
0436	82,5	70,0	100,0	77,5	25,0	90,0	55,9	63,3	80,0
0450	70,0	100,0	100,0	70,0	62,5	88,3	73,5	87,5	80,0
0454	58,8	100,0	100,0	100,0	57,5	96,7	64,7	73,3	60,0
0455	37,5	62,5	50,0	50,0	45,0	76,7	70,6	-	80,0
0456	-	-	-	-	-	-	-	-	70,0
0470	45,0	87,5	100,0	77,5	55,0	75,0	76,5	89,2	-
0487	43,8	95,0	75,0	67,6	42,5	78,3	88,2	-	60,0
0492	50,0	100,0	100,0	83,8	52,9	66,7	67,6	-	60,0
0501	58,8	95,0	100,0	59,5	10,0	56,7	76,5	65,0	60,0
0502	-	-	-	47,5	-	-	-	-	-
0526	62,5	95,0	40,0	47,5	47,5	78,0	82,4	-	70,0
0529	71,3	95,0	60,0	87,5	52,5	-	82,4	78,3	70,0
0541	62,5	65,0	-	85,0	47,5	66,7	73,5	73,3	70,0
0549	40,0	65,0	87,5	20,0	5,3	51,8	44,1	76,7	70,0
0561	47,5	50,0	100,0	85,0	38,9	76,7	73,5	93,3	70,0
0571	60,0	65,0	100,0	70,0	65,0	81,7	50,0	80,0	70,0
0577	76,3	80,0	100,0	80,0	63,9	65,0	67,6	45,8	70,0
0585	-	-	-	-	45,0	-	-	-	-
0586	47,5	-	100,0	62,5	37,5	68,3	70,6	81,7	80,0
0602	60,0	95,0	75,0	72,5	55,0	76,7	58,8	55,8	80,0
0606	52,5	70,0	75,0	55,0	26,3	66,7	57,1	70,0	60,0
0607	67,5	65,0	50,0	77,5	20,0	46,7	58,8	67,5	50,0
0621	65,0	50,0	100,0	62,5	22,5	54,2	73,5	75,8	50,0
0638	40,0	70,0	62,5	57,9	27,5	66,7	64,7	77,5	80,0
0644	45,0	51,0	25,0	45,0	2,5	45,8	29,4	80,8	70,0
0650	47,5	65,0	80,0	62,5	14,7	55,0	32,4	76,7	80,0
0663	70,0	75,0	62,5	52,6	37,5	83,3	67,6	77,5	70,0
0685	47,5	55,0	60,0	65,8	65,0	61,7	20,6	74,2	60,0
0686	42,5	100,0	100,0	92,5	20,0	63,6	55,9	72,5	70,0
0691	61,3	95,0	100,0	85,7	17,6	61,7	-	88,3	70,0
0694	-	-	-	-	-	-	44,1	-	-
0696	52,5	70,0	87,5	70,0	10,0	56,7	65,6	79,2	70,0
0700	-	-	-	-	-	-	32,4	-	-
0701	47,5	80,0	87,5	70,0	36,8	73,3	76,5	80,8	80,0
0706	45,0	50,0	87,5	72,5	37,5	66,7	-	70,0	70,0
0736	22,5	-	0,0	-	-	-	0,0	-	50,0
0741	47,5	65,0	60,0	92,5	50,0	70,0	50,0	68,3	60,0
0746	60,0	65,0	100,0	100,0	77,5	71,4	67,6	75,8	90,0
0751	68,8	65,0	100,0	100,0	39,5	71,7	73,5	73,3	80,0
0756	47,5	95,0	75,0	65,0	17,5	-	47,1	65,0	70,0
0786	57,5	65,0	80,0	57,5	2,5	53,3	50,0	81,7	60,0
0787	46,3	55,0	100,0	50,0	20,0	73,3	50,0	80,0	70,0
0829	58,8	65,0	75,0	70,0	12,5	56,7	32,4	-	50,0
0833	47,5	52,5	80,0	92,5	32,5	51,7	41,2	75,8	60,0
0834	40,0	90,0	70,0	35,0	0,0	38,3	61,8	-	50,0
0836	67,5	100,0	100,0	82,5	15,0	-	50,0	70,8	70,0

0847	50,0	55,0	75,0	77,5	50,0	-	67,6	-	60,0
0864	-	90,0	-	-	-	-	-	-	-
0869	48,8	-	75,0	70,0	43,8	-	59,4	90,0	90,0
0882	47,5	80,0	87,5	57,5	0,0	-	59,4	79,2	60,0
0884	46,3	85,0	37,5	55,0	27,8	-	82,1	90,8	50,0
0886	56,3	85,0	75,0	50,0	2,5	64,9	29,4	80,0	70,0
0901	38,8	95,0	50,0	55,0	18,8	71,7	68,8	80,0	60,0
0910	31,3	32,0	40,0	51,4	20,0	61,3	23,5	65,0	60,0
0912	-	-	40,0	-	-	-	-	-	-
0918	50,0	60,0	-	65,0	15,0	30,0	47,1	63,3	70,0
0935	57,5	70,0	62,5	40,0	30,0	48,3	70,6	62,5	60,0
0936	75,0	95,0	75,0	80,6	26,5	-	73,5	73,3	70,0
0937	72,5	92,5	75,0	55,0	20,0	-	73,5	77,5	60,0
0945	66,3	45,0	71,4	70,3	25,0	-	50,0	76,7	60,0
0946	70,0	61,0	87,5	51,4	15,0	-	55,9	80,0	90,0
0947	67,5	92,5	100,0	85,0	37,5	-	-	-	70,0
1003	60,0	90,0	87,5	-	42,1	80,0	64,7	87,5	80,0
1006	41,3	60,0	87,5	65,0	10,0	70,4	38,2	82,5	60,0
1009	52,5	55,0	40,0	40,0	10,0	58,6	61,8	76,7	70,0
1012	36,3	60,0	25,0	45,0	17,5	49,1	14,3	64,2	70,0
1015	55,0	80,0	50,0	65,0	17,5	68,3	-	80,0	60,0
1036	72,5	80,0	37,5	47,5	12,5	-	87,5	-	80,0
1038	67,5	75,0	50,0	85,0	50,0	-	82,4	-	70,0
1039	70,0	80,0	100,0	62,5	25,0	-	73,5	80,8	60,0
1065	82,5	65,0	-	50,0	17,5	-	76,5	65,0	60,0
1067	55,0	90,0	-	65,0	20,0	-	64,7	83,3	50,0
1103	-	-	-	77,5	-	-	-	-	-
5291	-	-	-	-	-	91,7	-	-	-
9471	-	-	-	-	-	-	44,1	-	-

Legende

- EP1 Expression écrite
- EP2 Compréhension de l'oral
- EP3 Compréhension de l'écrit
- EP4 Orthographe
- EP5 Mathématique
- EP6 Education scientifique et technologique
- EP7 Education sociale
- EP8 Education physique et sportive
- EP9 Dessin

Annexe « C » Coefficients de corrélation entre les résultats obtenus aux différentes épreuves

	ECRIT	ORAL	LECTURE	ORTHO	MATH	EST	SOCIALE	EPS	DESSIN
ECRIT	1,000	0,383	0,390	0,461	0,283	0,522	0,539	-0,305	-0,004
ORAL		1,000	0,316	0,332	0,217	0,425	0,502	-0,056	-0,090
LECTURE			1,000	0,338	0,203	0,343	0,341	-0,040	0,190
ORTHO				1,000	0,559	0,464	0,365	-0,077	-0,153
MATH					1,000	0,514	0,366	-0,009	0,209
EST						1,000	0,569	0,060	0,046
SOCIALE							1,000	0,051	0,029
EPS								1,000	0,021
DESSIN									1,000

Annexe « D » L'instrumentation

L'instrumentation utilisée pour réaliser une évaluation des apprentissages doit être préparée avec soin. Nous examinons ci-dessous la démarche d'élaboration d'une telle instrumentation. Nous compléterons en examinant les particularités des instruments d'évaluation utilisés pour la certification des études.

Démarche d'élaboration d'une épreuve

Voici les étapes habituellement présentes dans le cycle de développement d'une instrumentation utilisée pour évaluer les apprentissages.

La définition du domaine ou la préparation du tableau de spécification

Étape au cours de laquelle on identifie, décrit et pondère les objets d'évaluation propres à la discipline concernée.

La spécification de l'épreuve

On choisit le type d'épreuve, on en détermine les paramètres comme la durée, le matériel autorisé, le mode de correction, le type d'items, d'exercices ou de tâches.

La validation du domaine de l'épreuve

Il s'agit de vérifier, auprès d'un groupe de spécialistes, si le domaine retenu représente adéquatement ce que l'on veut évaluer, on examine aussi si les spécifications de l'épreuve sont pertinentes.

La spécification des items, des exercices ou des tâches

Étape au cours de laquelle on formule les caractéristiques des items ou des exercices à produire (consigne, stimulus, réponse).

La rédaction des items, des exercices ou des tâches

Les items, les exercices et les tâches doivent être rédigés conformément aux critères retenus aux autres étapes.

Le montage de l'épreuve

Étape au cours de laquelle on assemble et structure les items, les exercices ou les tâches, on produit aussi le guide d'utilisation, le cahier de l'élève, les consignes aux élèves, la clef de correction.

La validation de contenu et la révision docimologique

On devrait ici consulter deux types de spécialistes pour connaître d'une part, la congruence entre les items, les exercices ou les tâches produits et les dimensions à mesurer, d'autre part, le respect des exigences docimologiques

L'expérimentation

Il s'agit de vérifier la pertinence du matériel d'évaluation produit

La révision de l'instrumentation

On modifie le matériel d'évaluation, s'il y a lieu, à la lumière des recommandations ou des suggestions obtenues lors de l'expérimentation

Particularité de l'instrumentation utilisée pour réaliser une évaluation sommative

On sait que l'évaluation sommative est une démarche qui consiste à porter un jugement sur le degré d'atteinte des objectifs énoncés dans un programme, en se basant sur des données pertinentes, pour prendre des décisions relatives à la sanction des acquis scolaires ou extrascolaires. L'instrumentation utilisée pour évaluer chaque discipline doit respecter le tableau de spécification élaboré à partir du programme d'études en usage. Un tel tableau de spécification devrait contenir les éléments suivants

- les contenus notionnels, avec leur pondération,
- les habiletés intellectuelles reliées à un niveau taxonomique, avec leur pondération,
- les liens entre les éléments de contenus et les habiletés intellectuelles, c'est-à-dire les dimensions à mesurer

Nous traiterons à l'annexe « E » des concepts reliés au tableau de spécification qui devrait être préparé avant la rédaction d'un examen

Annexe « E » La définition de domaine / la préparation du tableau de spécification

Pour garantir la validité de contenu d'une épreuve, il faut, notamment, s'assurer que les questions qui la composent sont représentatives des apprentissages visés dans le programme. Le tableau de spécification permet de s'assurer de la représentativité des apprentissages mesurés dans une épreuve et de représenter les résultats sous forme de synthèse.

En outre, le tableau de spécification permet de vérifier la congruence entre le contenu d'une épreuve et les apprentissages devant être réalisés dans le cadre d'un programme. Si, dans un cours, on demande aux élèves de résoudre des problèmes ou d'appliquer des lois, l'épreuve doit mesurer toutes ces habiletés et non seulement celles qui portent sur l'acquisition de connaissances (c'est-à-dire, la mémorisation).

Analyser le programme en vue de construire un tableau de spécification

Pour bâtir un tableau de spécification, il faut d'abord analyser le programme ou la discipline. Cette analyse a pour but de déterminer quels sont les apprentissages qui feront objet de mesure et quelle est leur importance relative, c'est-à-dire leur pondération.

Il s'agit d'analyser les éléments constituant le programme. Toutefois, il ne faudrait pas se limiter aux composantes liées uniquement aux contenus notionnels. Il faut également déterminer quels sont les processus mentaux ou les habiletés intellectuelles que les élèves doivent exercer pour traiter ces contenus notionnels : doivent-ils les mémoriser ou en faire l'analyse? Doivent-ils appliquer des règles, résoudre des problèmes, évaluer des situations?

De plus, idéalement, il faut tenir compte des autres éléments qui constituent un programme, comme les caractéristiques de l'approche disciplinaire ou les buts de la formation. Ces éléments influent sur le format de l'épreuve ou des items.

L'analyse du programme se fait en fonction du but de l'épreuve et en sélectionnant, parmi les éléments repertoriés, ceux qui semblent être les plus significatifs, représentatifs et essentiels et qui feront l'objet de la mesure dans l'épreuve.

Il faut d'abord examiner les contenus notionnels (les notions, les concepts, les thèmes) et, ensuite, les habiletés intellectuelles que les élèves doivent exercer pour traiter ces contenus notionnels. De plus, il faut déterminer l'importance relative ou la pondération de ces éléments.

Les contenus notionnels

En fonction du but visé, les contenus notionnels devraient être analysés, sélectionnés puis regroupés en sous-ensembles représentatifs en vue d'établir une liste des notions à mesurer

Afin d'éviter de poser trop de questions sur un aspect secondaire de la discipline ou encore de négliger des aspects importants, il faut retenir ce qui est essentiel et représentatif du programme

Les contenus notionnels retenus et organisés doivent être ni trop généraux ni trop spécifiques. S'ils sont trop globaux, ils conduiront à des interprétations multiples, s'ils sont trop spécifiques, ils obligeront les concepteurs de l'épreuve à en retenir une trop grande quantité

Lorsqu'il s'agit d'une épreuve synthèse qui vise à dresser un bilan des apprentissages en fin de session, on privilégie les éléments englobants qui correspondent souvent mais pas toujours aux objectifs terminaux ou à un regroupement d'objectifs terminaux. Par contre, en cours d'année, si l'on souhaite élaborer une épreuve pour repérer les difficultés des élèves, on s'intéresse plus en détails aux notions ou aux étapes de la démarche intellectuelle

Les habiletés intellectuelles

Les processus mentaux ou les habiletés intellectuelles qui réfèrent à la discipline dont les élèves doivent aborder les contenus notionnels, doivent être déterminées et explicitées. Il s'agit de définir quels processus cognitifs les élèves doivent mettre en branle pour réaliser les apprentissages prévus dans un programme donné

À chaque habileté devrait correspondre une liste de manifestations ou de comportements observables. Les habiletés intellectuelles devraient se référer à des taxonomies. Si celles-ci ne sont pas présentes dans le programme, il faut en choisir une ou en déterminer une appropriée à la discipline

La pondération des contenus notionnels et des habiletés intellectuelles

Il faut déterminer la pondération ou l'importance relative des contenus notionnels et des habiletés intellectuelles. Le poids ainsi alloué sert à déterminer l'importance de chacun des contenus et de chacune des habiletés à l'intérieur de l'épreuve et le nombre de questions à concevoir pour chacun d'eux

L'importance relative des éléments du programme a une incidence directe sur l'importance qu'aura chaque question par rapport à l'ensemble de l'épreuve. Si on accorde le même nombre de points à toutes les questions, on retrouve un nombre proportionnel de questions correspondant aux pondérations

La pondération des éléments du programme peut déjà être indiquée dans le programme. Sinon, il faut la déterminer avec l'aide d'experts de contenu. Dans un cas comme dans l'autre, il peut être nécessaire de réviser cette pondération à la lumière du but de l'épreuve. C'est fondamentalement une question de consensus entre experts.

À la fin de l'étape qui consiste à analyser le programme, on devrait avoir une idée des éléments à mesurer dans une épreuve : quels sont les contenus notionnels à vérifier et comment les élèves doivent les traiter. Cette étape conduit à l'élaboration d'un tableau de spécification.

Constituer le tableau de spécification

Le tableau de spécification est un élément de planification qui permet de structurer les données pertinentes pour l'élaboration d'une épreuve. C'est, en quelque sorte, un plan d'examen en ce sens qu'il détermine l'importance relative des contenus notionnels et des niveaux de fonctionnement cognitif attachés à chaque élément de contenu et qui doivent faire objet de mesure dans une épreuve. Le tableau de spécification constitue un moyen efficace pour préciser l'échantillon de questions à inclure dans une épreuve.

Plus précisément, le tableau de spécification est un tableau ordinairement à double entrée : une pour la liste des contenus notionnels à maîtriser, l'autre pour la liste des habiletés intellectuelles à mesurer. Il indique aussi l'importance relative ou la pondération des contenus notionnels et des habiletés intellectuelles.

Le tableau de spécification illustre également les recoupements entre les contenus notionnels et les habiletés intellectuelles. On appelle « dimension » l'intersection entre une habileté et un contenu. Chaque dimension précise l'objet de mesure à partir duquel les questions seront rédigées.

Il est possible qu'une case d'un tableau de spécification soit vide, c'est-à-dire que l'intersection entre une habileté et une notion n'existe pas en fonction de la discipline à vérifier. Il s'agit donc d'une dimension qui ne fera pas l'objet de mesure dans l'épreuve.

L'utilisation du tableau de spécification lors de l'élaboration d'une épreuve permet de s'assurer du lien entre l'épreuve et le programme. En effet, si l'on rédige des questions selon les buts du programme en couvrant l'ensemble des dimensions du tableau de spécification tout en respectant les pondérations qui y sont indiquées, on s'assure que l'épreuve sera le portrait fidèle du programme. Le tableau de spécification constitue un ensemble de données à partir duquel plusieurs épreuves, portant sur un même programme, peuvent être élaborées.

En outre, le tableau de spécification est d'une grande utilité au moment de l'interprétation des résultats. Il permet de repérer les forces et les faiblesses d'un élève ou de l'ensemble des élèves qui ont subi l'épreuve. En effet, comme on connaît à quelle dimension se réfère chaque question de l'épreuve, on peut déterminer, à partir des questions échouées et des questions réussies, quels sont les contenus ou les habiletés intellectuelles qui ont posé problème ou lesquelles sont les mieux maîtrisées.

Annexe « F » La correction des items à réponses construites

Il arrive fréquemment que les réponses fournies par les élèves à une épreuve doivent être corrigées par du personnel enseignant. Il importe alors d'assurer la rigueur d'une telle opération si l'on veut garantir l'équité des résultats obtenus par chaque élève. Quelques mécanismes de contrôle de la qualité de la correction devraient être mis en place à cet effet. Au préalable, il aura fallu prévoir les critères de correction devant être utilisés. Des activités de formation destinées au personnel responsable de la correction devraient aussi être prévues.

Détermination des critères de correction

Au moment de l'élaboration d'une épreuve, il est nécessaire de prévoir la clé de correction, c'est-à-dire l'identification des réponses exactes. S'il s'agit d'un item dont la réponse fournie par les élèves se présente sous la forme d'un texte, il est important de prévoir la ou les réponses admissibles qui respectent les critères fixes par les concepteurs de l'instrumentation. Si l'épreuve a été expérimentée, il sera plus facile de disposer d'un éventail de réponses admissibles.

Formation à la correction

Les personnes assignées à la correction d'une épreuve devraient recevoir une information la plus complète possible sur les caractéristiques de l'épreuve objet de la correction ainsi que sur les critères de correction. L'examen de copies-types peut s'avérer une bonne stratégie pour identifier les réponses possibles fournies par les élèves. Au préalable, une sélection de telles copies devra avoir été faite dans le but d'identifier l'éventail le plus complet possible de réponses fournies par les élèves aux différentes questions de l'épreuve.

Des exercices de correction de la même copie par l'ensemble des correcteurs peuvent s'avérer une stratégie utile pour confirmer une compréhension univoque des critères de correction et détecter les correcteurs « déviants ».

Contrôle de la qualité de la correction

Plusieurs stratégies peuvent être retenues pour exercer le contrôle de la qualité de la correction. Nous en présentons quelques-uns ci-dessous.

Corrections de groupe

Comme nous l'indiquons plus haut, il est possible de réaliser la correction de la même copie par l'ensemble des correcteurs. On peut retenir cette stratégie au moment de la formation du personnel à la correction. Cependant, il peut être utile de faire de tels exercices de corrections de groupe de façon périodique afin de vérifier si l'ensemble des correcteurs demeurent constants dans leurs corrections. Il sera aussi possible d'identifier certains correcteurs ayant besoin d'un complément de formation dans le but de mieux respecter les critères de correction fixes au départ.

Double correction

On sait que tout mécanisme d'évaluation comporte des « erreurs de mesure ». Il importe de les minimiser surtout lorsque les résultats obtenus sont importants pour les personnes évaluées. La connaissance de la « marge d'erreur » permettra de mieux fixer les critères de succès rattachés aux résultats obtenus. La technique de la « double correction » d'un échantillon de copies permet de disposer de renseignements sur l'importance de cette « marge d'erreur ».

Identification des « deviances »

Il peut être possible d'identifier les correcteurs ou groupes de correcteurs dont la correction s'écarte de façon significative des critères fixés. Dans un tel cas, il sera possible de reprendre la correction de certaines copies pour vérifier si les résultats obtenus respectent les exigences définies. Une telle opération est plus facile à exécuter si la correction est réalisée centralement. On doit cependant disposer de moyens techniques permettant de disposer rapidement des résultats obtenus à la suite de la correction. Il est enfin souhaitable que les copies à corriger aient été assignées aux correcteurs de façon aléatoire.

Annexe G Le contrôle de la qualité

Voici un sommaire des activités de contrôle de la qualité d'une évaluation par rapport aux principales étapes de réalisation d'une telle opération

1- Planification de l'opération d'évaluation et préparation de l'instrumentation	<ul style="list-style-type: none"> • préparation du devis • préparation du devis d'échantillonnage (si pertinent) • expérimentation • révision des procédures et des instruments • traduction à rebours (si plusieurs langues) • contrôle de la qualité de présentation des instruments de collecte de données
2- Administration de l'instrumentation	<ul style="list-style-type: none"> • contrôle de l'administration (par des visites sur place par exemple)
3 Notation et codification des informations	<ul style="list-style-type: none"> • définition des règles de codification et de notation • formation du personnel à la codification et à la notation • double correction (s'il y a notation) • correction multiple (s'il y a notation)
4 Préparation des bases de données	<ul style="list-style-type: none"> • choix du type de saisie • comparaison avec un fichier échantillonné • fabrication de distributions de fréquences en relation avec les valeurs valides • examen visuel des fichiers (comparaison entre les données consignées dans le formulaire de collecte et le fichier de données)
5- Réalisation des analyses statistiques	<ul style="list-style-type: none"> • analyse classique des items (s'il s'agit d'une épreuve) • analyse de biais (technique DIF) • analyse à l'aide de la théorie des réponses aux items (s'il s'agit d'une épreuve) • examen de l'unidimensionalité de l'épreuve utilisée • utilisation de Bilog pour calcul des paramètres d'items permettant de mettre à niveau les items, calculer des scores d'habileté, d'ancrer les items
6- Rédaction et fabrication des rapports	<ul style="list-style-type: none"> • identification des publics cibles • compréhension par les rédacteurs des résultats des analyses statistiques • établissement de liens entre différentes données • consultation de spécialistes du domaine pour aider à l'interprétation des résultats • utilisation de techniques diverses pour présenter les résultats (tableaux, schémas graphiques)