Dépôt Institutionnel de l'Université libre de Bruxelles /
Université libre de Bruxelles Institutional Repository
Thèse de doctorat/ PhD Thesis

Citation APA:

Lebrun, B. (1977). *Une évaluation informative* (Unpublished doctoral dissertation). Université libre de Bruxelles, Faculté des Sciences psychologiques et de l'éducation, Bruxelles.

Disponible à / Available at permalink : https://dipot.ulb.ac.be/dspace/bitstream/2013/214257/6/562dd7ef-bd9d-4264-bded-8fcab3d15952.txt

(English version below)

Cette thèse de doctorat a été numérisée par l'Université libre de Bruxelles. L'auteur qui s'opposerait à sa mise en ligne dans DI-fusion est invité à prendre contact avec l'Université (di-fusion@ulb.be).

Dans le cas où une version électronique native de la thèse existe, l'Université ne peut garantir que la présente version numérisée soit identique à la version électronique native, ni qu'elle soit la version officielle définitive de la thèse.

DI-fusion, le Dépôt Institutionnel de l'Université libre de Bruxelles, recueille la production scientifique de l'Université, mise à disposition en libre accès autant que possible. Les œuvres accessibles dans DI-fusion sont protégées par la législation belge relative aux droits d'auteur et aux droits voisins. Toute personne peut, sans avoir à demander l'autorisation de l'auteur ou de l'ayant-droit, à des fins d'usage privé ou à des fins d'illustration de l'enseignement ou de recherche scientifique, dans la mesure justifiée par le but non lucratif poursuivi, lire, télécharger ou reproduire sur papier ou sur tout autre support, les articles ou des fragments d'autres œuvres, disponibles dans DI-fusion, pour autant que :

- Le nom des auteurs, le titre et la référence bibliographique complète soient cités;
- L'identifiant unique attribué aux métadonnées dans DI-fusion (permalink) soit indiqué;
- Le contenu ne soit pas modifié.

L'œuvre ne peut être stockée dans une autre base de données dans le but d'y donner accès ; l'identifiant unique (permalink) indiqué ci-dessus doit toujours être utilisé pour donner accès à l'œuvre. Toute autre utilisation non mentionnée ci-dessus nécessite l'autorisation de l'auteur de l'œuvre ou de l'ayant droit.

---------- **English Version** ----------

This Ph.D. thesis has been digitized by Université libre de Bruxelles. The author who would disagree on its online availability in DI-fusion is invited to contact the University (di-fusion@ulb.be).

If a native electronic version of the thesis exists, the University can guarantee neither that the present digitized version is identical to the native electronic version, nor that it is the definitive official version of the thesis.

DI-fusion is the Institutional Repository of Université libre de Bruxelles; it collects the research output of the University, available on open access as much as possible. The works included in DI-fusion are protected by the Belgian legislation relating to authors' rights and neighbouring rights. Any user may, without prior permission from the authors or copyright owners, for private usage or for educational or scientific research purposes, to the extent justified by the non-profit activity, read, download or reproduce on paper or on any other media, the articles or fragments of other works, available in DI-fusion, provided:

- The authors, title and full bibliographic details are credited in any copy;
- The unique identifier (permalink) for the original metadata page in DI-fusion is indicated;
- The content is not changed in any way.

It is not permitted to store the work in another database in order to provide access to it; the unique identifier (permalink) indicated above must always be used to provide access to the work. Any other use not mentioned above requires the authors' or copyright owners' permission.

UNIVERSITE LIBRE DE BRUXELLES
Faculté des Sciences Psychologiques et Pédagogiques

EVALUATION INFORMATIVE

ANNEXE 5

Utilisation des méthodes d'évaluation informative
dans des recherches expérimentales

Bernard LEBRUN

UNIVERSITAS BRUXELLENSIS

OCTOBRE 1977

H354

UNIVERSITE LIBRE DE BRUXELLES
Faculté des Sciences Psychologiques et Pédagogiques

EVALUATION INFORMATIVE

****

ANNEXE 5

Utilisation des méthodes d'évaluation informative
dans des recherches expérimentales

695.760
[v.6]
cp1

Bernard LEBRUN

BIBLIOTHEQUE UNIVERSITAIRE BRUXELLES

OCTOBRE 1977

289 1995

INTRODUCTION

Parallèlement à nos recherches, des expérimentateurs ont exploité les techniques que nous mettons au point pour traiter de problèmes particuliers. En utilisant l'évaluation informative, ces chercheurs visent la mise en évidence d'effets significatifs en vue de dégager des conclusions dans les travaux qui leur sont propres.

Nous retiendrons trois de ces recherches.

1° Celles de BAMBA, concrétisées par une dissertation doctorale intitulée : "Le recours à une langue vernaculaire comme moyen d'enseignement à l'Ecole Primaire au Zaïre" déposée à l'Université Libre de Bruxelles en novembre 1976.

2° Celles de Marcel FRYDMAN, dont la dissertation doctorale défendue à l'Université de l'Etat à Mons en 1976 a pour objet : "Le développement de l'aptitude à exploiter un matériel de références".

3° Celles de Samuel AMEGAN, en cours à l'Université du Québec, portant sur les interventions pédagogiques susceptibles d'entraîner certains comportements créatifs.
Michel DEPASSE a analysé les données de S. AMEGAN par les techniques d'évaluation informative.
Il a présenté ses résultats dans un mémoire intitulé : "Traitement de données relatives aux effets de deux systèmes éducatifs. Utilisation de l'évaluation informative "défendu en juin 1977 à la Faculté des Sciences Psychologiques et Pédagogiques de l'U.L.B.

Il ne nous appartient pas à propos de ces travaux, de discuter les hypothèses émises par leurs auteurs, mais bien de nous assurer dans quelle mesure les techniques d'évaluation informative concourent à leur vérification. C'est l'aspect méthodologique beaucoup plus que l'aspect technique qui nous importe.

478170

1° RECHERCHES DE BAMBA

Ce travail étudie l'impact de la langue d'enseignement sur le rendement scolaire en première année primaire au Zaïre. La langue maternelle des enfants testés étant le lingala, il s'agit de rechercher si l'utilisation du lingala comme langue d'enseignement donne de meilleurs résultats que l'emploi du français.

L'hypothèse de travail se présente donc sous un double aspect (1)

" 1. Nous supposons que l'enfant peut être indifféremment enseigné soit en français, soit en lingala. Dans ce cas, nous supposons qu'il ne s'observe pas de différence significative entre les deux méthodes d'enseignement.

2. Au cas où les résultats obtenus par les enfants répartis entre les deux principaux groupes expérimentaux (francophone et lingalaphone) présentent des différences et n'admettent pas, par voie de conséquence, l'équivalence des moyennes, nous admettrons que l'hypothèse nulle (d'équivalence) posée en premier lieu est impertinente. Dès lors, nous ne pourrons que la rejeter purement et simplement et serons obligé d'admettre, l'existence de facteurs favorables ou défavorables liés aux circonstances expérimentales définies : la langue d'enseignement."

---

(1) BAMBA
"Le recours à une langue vernaculaire comme moyen d'enseignement à l'Ecole primaire au Zaïre". U.L.B. - 1976 - P.166

Le plan expérimental prévoit la comparaison des résultats de 10 classes prélevées de manière occasionnelle dans l'agglomération de Kinshasa. Ces classes se répartissent comme suit :

- 5 classes lingalaphones (n°2,3,7,8,9)
- 4 classes francophones (n°1,4,6,10)
- 1 classe bilingue (n°5)

A. Etude de l'homogénéité initiale du groupe expérimental.

Il s'agit de vérifier si les classes étudiées sont équivalentes. Sur la base d'une épreuve d'entrée de maîtrise de vocabulaire, BAMBA a comparé les résultats d'ensemble des différentes classes et sous-groupes lingala et français constitués a posteriori selon les techniques d'évaluation informative "arc sinus" et "XHI2".

L'épreuve d'entrée compte 16 items. L'expérimentateur prononce 8 mots en lingala et 8 mots en français de manière alternée. Sur une feuille comportant cinq dessins, chaque élève coche celui qui correspond au mot prononcé.

1. Technique XHI2

a) Vérification de l'homogénéité de l'ensemble de population

	N	Σx	$\bar{x}$	s	$\left(\frac{\bar{X} - \bar{x}}{s/\sqrt{N}}\right)^2$
C 01	40	463	11,5750	1,7900	0,0560
C 02	30	335	11,1666	3,0568	0,7276
C 03	40	483	12,0750	3,0160	0,8222
C 04	60	727	12,1166	2,4776	2,1966
C 05	30	347	11,5666	2,0791	0,0400
C 06	44	526	11,9545	2,2303	0,8611
C 07	37	440	11,8918	1,8971	0,6400
C 08	40	450	11,2500	2,3938	1,0764
C 09	30	351	11,7000	2,0367	0,0238
C 10	30	334	11,1333	2,6487	1,1094
	381	4456	$\bar{X}$ = 11,6426		$\Sigma\chi^2$ = 7,5331

XHI² calculée : 7,5331

XHI² tabulée : 18,307 à 10 dl

Ces dix classes sont donc considérées commme présentant les caractéristiques d'un échantillon aléatoire simple.

b) Etude de l'homogénéité des sous-groupes

1. Sous-groupe lingalaphone

	N	Σx	$\bar{x}$	s	$\left(\frac{\bar{X} - \bar{x}}{s/\sqrt{N}}\right)^2$
C 02	30	335	11,1666	3,0569	0,6501
C 03	40	483	12,0750	3,0160	0,9242
C 07	37	440	11,8918	1,8976	0,7784
C 08	40	450	11,2500	2,3938	0,9385
C 09	30	351	11,7000	2,0367	0,0503
	177	2059	$\bar{X}$ = 11,6166		$\Sigma\chi^2$ = 3,3415

XHI² calculée : 3,3415

XHI² tabulée : 11,070 à 5 dl.

2. Sous-groupe francophone

	N	Σx	$\bar{x}$	s	$\left(\frac{\bar{X} - \bar{x}}{s/\sqrt{N}}\right)^2$
C 01	40	463	11,5750	1,7900	0,1791
C 04	60	727	12,1166	2,4776	1,7394
C 06	44	526	11,9545	2,2303	0,5966
C 10	30	334	11,1333	2,6487	1,3486
	174	2050	$\bar{X}$ = 11,6948		$\Sigma\chi^2$ = 3,8637

2. Technique arc sinus

Les résultats sont synthétisés dans le tableau ci-dessous. Les items sont en abscisse (F: français , L: lingala) et les classes en ordonnée. Une flèche vers le haut indique que la classe est supérieure à l'intervalle pour cet item et une flèche vers le bas qu'elle est inférieure.

	L	F	L	F	L	F	L	F	L	F	L	F	L	F	L	F	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
1			↑			↑				↓							3
2									↓	↓					↑		3
3								↑							↑		2
4						↓										↓	2
5			↑									↓		↓			3
6																	0
7										↑					↓		2
8		↓			↓												2
9			↑										↓			↓	3
10		↓		↓	↓			↓		↑						↓	6

La proportion de rejets est très faible (16%). Elle se répartit uniformément sur les classes et sur les items, ce qui confirme l'homogénéité constatée par l'application de la technique XHI². Seule la classe 10 (francophone) est faible pour 5 des 16 items (dont 4 items français).

3. Technique XHI² classique.

L'application de la technique XHI² d'indépendance sur ces données a permis à BAMBA de confirmer les informations recueillies ci-avant.

a) pour la comparaison des proportions de réussite dans les deux sous-groupes vis-à-vis des items prononcés en lingala d'une part, et des items prononcés en français d'autre part.

b) pour la comparaison des classes en regroupant les notes selon les critères ci-après :

1. notes de 00 à 06
2. note 07
3. note 08
4. notes de 09 à 16.

Toutes les classes sont homogènes, que l'on étudie la population totale ou les sous-groupes.

c) pour vérifier si les "maintiens" ou les "exclusions" observés par la technique arc sinus différent en fonction de la nature linguistique des items.
Les items "français" et les items "lingala" sont également stables.

B. Etude des épreuves expérimentales (1)

BAMBA a construit des épreuves spécialement adaptées au niveau des élèves et aux objectifs de son expérimentation.
"Eu égard au niveau scolaire des enfants, nous ne pouvions guère apprécier que certaines manifestations de maîtrise de la langue orale, quelques déchiffrements élémentaires de lecture et les premiers rudiments du calcul mental " (2)

1. Contenu des épreuves.

a) épreuves de vocabulaire (N° 1 et 2)
Il s'agit d'une identification d'images à partir d'une description orale.

b) épreuve de lecture (N° 3)
Les élèves sont appelés à reconnaître huit éléments (lettres ou mots en deux syllabes au maximum)

(1) Cette étude a été réalisée avec la collaboration de Madame CHAUVAUX au cours d'un stage de didactique expérimentale.

(2) BAMBA - idem page 175

c) épreuve de calcul (N° 4)

Il s'agit de cinq additions et de cinq soustractions de nombres inférieurs à 10.

d) épreuve de langage oral (N° 5, 6 et 7)

Cette épreuve s'évalue selon plusieurs dimensions : la production verbale, la correction syntaxique et le maniement oral de la langue.

Bamba effectue trois observations à partir d'une même production verbale :

- un inventaire du vocabulaire (N° 5),
- un indice de la correction syntaxique (N° 6),
- un indice de maniement oral de la langue combinant les épreuves 5 et 6 (N° 7).

e) épreuve de compréhension verbale (N° 8)

L'enfant doit distinguer, parmi cinq images, celle qui correspond à la séquence décrite par l'examinateur.

2. Analyse des résultats.

a) Epreuves de vocabulaire

Liste des items du test n° 1

Groupe français	Groupe lingala
1. une jambe	1. lokolo
2. une tasse (ce qu'on emploie pour boire du thé)	2. kopo bamelilaka ti.
3. un pot de fleurs	3. kopo ya fulele.
4. ce qui sert à labourer de la terre	4. eloko balonaka elanga, to mabele.
5. un homme en train de fumer une cigarette.	5. moto azali komela likaya.
6. un enfant endormi dans un lit.	6. muana alali na mbetu.
7. un enfant en train de courir	7. muana azali kokima mbangu.
8. un homme en train de cueillir des mangues.	8. moto azali kobuka manga.
9. ce qu'on emploie pour boire de l'eau.	9. eloko bamelilaka mayi.
10. ce qui sert à puiser de l'eau.	10. eloko batokilaka mayi.

- Liste des items du test n° 2

Groupe français	Groupe lingala
1. Un enfant qui porte un seau sur la tête.	1. muana moko amemi katini na moto.
2. Une femme qui se promène avec son enfant.	2. muasi moko akokende na muana naye.
3. Un enfant en train de boire de l'eau, avec un verre.	3. muana moko akomela mayi na kopo.
4. Deux enfants en train de jouer au football.	4. Bana mibale bazali kobeta ndembo.
5. Un poisson.	5. Mbisi ya mayi.
6. Une femme avec un bébé au dos.	6. Muasi moko amemi muana na mokongo.
7. Une femme en train de piler du manioc	7. Muasi moko akotuta fufu.
8. Une femme en train de cultiver de la terre.	8. Muasi moko akolona elanga.
9. Ce qui sert à tuer des animaux	9. Eloko babominaka niama.
10. Ce qui sert à écrire.	10. Eloko bakomelaka mokanda.

Les deux premiers tests sont des épreuves de vocabulaire. Ils comprennent chacun 10 items. L'enfant doit identifier, parmi 5 images, celle qui correspond à ce que lui désigne oralement l'examinateur. Il s'agit tantôt de l'objet désigné par son nom, tantôt de l'objet désigné par un de ses usages, tantôt d'un état, tantôt d'une action.

Comportement de l'ensemble de la population et des différents sous-groupes.

Dans les tableaux de synthèse ci-après, nous utilisons le code suivant :

Colonnes : - % = les % de réussites dans le groupe considéré.
- IDS = indice de différenciation spécifique
- T = type de question
  - D : de différenciation
  - I : d'interaction
  - inf : d'enquête inférieure
  - S : d'enquête supérieure

	Groupe total			Sous-groupe 1			Sous-groupe 2		
Test 1	n = 377			n = 171			n = 176		
Item	%	IDS	T	%	IDS	T	%	IDS	
1	.71	.53	D	.51	.42	I	.95		S
2	.85		S	.79		S	.92		S
3	.88		S	.86		S	.91		S
4	.58	.61	D	.32	.25	I	.89		S
5	.53	.61	D	.23		inf	.87		S
6	.70	.54	D	.50	.56	D	.96		S
7	.49	.63	D	.23		inf	.83		S
8	.72	.45	I	.60	.39	I	.92		S
9	.48	.65	D	.20		inf	.81		S
10	.54	.46	I	.37	.43	I	.72		I
Test 2	n = 367			n = 169			n = 168		
Item	%	IDS	T	%	IDS	T	%	IDS	T
1	.63	.57	D	.39	.18	I	.89		S
2	.39	.56	D	.14		inf	.69		I
3	.62	.63	D	.36	.60	D	.92		S
4	.71	.50	D	.52	.61	D	.93		S
5	.88		S	.78		S	.98		S
6	.57	.72	D	.30	.24	I	.91		S
7	.77		S	.60	.51	D	.96		S
8	.61	.66	D	.37	.56	D	.93		S
9	.58	.72	D	.24	.22	I	.97		S
10	.53	.69	D	.23		inf	.90		S

Le sous-groupe 1 (francophone) a un taux de réussite variable selon les items alors que le sous-groupe 2 (langalaphone) est remarquablement compétent.
Seuls trois items (test 1, items 2 et 3 - test 2, item 5) ne distinguent pas les sous-groupes.
Ils concernent des objets familiers (la tasse, le pot de fleur, un poisson).
La connaissance d'éléments isolés de ce type est donc acquise aussi bien en lingala qu'en français. Il n'en va plus de même lorsqu'interviennent des relations linguistiques plus complexes.
La différenciation du groupe total (14 items sur les 20!) s'effectue entre les deux sous-groupes.

Comportement des classes séparées.

	1	4	6	10	2	3	7	8	9	5
Test 1										
Item										
1	I	I	D	S	S	S	S	S	S	I
2	S	S	S	I	S	S	S	S	S	I
3	S	I	S	S	S	I	S	S	S	S
4	inf	inf	S	inf	S	S	S	S	S	I
5	inf	I	inf	I	S	S	S	S	S	I
6	I	I	D	I	S	S	S	S	S	I
7	inf	inf	I	inf	S	I	S	S	D	inf
8	I	I	S	I	S	S	S	S	D	I
9	I	inf	inf	I	I	D	S	S	S	inf
10	I	I	D	I	S	D	S	I	S	I
Test 2										
Item										
1	I	I	inf	D	S	S	S	S	S	I
2	I	inf	inf	inf	I	I	I	I	I	inf
3	D	inf	S	inf	S	S	S	S	S	I
4	I	I	S	I	S	S	S	S	S	D
5	S	D	S	I	S	S	S	S	S	S
6	D	D	inf	D	S	S	S	S	S	I
7	S	D	S	inf	S	S	S	S	S	D
8	D	inf	S	inf	S	S	S	S	S	inf
9	inf	I	inf	D	S	S	S	S	S	D
10	inf	inf	I	inf	S	S	S	S	S	inf

Les classes du sous-groupe lingalaphone réussissent massivement les deux tests. Seul l'item 2 du test 2 est moins bon. Nous mettons en cause la qualité du dessin proposé.
La situation dans les classes du sous-groupe francophone est bien moins homogène, avec un nombre très important d'items d'interaction pour lesquels on peut soupçonner le devinement puisque la forme de réponse est un choix multiple.
Le comportement de la classe 6 est spécifique :
elle réussit très bien 9 items et très mal 4 items.

Etude des résultats par la XHI² classique

BAMBA applique le test XHI² d'indépendance entre les résultats observés dans les deux sous-groupes.
D'après lui, tous les items, à l'exception de l'item 3, différencient les deux sous-groupes à l'avantage du sous-groupe lingalaphone.

b) Epreuve de lecture.

Liste des items du test n° 3

1. u o a i e
2. tate tati titi tato
3. pape pepa papa papi
4. pato tape pipe pate
5. mame même mimi mimo
6. mapi mepa mipi pema
7. lolo lala lili léa
8. malo mili mele mila

L'enfant doit identifier le mot prononcé par l'examinateur (souligné ci-dessus) sur une liste de 4 mots. La plupart de ces mots n'ont pas de signification. Le premier item est composé exclusivement de voyelles.

Comportement de l'ensemble de la population et des différents sous-groupes.

Item	Groupe total n = 373 %	IDS	T	Sous-groupe 2 n = 174 %	IDS	T	Sous-groupe 1 n = 169 %	IDS	T
1	.81		S	.79		S	.83		S
2	.53	.56	D	.63	.64	D	.45	.50	D
3	.63	.69	D	.66	.56	D	.59	.72	D
4	.60	.67	D	.62	.64	D	.59	.69	D
5	.44	.51	D	.46	.50	D	.43	.60	D
6	.35	.51	D	.37	.61	D	.31	.54	D
7	.03		inf	.04		inf	.02		inf
8	.43	.52	D	.45	.65	D	.43	.46	D

6 items sur 8 sont réussis par une proportion d'élèves assimilable à 50% et permettent une différenciation, aussi bien dans le groupe total que dans chacun des sous-groupes. La langue d'enseignement n'est donc pas en cause ici, mais la didactique qui conduit à une différenciation systématique au niveau de la technique de déchiffrement.

L'item 1, d'enquête supérieure partout, montre que l'élément "voyelle" est assimilé.

Quant à l'item 7, inférieur partout, nous suspectons la très grande ressemblance euphonique entre "lala", terme à identifier, et "lea", figurant parmi les autres termes.

Comportement des classes séparées.

Item	1	4	6	10	2	3	7	8	9	5
1	S	S	I	S	S	D	S	I	S	S
2	I	I	D	D	D	D	D	D	D	D
3	D	S	D	D	S	D	S	D	D	I
4	D	S	D	D	D	D	D	D	S	D
5	D	D	I	D	D	I	D	I	D	I
6	inf	D	I	inf	D	D	I	I	D	I
7	inf	inf	inf	inf	inf	inf	inf	inf	inf	inf
8	D	I	D	D	D	I	I	inf	D	I

Sur les 6 items différenciant le groupe total, nous retrouvons 38 différenciations sur 60 possibles dans les 10 classes. Les variations individuelles sont donc peu nombreuses.
Nous retrouvons, au sein des classes, les mêmes constatations que plus haut à propos de l'item 1 (7 fois supérieur sur 20) et de l'item 7 (toujours inférieur).
Nous mettons donc en cause les effets qu'a eu le traitement linguistique des élèves sur l'aptitude sollicitée ici.

Etude des résultats par la XHI² classique.
Toutes les comparaisons sont non significatives, sauf pour l'item 2 (?)

c) Epreuve de calcul.

Liste des items du test n° 4

1 : 1 + 1 =

2 : 2 + 1 =

3 : 3 + 1 =

4 : 5 + 1 =

5 : 6 + 1 =

6 : 2 - 1 =

7 : 3 - 1 =

8 : 5 - 1 =

9 : 6 - 1 =

10 : 8 - 2 =

Comportement de l'ensemble de la population et des différents sous-groupes

	Groupe total			Sous-groupe 1			Sous-groupe 2		
	n = 351			n = 156			n = 172		
Item	%	IDS	T	%	IDS	T	%	IDS	T
1	.86		S	.90		S	.83		S
2	.76	.37	I	.76	.40	I	.76	.39	I
3	.73	.41	I	.69	.52	D	.74	.35	I
4	.67	.50	D	.67	.54	D	.66	.54	D
5	.63	.61	D	.62	.60	D	.63	.69	D
6	.55	.75	D	.53	.71	D	.55	.74	D
7	.46	.89	D	.44	.84	D	.49	.87	D
8	.49	.87	D	.46	.74	D	.52	.90	D
9	.39	.77	D	.36	.71	D	.42	.79	D
10	.41	.69	D	.42	.65	D	.42	.75	D

Les 7 derniers items introduisent une différenciation au sein du groupe total et des deux sous-groupes.
L'item 1 est connu d'une majorité des élèves (enquête supérieure).
Les items 2 et 3 sont presque toujours en interaction. Nous attendons l'analyse par classe pour nous prononcer sur ce phénomène.

Comportement des classes séparées.

Item	1	4	6	10	2	3	7	8	9	5
1	S	S	S	S	S	S	S	S	S	S
2	S	S	D	D	S	I	S	S	D	S
3	S	S	D	I	S	I	S	D	D	S
4	I	D	D	D	S	I	I	D	D	S
5	I	S	D	I	D	D	I	D	D	I
6	D	D	I	D	D	D	D	D	D	D
7	D	D	D	D	D	D	D	D	D	D
8	D	D	I	D	D	D	D	D	D	D
9	D	D	D	D	D	D	D	D	D	D
10	D	D	I	D	D	D	D	D	D	D

Dans certaines classes, les items 2 et 3 sont supérieurs, dans d'autres ils sont de différenciation. C'est donc sur la technique qu'il faut rejeter la responsabilité de cette interaction. Lorsque les pourcentages de réussite s'éloignent de 50% (sans être significativement différents), on rejette parfois à tort des items en interaction.
(Nous nous rapprochons en fait d'items hiérarchisés qui ne sont pas décelés par cette technique).

Etude des résultats par la XHI² classique.
Toutes les comparaisons sont non significatives.
Ici aussi, nous nous interrogeons sur l'effet qu'a pu avoir le traitement linguistique proposé aux enfants sur la qualité de leur performance à ce test.

d) Epreuve de langage oral (n°5,6 et 7)

- Liste des items du test n°5 (production verbale)

Une série de saynètes est présentée aux élèves individuellement Leurs réponses sont enregistrées. Elles sont évaluées à partir de la grille ci-dessous (seuls les mots soulignés sont pris en considération).

1. Mambu va à l'école.
2. Il porte ses cahiers sous le bras.
3. Debout, il lit devant ses amis en classe.
4. La classe terminée il retourne à la maison et salue maman et papa.
5. Il joue avec sa soeur, à la petite voiture.
6. Il parle avec son père.
7. Sa mère fait la vaisselle.
8. Il mange à table, assis sur une chaise, à la maison.
9. Son papa fume la pipe.
10. Sa mère pile du manioc, le bébé au dos.
11. Ils sont ensemble (papa, maman et Mambu) autour du feu, le soir.
12. Il enlève la chemise, dans sa chambre.
13. Il dort dans son lit, la nuit.

Comportement de l'ensemble de la population et des différents sous-groupes.

	Groupe total			Sous-groupe 1			Sous-groupe 2		
	n = 367			n = 163			n = 176		
Item	%	IDS	T	%	IDS	T	%	IDS	T
1	.74	.45	I	.48	.64	D	.96		S
2	.33	.60	D	.01		inf	.64	.33	I
3	.36	.61	D	.05		inf	.67	.24	I
4	.60	.63	D	.24	.39	I	.92		S
5	.50	.76	D	.08		inf	.89		S
6	.37	.68	D	.02		inf	.72	.33	I
7	.38	.73	D	.01		inf	.77	.32	I
8	.55	.76	D	.17		inf	.94		S
9	.67	.56	D	.37	.75	D	.96		S
10	.52	.77	D	.13		inf	.91		S
11	.35	.64	D	.01		inf	.68	.36	I
12	.42	.82	D	.01		inf	.84		S
13	.53	.90	D	.07		inf	.99		S

12 des 13 items différencient les élèves du groupe total. Cette différenciation se concrétise par une situation très différente selon les deux sous-groupes.
10 des 13 items sont d'enquête inférieure pour le sous-groupe francophone, alors que 8 items sont d'enquête supérieure pour le sous-groupe lingalaphone.
Les items à 50% pour le sous-groupe lingalaphone (2,3,6,7 et 11) sont tous inférieurs pour le sous-groupe francophone. Remarquons que ces 5 items sont le siège de phénomènes d'interaction au sein de ce sous-groupe.

Comportement des classes séparées.

Item	1	4	6	10	2	3	7	8	9	5
1	I	D	I	D	S	S	S	S	S	S
2	inf	inf	inf	inf	D	D	I	S	S	I
3	inf	inf	inf	inf	I	I	I	D	I	inf
4	inf	inf	D	D	S	S	S	S	S	I
5	inf	inf	inf	inf	S	S	S	S	S	I
6	inf	inf	inf	inf	S	S	I	D	D	inf
7	inf	inf	inf	inf	D	D	S	S	S	inf
8	inf	inf	I	inf	S	S	S	S	S	I
9	D	D	D	D	S	S	S	S	S	I
10	D	inf	inf	inf	S	I	S	S	S	I
11	inf	inf	inf	inf	S	I	I	D	D	I
12	inf	inf	inf	inf	S	D	S	S	S	inf
13	inf	inf	inf	inf	S	S	S	S	S	I

Cette situation est confirmée par l'analyse des résultats des classes séparées. A part l'item 3, l'interaction observée dans le sous-groupe lingalaphone se situe davantage entre les classes (certaines sont en différenciation, d'autres supérieures) qu'entre les individus.
La classe bilingue révèle 8 items d'interaction : les aptitudes individuelles l'emportent sur la stratégie didactique .

Etude des résultats par la XHI² classique

Toutes les comparaisons sont significatives.

Présentation des items du test n° 6

(correction syntaxique)

L'examinateur considère , dans les réponses des élèves à l'épreuve n° 5, les "mots significatifs", c'est-à-dire les mots-outils, les mots de liaison. La correction syntaxique est évaluée sur la base de la différence entre la syntaxe attendue et la syntaxe observée.

Comportement de l'ensemble de la population et des différents sous-groupes.

	Groupe total n = 367			Sous-groupe 1 n = 163			Sous-groupe 2 n = 176		
Item	%	IDS	T	%	IDS	T	%	IDS	T
1	.74	.51	D	.48	.91	D	.99		S
2	.39	.76	D	.00		inf	.78		S
3	.49	.83	D	.06		inf	.93		S
4	.58	.75	D	.18		inf	.98		S
5	.48	.78	D	.05		inf	.88		S
6	.43	.77	D	.02		inf	.82		S
7	.49	.91	D	.01		inf	.96		S
8	.56	.79	D	.14		inf	.98		S
9	.65	.65	D	.29	.56	D	.99		S
10	.56	.86	D	.12		inf	1.		S
11	.46	.85	D	.01		inf	.89		S
12	.49	.88	D	.04		inf	.94		S
13	.53	.93	D	.04		inf	1.		S

Les 13 items différencient l'ensemble des élèves, et les deux sous-groupes s'opposent : le sous-groupe lingalaphone est particulièrement compétent, alors que le sous-groupe francophone est particulièrement faible.

Comportement des classes séparées

Item	1	4	6	10	2	3	7	8	9	5
1	D	I	D	D	S	S	S	S	S	D
2	inf	inf	inf	inf	S	I	S	S	S	inf
3	inf	inf	inf	inf	S	S	S	S	S	I
4	inf	inf	D	D	S	S	S	S	S	I
5	inf	inf	inf	inf	S	S	I	S	S	D
6	inf	inf	inf	inf	S	S	S	I	S	I
7	inf	inf	inf	inf	S	S	S	S	S	I
8	I	inf	inf	inf	S	S	S	S	S	I
9	D	inf	inf	D	S	S	S	S	S	D
10	I	inf	inf	inf	S	S	S	S	S	D
11	inf	inf	inf	inf	S	S	S	S	S	I
12	inf	inf	inf	inf	S	S	S	S	S	D
13	inf	inf	inf	inf	S	S	S	S	S	I

La même constatation est observable au niveau des classes séparées. On remarque que la classe bilingue se situe à mi-chemin entre les deux populations, mais avec plus de 50 % d'items en interaction.

Etude des résultats par la XHI² classique.

Toutes les comparaisons possibles (lorsque les effectifs sont supérieurs à 5) sont évidemment significatives.
De toute manière, des résultats aussi nets se passent de commentaires...

Présentation des items du test n° 7

(maniement oral de la langue)
Le maniement oral de la langue est évalué à partir de la note de vocabulaire (test n°5) et de la note de correction syntaxique (test n° 6). Ces deux épreuves différenciant nettement les deux sous-groupes, nous devons nous attendre à ce que la combinaison des deux aille dans le même sens.

Comportement de l'ensemble de la population et des différents sous-groupes.

	Groupe total n = 367			Sous-groupe 1 n = 163			Sous-groupe 2 n = 176		
Item	%	IDS	T	%	IDS	T	%	IDS	T
1	.77		S	.52	.84	D	.99		S
2	.37	.69	D	.00		inf	.73		I
3	.44	.73	D	.06		inf	.81		S
4	.60	.66	D	.23		inf	.94		S
5	.49	.72	D	.08		inf	.86		S
6	.41	.77	D	.02		inf	.80		S
7	.44	.81	D	.01		inf	.86		S
8	.57	.78	D	.17		inf	.98		S
9	.70	.57	D	.39	.62	D	.99		S
10	.55	.85	D	.12		inf	.99		S
11	.45	.83	D	.01		inf	.88		S
12	.47	.89	D	.02		inf	.92		S
13	.53	.91	D	.05		inf	.99		S

Comportement des classes séparées.

TABLEAU 12

Item	1	4	6	10	2	3	7	8	9	5
1	D	D	I	D	S	S	S	S	S	S
2	inf	inf	inf	inf	D	I	I	S	S	inf
3	inf	inf	inf	inf	S	S	S	I	S	I
4	inf	inf	D	I	S	S	S	S	S	I
5	inf	inf	inf	inf	S	S	I	S	S	S
6	inf	inf	inf	inf	S	S	S	I	S	inf
7	inf	inf	inf	inf	D	S	S	S	S	I
8	I	inf	I	inf	S	S	S	S	S	I
9	D	D	D	D	S	S	S	S	S	S
10	D	inf	inf	inf	S	S	S	S	S	I
11	inf	inf	inf	inf	S	S	S	S	S	S
12	inf	inf	inf	inf	S	S	S	S	S	I
13	inf	inf	inf	inf	S	S	S	S	S	I

Cette évaluation n'apporte aucun élément neuf d'information par rapport à l'épreuve précédente.
Nous ne voyons donc pas son intérêt.

e) Epreuve de compréhension verbale (n°8)

L'enfant doit identifier, parmi 5 images, celle qui correspond à la phrase lue par l'examinateur.
A notre avis cette épreuve ne diffère de l'épreuve de vocabulaire (a) que sur deux points :

1. elle porte sur des saynètes (compréhension de situations complexes).
2. elle se situe beaucoup plus tard dans l'année scolaire, fin décembre, alors que les deux premiers tests avaient eu lieu en septembre-octobre.

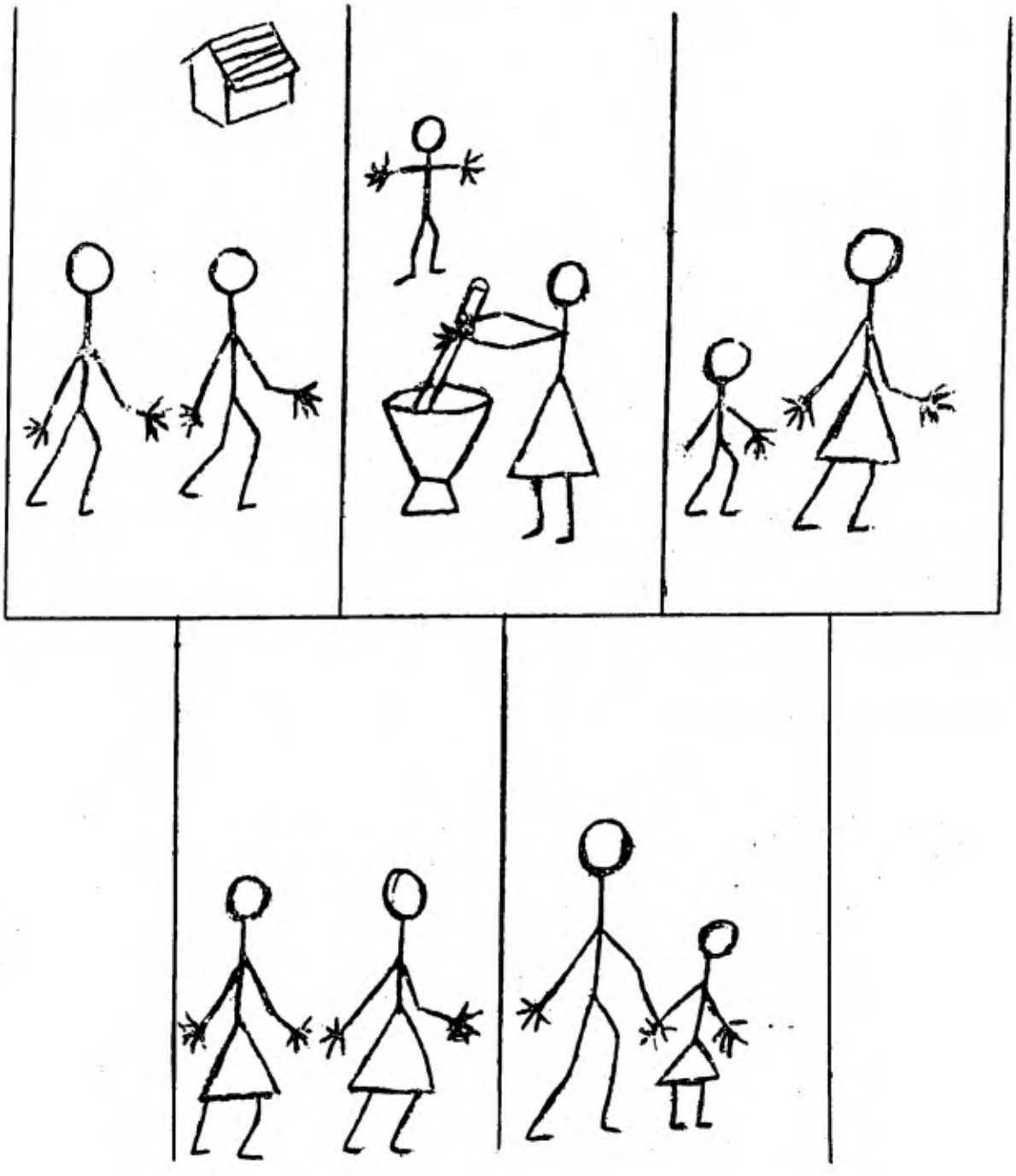
Liste des items du test n° 8

1. Mambu vient de l'école, fatigué, il est assis sur la chaise.
2. Mambu rentre très content à la maison car il a attrapé un poisson à la pêche.
3. Mambu a tué un oiseau avec son lance-pierre et cet oiseau est tombé à côté de l'arbre.
4. Le papa de Mambu voudrait aller chasser (tuer) un grand animal : que va-t-il apporter à la chasse?
5. Un jour, Mambu rentrait à la maison quand il pleuvait.
6. Mambu et sa maman se promènent.
7. Sous l'arbre, Mambu est en train de lire, assis sur une chaise.
8. Il fait nuit dans le village de Mambu.
9. Les deux amis de Mambu sont arrivés en retard en classe. Ils sont punis à genoux.
10. La maman de Mambu s'en va à la rivière avec une dame-jeanne sur la tête.
11. Après avoir étudié, Mambu se promène avec son chien.
12. Mambu avance vers son père qui est en train de fumer la pipe.

Comportement de l'ensemble de la population et des différents sous-groupes.

	Groupe total n = 378			Sous-groupe 1 n = 178			Sous-groupe 2 n = 172		
Item	%	IDS	T	%	IDS	T	%	IDS	T
1	.66	.50	D	.43	.22	I	.95		S
2	.66	.48	D	.49	.53	D	.84		S
3	.59	.61	D	.34	.43	I	.90		S
4	.54	.58	D	.30	.22	I	.84		S
5	.51	.69	D	.16		inf	.86		S
6	.13		inf	.09		inf	.15		inf
7	.50	.56	D	.20		inf	.81		S
8	.28	.37	I	.16		inf	.42		I
9	.86		S	.74	.29	I	.98		S
10	.64	.46	I	.45	.52	D	.82		S
11	.70	.41	I	.56	.50	D	.84		S
12	.73	.42	I	.56	.48	D	.90		S

L'item 6 est inférieur pour toute la population et pour chacun des sous-groupes. Nous reproduisons ci-après les images parmi lesquelles l'élève doit choisir celle qui correspond à la phrase prononcée. On remarquera l'ambiguité de plusieurs des dessins.



L'item 9 est supérieur pour le groupe total, mais entre les deux sous-groupes, nous observons une différence au profit du sous-groupe lingalaphone.
Nous nous intéressons particulièrement aux items 2,10,11 et 12, qui sont supérieurs pour les élèves lingalaphones alors qu'ils différencient les élèves francophones.
Cela signifie qu'une fraction importante de ces derniers atteint la compétence.
L'analyse par classe devrait nous renseigner à ce sujet.
L'étude du contenu de ces items nous révèle la présence pour chacun d'eux de deux éléments bien définis et facilement identifiables parmi les autres dessins :

- item 2 : Mambu et le poisson
- item 10 : Maman et la dame-jeanne
- item 11 : Papa et le chien
- item 12 : Papa et la pipe

Alors que les autres items font appel à la compréhension d'une structure complexe, ces quatre items sont résolus par une compréhension d'éléments juxtaposés.

Comportement des classes séparées.

Item	1	4	6	10	2	3	7	8	9	5
1	D	I	I	I	S	S	S	S	S	I
2	I	D	S	I	S	S	S	S	I	D
3	inf	inf	S	inf	S	S	S	S	I	I
4	I	I	inf	D	S	S	S	S	D	inf
5	inf	inf	inf	D	S	S	S	S	S	D
6	inf	inf	inf	inf	inf	inf	inf	inf	inf	inf
7	I	inf	inf	D	I	D	S	S	D	D
8	inf	inf	D	inf	D	D	D	I	D	inf
9	S	S	S	I	S	S	S	S	S	S
10	D	I	D	I	S	S	S	D	D	I
11	I	D	D	D	S	D	S	S	S	I
12	I	D	S	I	S	S	S	S	I	I

En analysant le comportement des 4 items dont nous avons parlé plus haut, nous constatons pour le sous-groupe francophone :

- 7 différenciations sur 16 possibles;
- 2 supérieurs pour la classe 6, qui démontre ici sa singularité en s'approchant, au moins pour une partie des élèves, des compétences sous-groupe lingalaphone.

La classe n°9, par contre, éprouve de sérieuses difficultés pour cet exercice avec 4 items supérieurs et 7 items à 50%, dont 4 de différenciation.

Là, nous observons le phénomène inverse : une partie de la classe n'atteint pas systématiquement la compétence escomptée.

Etude des résultats par XHI² classique.

Toutes les comparaisons sont significatives, même pour l'item 6!

2° RECHERCHES DE FRYDMAN

1) Objectifs de la recherche.

Frydman (1) se proposait de tester l'éducabilité de l'aptitude à utiliser un référentiel. Son hypothèse reposait sur l'alternative suivante : la simple mise en contact avec des problèmes exigeant le recours à un matériel de référence suffit-elle, ou faut-il mettre au point un entraînement systématique et une technique d'apprentissage?

2) Dispositif expérimental.

Le dispositif expérimental (2), représenté schématiquement ci-après, se composait de quatre échantillons comprenant chacun 12 classes : les trois premiers constitués de classes de 6ème primaire, le quatrième (D) de classes de 3ème année de l'enseignement secondaire. Les groupes expérimentaux (A et B) ont été soumis à des traitements différents et leurs performances comparées à celles des groupes de contrôle (C et D). Toutes les observations ont été représentées sur le schéma par la lettre O, qu'elles fussent pré-expérimentales, expérimentales ou post-expérimentales.

Les symboles X et X' désignent les deux variables indépendantes correspondant aux deux traitements qui furent administrés aux groupes A et B.

Trente six classes de 6e primaire ont été soumises à deux prétests basés sur l'utilisation de référentiels , et dont le niveau d'exigence était indiscutablement élevé. Nous étudierons plus loin l'application de la technique "arc sinus" sur les résultats enregistrés à ces épreuves, qui a permis de constituer de manière aléatoire trois groupes équivalents de 12 classes.

---

(1) FRYDMAN, Marcel.
Le développement de l'aptitude à exploiter un matériel de références. Etude expérimentale des conditions requises par un curriculum. Université de l'Etat à Mons - 1975

(2) FRYDMAN, Marcel.
L'utilisation d'un matériel de références. Département de Psychologie - Université de l'Etat à Mons - Décembre 1976.

REPRESENTATION SCHEMATIQUE DU DISPOSITIF

Groupes	PRETESTS	TRAITEMENTS																							POSTTESTS
A	O_{p1} O_{p2}		X_1O_1				X_3O_3				X_5O_5				X_7O_7				X_9O_9			$X_{11}O_{11}$			OP_1 OP_2 OP_3
B	O'_{p1} O'_{p2}	Z_1	$X'_1O'_1$	Y_1	$X'_2O'_2$	Y_2	$X'_3O'_3$	Y_3	$X'_4O'_4$	Y_4	$X'_5O'_5$	Y_5	$X'_6O'_6$	Y_6	$X'_7O'_7$	Y_7	$X'_8O'_8$	Y_8	$X'_9O'_9$	Y_9	Z_2	$X'_{11}O'_{11}$	Y_{11}		$O'P_1O'P_2O'P_3$
C	O''_{p1} O''_{p2}												W												$O''P_1O''P_2O''P_3$
D													T												$O'''P_1O'''P_2O'''P_3$

<u>Traitements</u> : Z_1 = premier entraînement à l'utilisation d'un référentiel comportant l'exécution individuelle, l'auto-correction et la correction collective;

X et X' = exécution indiduelle;

Y = entraînement assisté et renforcement;

Z_2 = correction de O_{p2};

W = activité traditionnelle;

T = activité traditionnelle + expérience personnelle + maturation.

3. Traitement.

Les classes du groupe A (traitement X) ont bénéficié de six épreuves d'entraînement comportant chacune dix items. Les sujets ont eu à résoudre, à l'aide de deux référentiels, 60 problèmes présentés selon un ordre de difficulté croissante. Ils travaillaient de manière strictement individuellle et aucune démarche ne leur a été suggérée par l'instituteur qui n'avait d'ailleurs aucune information sur le contenu des tests. A la fin de la séance, les feuilles réponses étaient enlevées aux élèves qui ne recevaient aucun feed-back. On peut dire qu'ils furent placés dans une situation de "learning by doing".

Les classes du groupe B (traitement X') ont reçu les mêmes tests. Ici aussi les feuilles réponses étaient retirées et notées par un évaluateur extérieur. Toutefois, les maîtres - qui avaient bénéficié de plusieurs séances de formation organisées par l'expérimentateur - dirigeaient la leçon de correction collective au cours de laquelle ils étaient chargés de multiplier les renforcements. La semaine suivante, une nouvelle épreuve d'entraînement, basée sur l'utilisation d'un référentiel différent, était appliquée aux seules classes B et également suivie d'une correction simultanée. Huit jours plus tard, les élèves retrouvaient un test du premier type, parallèlement à leurs compagnons de A, et le cycle recommençait pour se terminer à la fin mai.
Pendant quatre mois, les sujets furent donc soumis à un traitement intensif impliquant, sous une forme ou une autre, un entraînement par semaine, en moyenne, les tests alternant avec les séances de renforcement.

4. Posttests.

Les élèves des trois échantillons de classes ont été soumis à trois posttests au début du mois de juin. Le premier était identique au premier prétest et composé des huit mêmes questions. Le deuxième exigeait le recours aux référentiels que les élèves de A et B avaient manipulés à six reprises mais les problèmes à résoudre étaient relativement plus exigeants que ceux des tests précédents. Le troisième, enfin, impliquait l'exploitation d'une série de documents qui n'avaient jamais fait leur apparition antérieurement. La première et la troisième épreuves d'évaluation finale devaient permettre la mise en évidence éventuelle d'un effet de transfert d'apprentissage. L'une et l'autre étaient basées sur un matériel de référence qui n'avait pas été utilisé au cours de la phrase d'apprentissage.

5. Résultats.

Après avoir traité les résultats aux posttests par les techniques d'analyse de variance hiérarchisées, FRYDMAN a constaté que, dans aucun cas, il n'y a de différence significative entre A et C. Il a donc fallu conclure à l'inefficacité des conditions expérimentales du groupe A, c'est-à-dire du "learning by doing".

Par contre, le groupe B s'est révélé significativement supérieur aux deux autres échantillons à chacune des trois épreuves d'évaluation finale. Les résultats enregistrés montrent donc que l'aptitude à utiliser un référentiel peut être développée dès la sixième primaire et qu'elle s'accompagne d'un transfert.

Il était cependant indispensable d'évaluer la supériorité des classes B avec autant d'objectivité que possible. Dans ce but, on a comparé leurs performances avec celles d'un échantillon de classes provenant de l'enseignement secondaire (groupe D), auxquelles ne furent administrés que les posttests.

Si, à la troisième de ces épreuves, les scores obtenus par les deux groupes doivent être considérés comme équivalents, aux deux autres en revanche - les plus exigeantes aussi du point de vue de l'exploitation de références - c'est l'échantillon B qui s'avère significativement supérieur.

A. Constitution des échantillons expérimentaux.

1. Analyse du groupe total.

Trente-six classes de 6ème primaire de la région bruxelloise et du Brabant wallon ont été retenues et soumises aux deux épreuves préexpérimentales.

Le premier (8 items) avait pour objet l'aptitude à consulter un indicateur de chemins de fer.

Le second (9 items) était moins difficile et portait sur une matière (la langue maternelle) plus familière à l'élève.

Les données nécessaires à la résolution des items n'ayant cependant pas encore été abordées au début de la sixième primaire, les réponses ne pouvaient être obtenues que par une exploitation correcte du matériel de référence.

Les conditions du testing ont été identiques pour les deux prétests. Les élèves travaillaient seuls.

L'analyse du tableau synthétique ci-après par la technique "arc sinus" révèle 8 items stables.

L'item stable peut être défini comme celui pour lequel l'ensemble des classes se comporte de manière homogène.

Les instabilités occasionnellles sont inévitables dans une telle épreuve. Les déviations systématiques de certaines classes ont permis de situer celles-ci de la manière suivante :

- six classes inférieures à l'intervalle ;
- vingt-quatre classes situées dans l'intervalle ;
- six classes supérieures à l'intervalle.

Il devenait donc possible de constituer trois groupes de douze unités de façon aléatoire en tirant au sort, sans replacement, deux classes supérieures, deux classes inférieures et huit classes moyennes dans chaque cas.

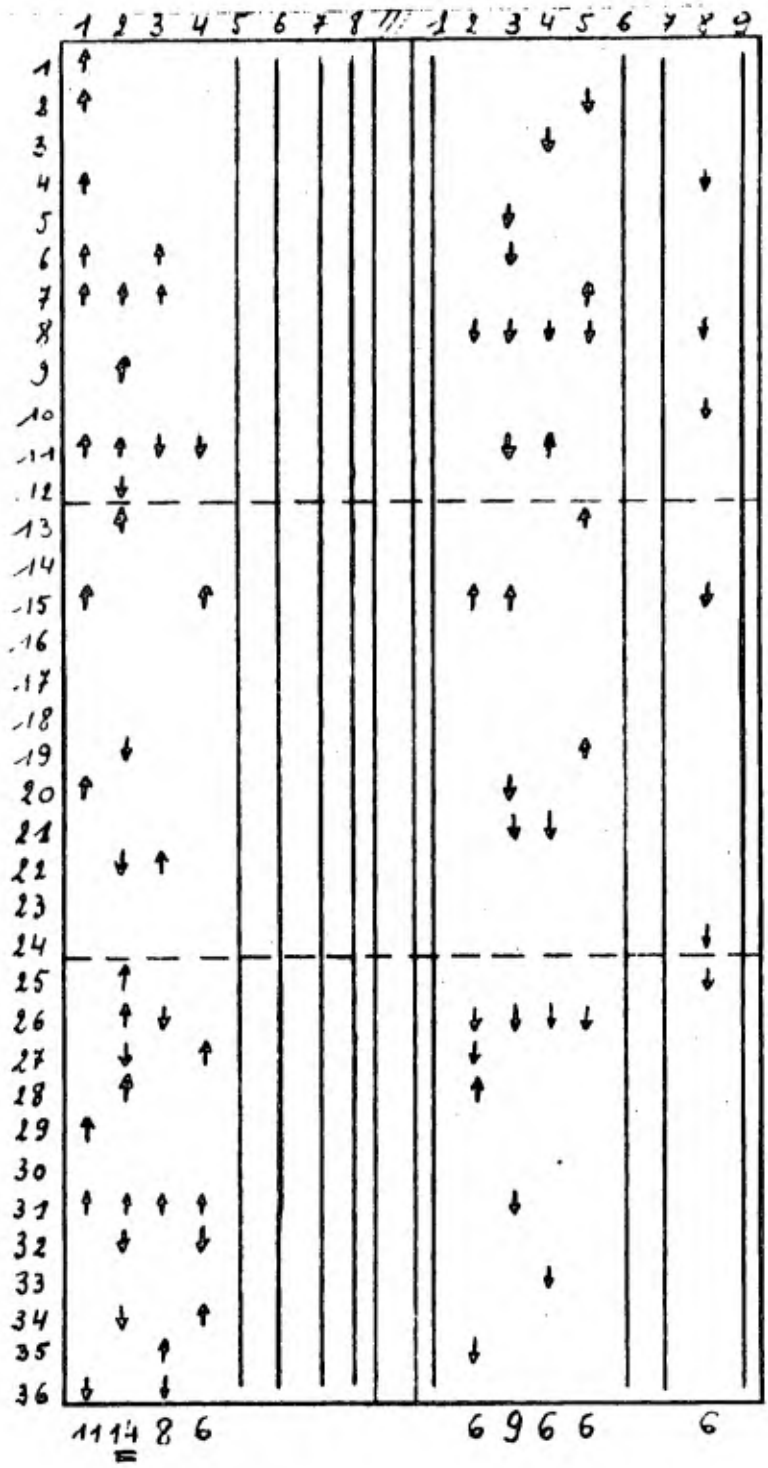
C'est également de manière aléatoire qu'ont été désignés les deux groupes expérimentaux (A et B) et le groupe de contrôle (C) en numérotant les classes respectivement de 1 à 12 (groupe A), de 13 à 24 (groupe B) et de 25 à 36 (groupe C).

Tableau synthétisant les observations sur l'ensemble de la populatio

Les items sont en abscisse et les classes en ordonnée.

Une flèche vers le haut indique que la classe est supérieure à l'intervalle et une flèche vers le bas qu'elle est inférieure.

Une ligne verticale signifie que le trop faible pourcentage de réussite (moins de 5%) interdit l'application de la technique "arc sinus'



Les huit items présentant moins d'1/3 de rejets ont été retenus pour la répartition des classes.

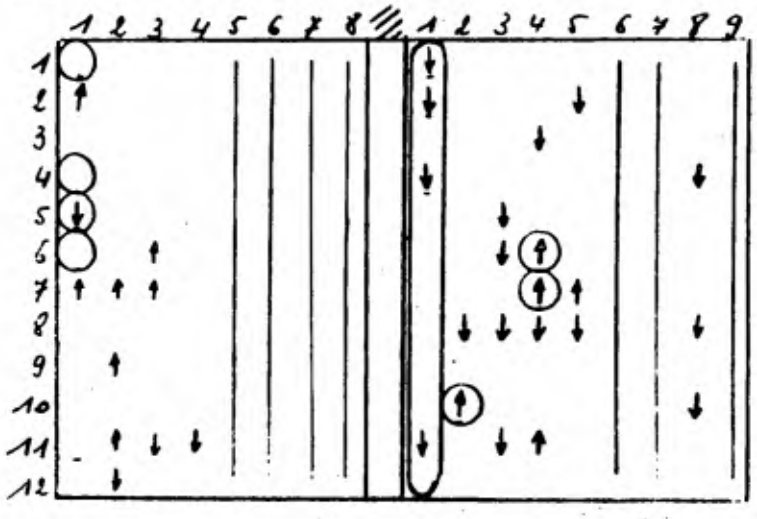
2. Analyse des sous-groupes.

Les sous-groupes ayant été constitués, il fallait vérifier leur cohérence interne et leur équivalence.

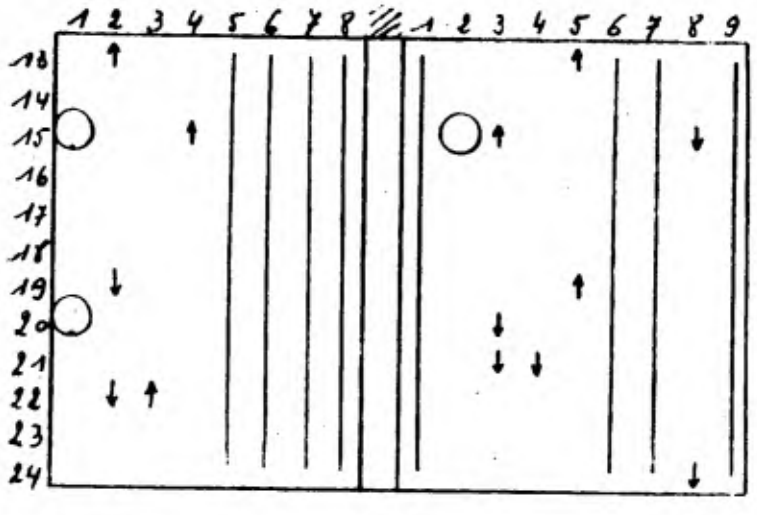
a) Etude de la cohérence interne des sous-groupes

En appliquant la technique arc sinus pour les items de chacun des sous-groupes, on pouvait comparer les résultats sur l'ensemble de la population et sur chacun des sous-groupes. Dans les tableaux ci-dessous, les divergences sont entourées d'un cercle.

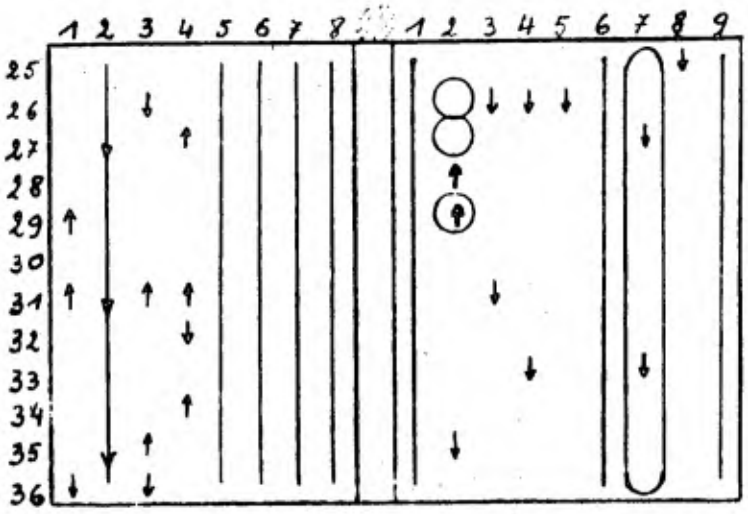
Groupe A



Groupe B



Groupe C



Conclusions :

Les divergences sont peu nombreuses.

Le nombre de rejets est moins important dans le groupe B (13/108) que dans le groupe A (32/120). Le groupe B est plus homogène que les autres.

b) Application d'autres techniques.

L'application de la technique XHI² entre les moyennes des classes confirme la répartition en classes supérieures, moyennes, inférieures.

L'analyse de variance hiérarchisée a également montré qu'il n'y avait aucune différence significative entre les trois groupes ainsi formés, non seulement pour les résultats obtenus au premier comme au second prétest, mais également en fonction des scores enregistrés à chacune des dix-sept questions envisagées séparément .

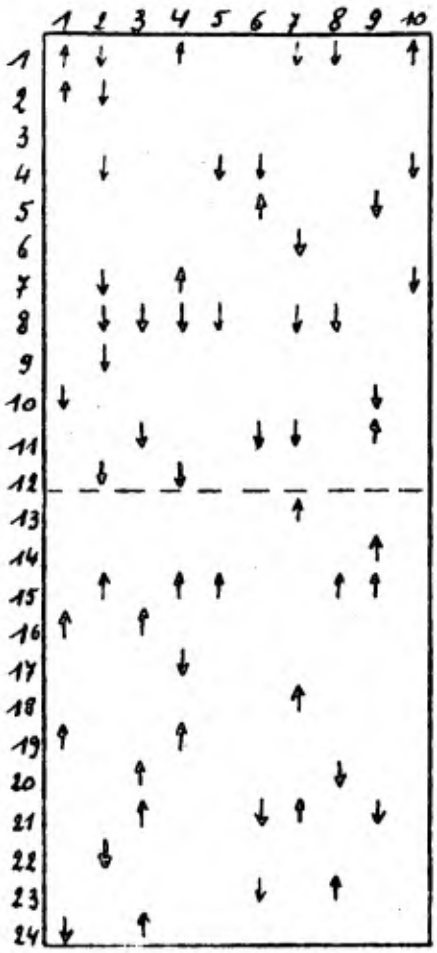
B. Analyse des résultats du traitement.

Les classes des groupes A et B ayant passé simultanément les six épreuves-tests du traitement, il est possible, par la technique "arc sinus" de comparer leurs résultats.

Pour chaque épreuve, nous analyserons les résultats globaux (comme si les deux groupes ne formaient qu'une seule population) et les résultats de chacun des groupes séparés.

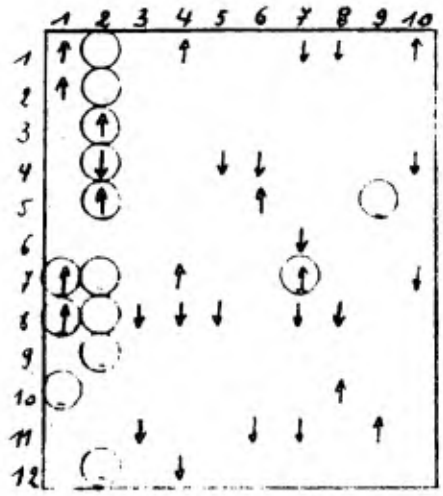
1) Première épreuve.

Groupe total

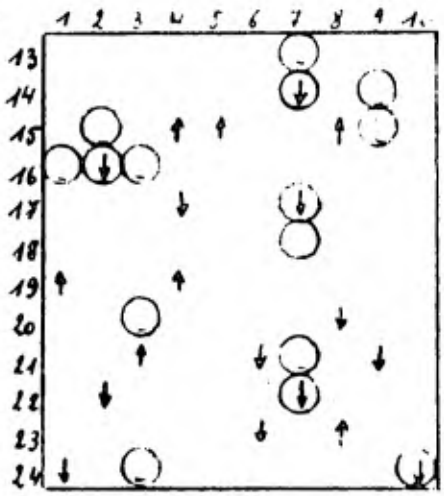


Groupes séparés

A



B



L'analyse de la population totale montre une nette différence entre les deux groupes :

Groupe A : 26 rejets négatifs et 7 rejets positifs /120

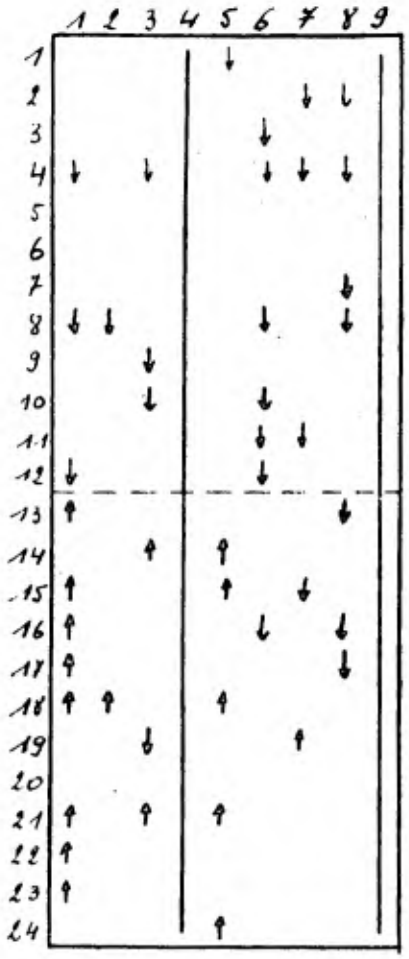
Groupe B : 7 rejets négatifs et 17 rejets positifs/120

L'item 2 oppose les deux groupes nettement. Il s'agit d'une question d'application relative à la construction d'adverbes en -ment à partir d'adjectifs. Sans doute les élèves du groupe A ont-ils été tentés de répondre en puisant exclusivement dans leurs connaissances orthographiques plutôt qu'en consultant le référentiel.

L'item 7 perturbe le groupe B tout en lui laissant un net avantage. Cette question nécessite le recours à trois références en cascade.

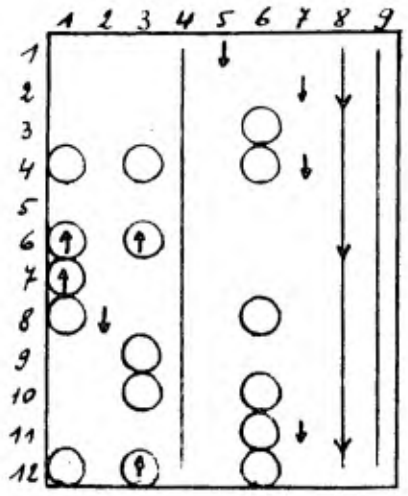
2) Deuxième épreuve.

Groupe total

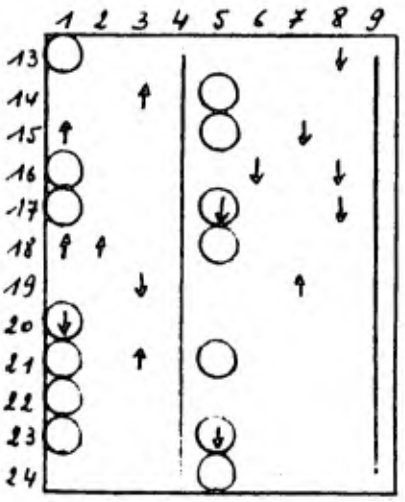


Groupes séparés

A



B



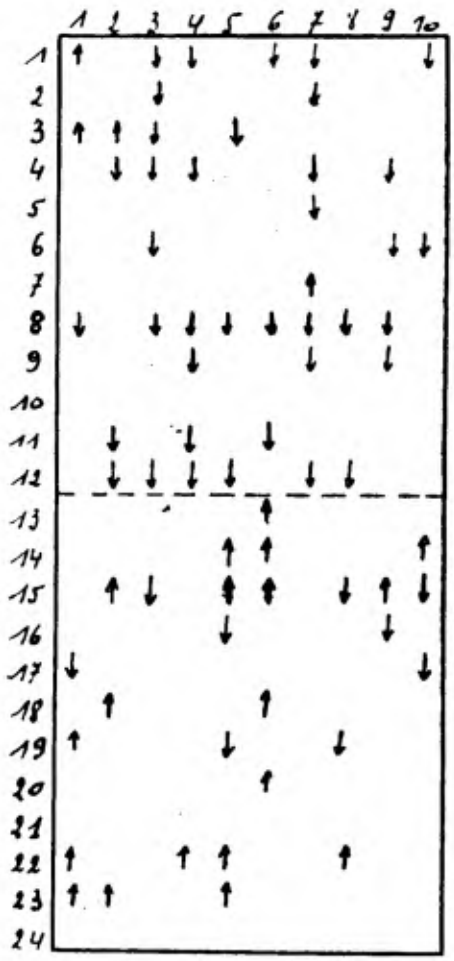
Groupe A : 21 rejets négatifs et 0 rejet positif /84

Groupe B : 6 rejets négatifs et 17 rejets positifs /84

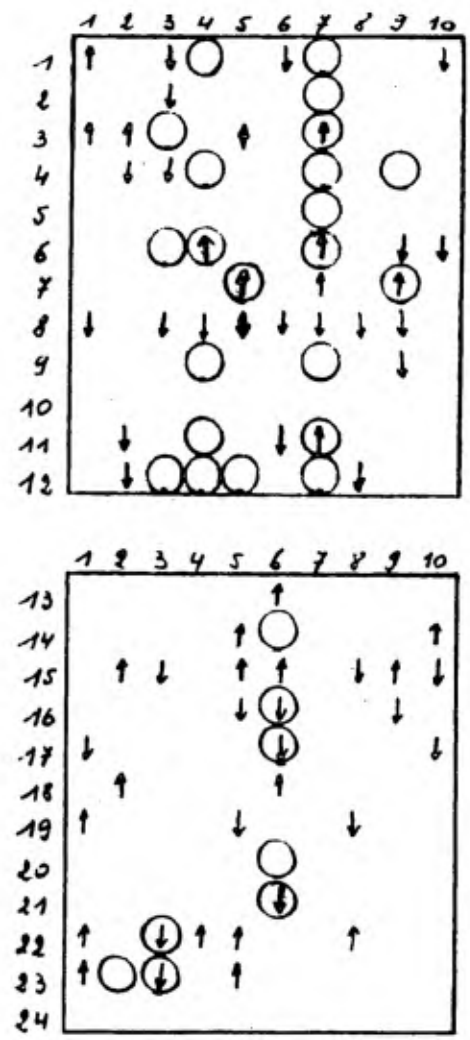
La différence entre les deux groupes s'accentue.
Elle épargne cependant les items 2 et 7.
Il est difficile de se prononcer au sujet des items 4,8 et 9 qui présentent un trop faible pourcentage de réussite.
On remarque que les classes 5 et 6 ne se distinguent aucunement de la moyenne, comme la classe 20...

3) Troisième épreuve.

Groupe total



Groupes séparés



Groupe A : 38 rejets négatifs et 4 rejets positifs /120

Groupe B : 9 rejets négatifs et 19 rejets positifs /120

Les items 3,4,6 et 7 opposent le plus ces deux groupes. Surtout le 7, qui joue un rôle de provocateur: "Comment forme-t-on les temps composés des verbes?"

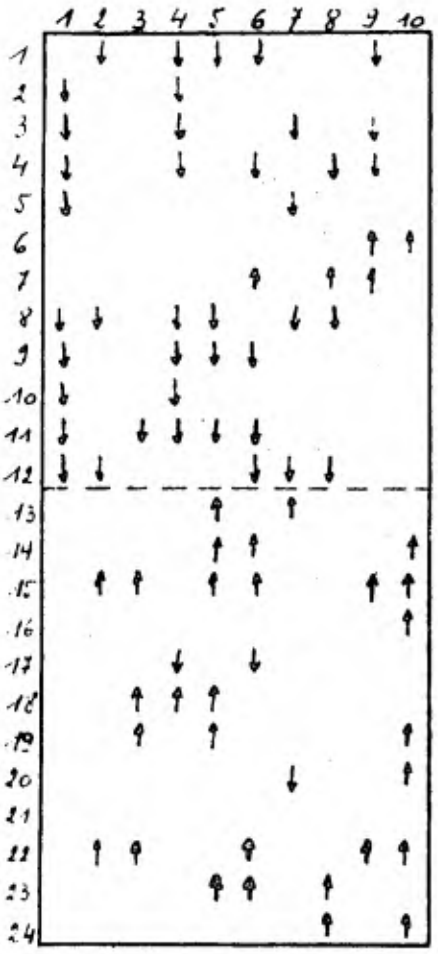
Qui ne serait tenté de répondre à l'aide de ses connaissances à cette question?

Les classes 7 et 10 du groupe A se comportent très bien et dament le pion à plusieurs classes du groupe B (16, 17, 21, 24) pour lesquelles on peut suspecter un certain essoufflement.

La classe 8, faible dès le départ, s'enlise de plus en plus, tandis que la classe 15 a un comportement plutôt anarchique.

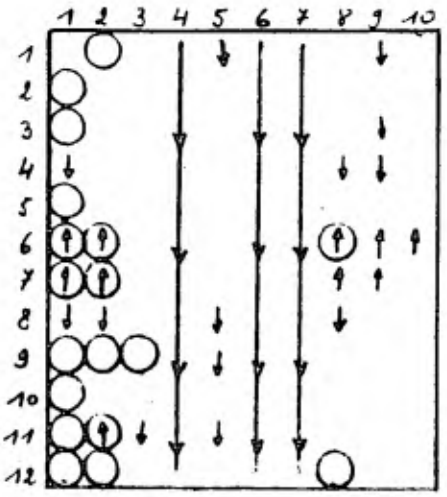
4) Quatrième épreuve

Groupe total

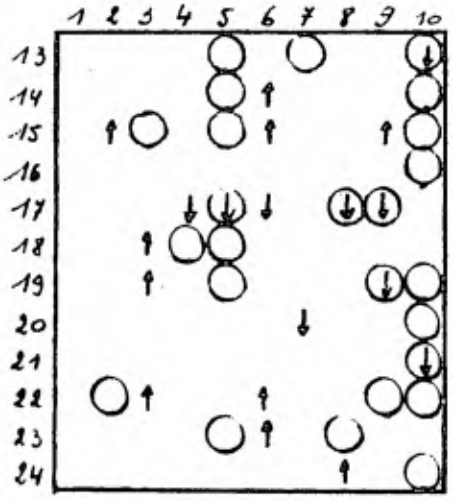


Groupes séparés

A



B



Groupe A : 40 rejets négatifs et 5 rejets positifs /120

Groupe B : 3 rejets négatifs et 29 rejets positifs /120

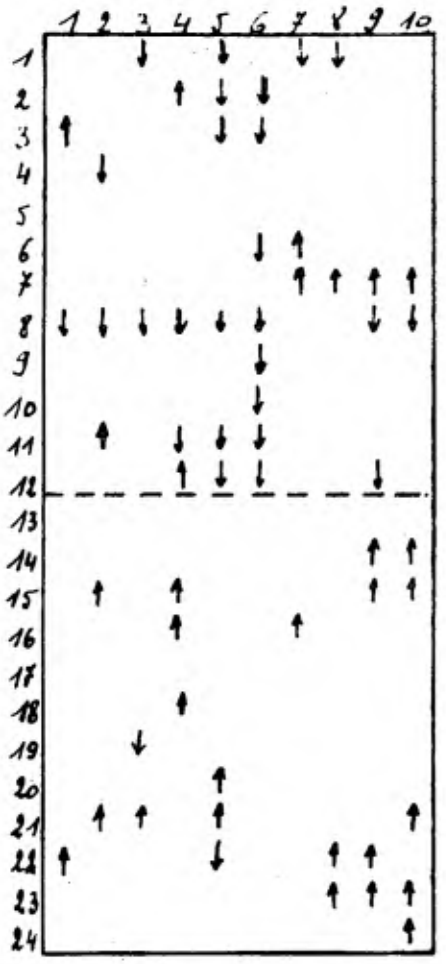
La supériorité du groupe B s'accentue. Elle est cependant beaucoup plus faible pour certains items (2,3,7,8 et 9).

Les classes 6 et 7 du groupe A montrent une compétence comparable à une bonne partie des classes du groupe B.

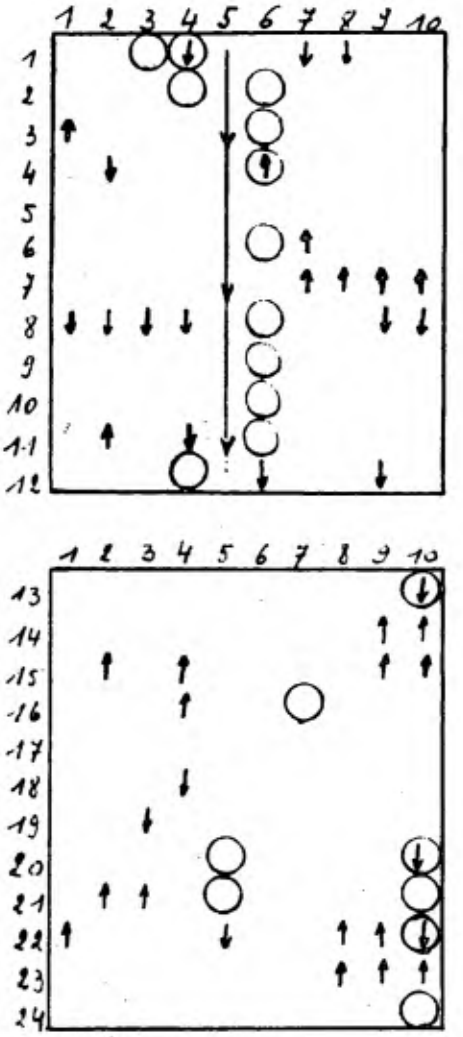
La classe 8 est toujours aussi faible, alors que la classe 17 perd pied.

5) Cinquième épreuve.

Groupe total



Groupes séparés



Groupe A : 26 rejets négatifs et 9 rejets positifs/120

Groupe B : 2 rejets négatifs et 21 rejets positifs/120

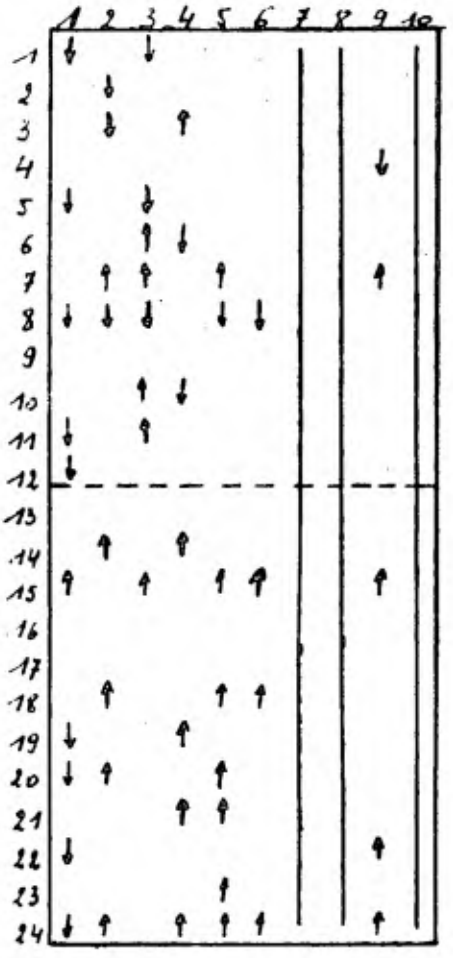
La supériorité du groupe B est beaucoup moins nette. Elle ne se manifeste que pour les items 5, 6 et dans une moindre mesure 10.

Ce test est fondé sur la lecture d'un texte, et plusieurs questions s'apparentent davantage à la "lecture silencieuse" classique qu'à la recherche de références.

La classe 7 confirme sa bonne tenue : elle figurerait parmi les meilleures du groupe B. La classe 17 semble avoir surmonté sa défaillance.

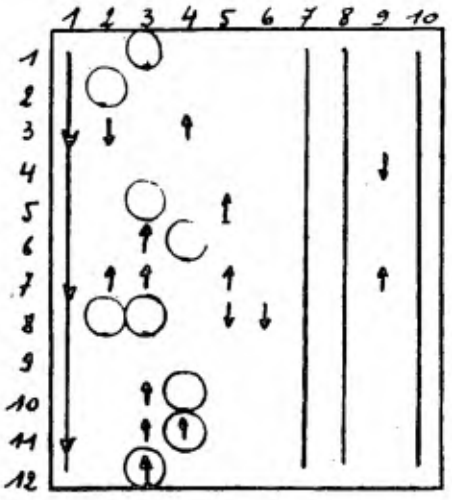
6) Sixième épreuve.

Groupe total

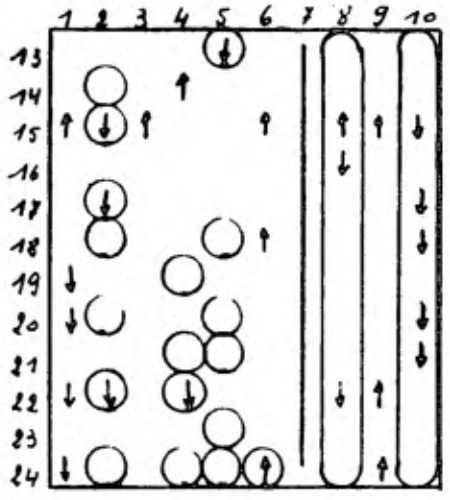


Groupes séparés

A



B



Groupe A : 17 rejets négatifs et 8 rejets positifs /84

Groupe B : 4 rejets négatifs et 22 rejets positifs/84

Si la supériorité du groupe B est manifeste sur l'ensemble de l'épreuve, elle ne s'exprime nettement qu'au niveau des items 2,4 et 5.

La classe 7 confirme sa bonne tenue.

7) Evaluation hiérarchique.

a) Matrices de concordance

Cette évaluation n'a de sens que si nous la menons au travers de l'entièreté du curriculum. Nous avons donc regroupé les 59 items en une suite continue de données, en indiquant les absences d'élèves à certaines épreuves par des perforations 5. C'est donc le programme spécial CØIDS08 qui a été utilisé pour traiter les données.

Nous avons étudié chaque sous-groupe séparément. Au prétest, aucune interconcordance ne s'est manifestée entre les items significatifs (assimilable à 50% de réussite), ce qui montre bien que la population n'avait nullement été entraînée à ce genre de travail.

Les (faibles) compétences individuelles ont seules joué, et aucun effet didactique antérieur n'est suspecté.

Etude des items du traitement

Groupe A

Trois interconcordances significatives sont relevées :

items 41-42 : concernent tous deux un texte de Léautaud. Il est possible d'y répondre sans recours aux références.

items 15-55 : ne prévoient pas de recours aux références.

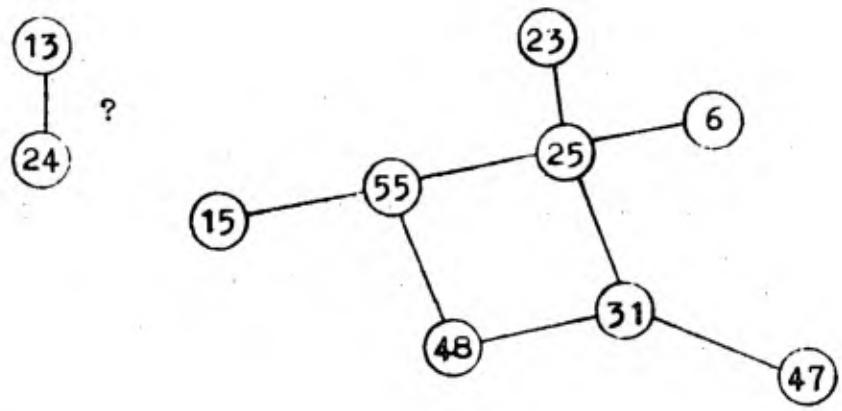
items 7 - 55 : pas de référence pour le 55; l'item 7 demandait des références faciles, mais en cascade...

Groupe B

On relève 9 concordances significatives :

6-25/13-24/15-55/23-25/25-31/25-55/31-47/31-48/48-55.

Nous schématisons les "chaînes" d'items observés ci-dessous :



Nous pouvons répartir ces items selon deux critères :

les items 6-23-25 sont des items d'application (300).

Il n'est pas impossible que les matières qu'ils traitent aient été abordées en classe.

Les items 15-55-48-47 se résolvent sans recours aux références (Ro) et s'apparentent à la lecture silencieuse.

L'item 31 concerne une production (500) classique.

Ce qui relie les individus concerne essentiellement des acquis scolaires inscrits au programme de 6ème année primaire.

b) Etude des hiérarchies.

La diversité des objectifs poursuivis par les différents items rend très difficile l'interprétation des hiérarchies.
N'oublions pas non plus que l'étalement des six épreuves sur l'année scolaire ne favorise pas la concordance entre les résultats.
Cependant, certains items reviennent régulièrement dans les hiérarchies :

1. Les items 18,19,50,56,59, caractérisés par un pourcentage de réussite très bas (moins de 10%).
   Il s'agit d'items très exigeants, qui poussent les individus dans leurs derniers retranchements.

2. Les items 24,25,27 et 35, qui touchent des domaines grammaticaux familiers.

3. Les items 47,48 et 55 qui ne font pas appel aux références.
   L'item 36, quant à lui, ne peut être rattaché à aucune des familles ci-dessus.

8) Application d'autres techniques.

FRYDMAN a appliqué l'analyse de variance hiérarchisée sur ces mêmes données.

Voici ses conclusions :

"Malgré une variabilité entre classes importante et significative, la valeur du F entre traitements dépasse chaque fois le niveau de p = 0,01 (sauf au cinquième test, où F est significatif à .05). Cette supériorité constante et significative est sans doute le meilleur témoignage du savoir-faire qui a progressivement été développé dans le groupe B et clairement démontré par les résultats enregistrés aux posttests."

FRYDMAN a remarqué les différences d'évolution des classes, notamment en ce qui concerne la classe 7. Cependant la technique qu'il a utilisée ne lui permet pas de désigner les items responsables des différences observées.

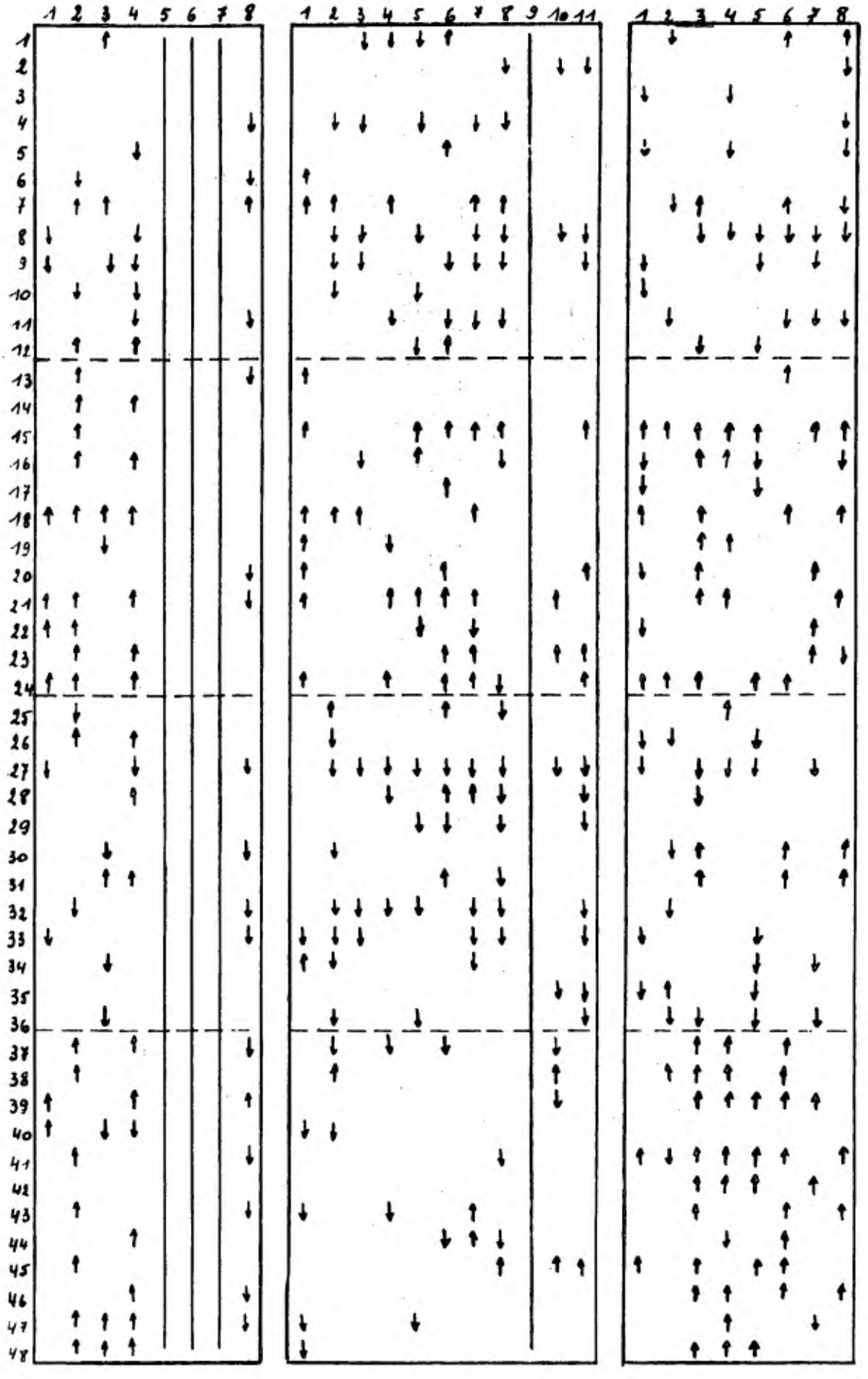
Chaque épreuve est considérée comme un tout homogène. Or, nous avons vu que c'était loin d'être exact.

9) Conclusion.

Il n'est pas étonnant, et peut-être est-ce même inhérent à la nature de la didactique, que nous observions des résultats qui ne présentent qu'une faible cohérence. Tout en recherchant un effet d'apprentissage, que FRYDMAN a mis en évidence par les techniques statistiques classiques, le recours à des références provoque des démarches divergentes chez les individus. Globalement, les résultats sont meilleurs, mais il est difficile, voire impossible de déterminer avec précision la nature des changements intervenus. Est-ce une raison pour abandonner la poursuite de ces objectifs?

C. Analyse des résultats aux posttests.

1. Homogénéité du groupe total (Arc sinus)



Interprétation.

L'application de la technique "arc sinus" ne se justifie que si l'on émet l'hypothèse de départ de l'homogénéité du groupe observé. Cela signifie, dans le cas présent, que l'on considère chacun des sous-groupes à un niveau équivalent. Or, nous savons que le groupe D est composé d'élèves de 3ème année de l'enseignement secondaire. Outre la maturation liée à l'âge, il faut tenir compte de l'enseignement supplémentaire qu'ils ont reçu.

Cependant, cette répartition est plus équilibrée qu'il ne paraît à première vue : deux groupes (A et C) n'ont reçu aucun entraînement particulier, les deux autres (B et D) ont subi des traitements, soit dans un curriculum expérimental spécifique (B), soit dans le cadre institutionnel normal (D).

Comparaison pour chacun des subtests.

1er subtest (identique au premier prétest)

Groupe A : 13 rejets négatifs et 6 rejets positifs /60

Groupe B : 4 rejets négatifs et 20 rejets positifs /60

Groupe C : 12 rejets négatifs et 4 rejets positifs /60

Groupe D : 7 rejets négatifs et 19 rejets positifs /60

Globalement, le groupe B est légèrement supérieur au groupe D, mais très supérieur aux groupes A et C.

Ce sont les items 2 et 4 qui supportent le plus cette opposition entre les groupes (B-D) et (A-C).

L'item 1 les oppose plus faiblement, les items 3 et 8 aucunement, et l'information liée aux items 5,6 et 7 n'est pas exploitable.

TOURNEUR Y.et FRYDMAN M.(1) ont appliqué la théorie de la généralisabilité à l'estimation d'erreurs sur cette épreuve en comparant les résultats au prétest et au posttest. Ils observent que la composante interquestion est la source de variance la plus importante.

(1) Y.TOURNEUR et M.FRYDMAN.
Une application de la théorie de la généralisabilité à l'estimation d'erreurs.
Université de l'Etat à Mons. Doc. 770-620 - 1977 - 17p

" Pour accroître la précision de la différenciation sur la facette M (moment),il serait nécessaire de modifier le dispositif en introduisant l'un ou l'autre des deux changements possibles :

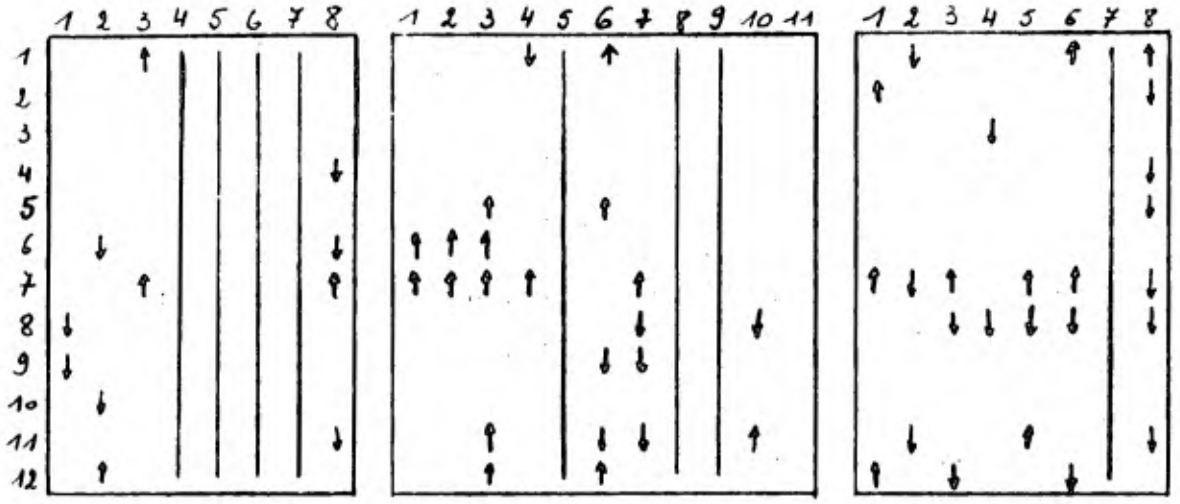
1. augmenter le nombre de questions dans le test ;
2. grouper les questions en objectifs. "

Leur préférence va à la seconde solution, qui devrait réduire l'hétérogénéité de difficulté des questions. Ils rejoignent ainsi nos conclusions sur le manque d'homogénéité du matériel utilisé et sur la difficulté de déterminer a posteriori avec précision la nature des modifications intervenues à la suite de cet apprentissage.

2e subtest (comparable aux épreuves d'entraînement proposées aux groupes A et B)
Groupe A : 31 rejets négatifs et 9 rejets positifs /120
Groupe B : 6 rejets négatifs et 32 rejets positifs /120
Groupe C : 38 rejets négatifs et 6 rejets positifs /120
Groupe D : 15 rejets négatifs et 7 rejets positifs /120
L'opposition entre les groupes est mieux répartie, aussi bien sur les classes que sur les items.
L'effet global est beaucoup plus net sur le groupe B que sur le groupe D, les groupes A et C sont aussi faibles l'un que l'autre.

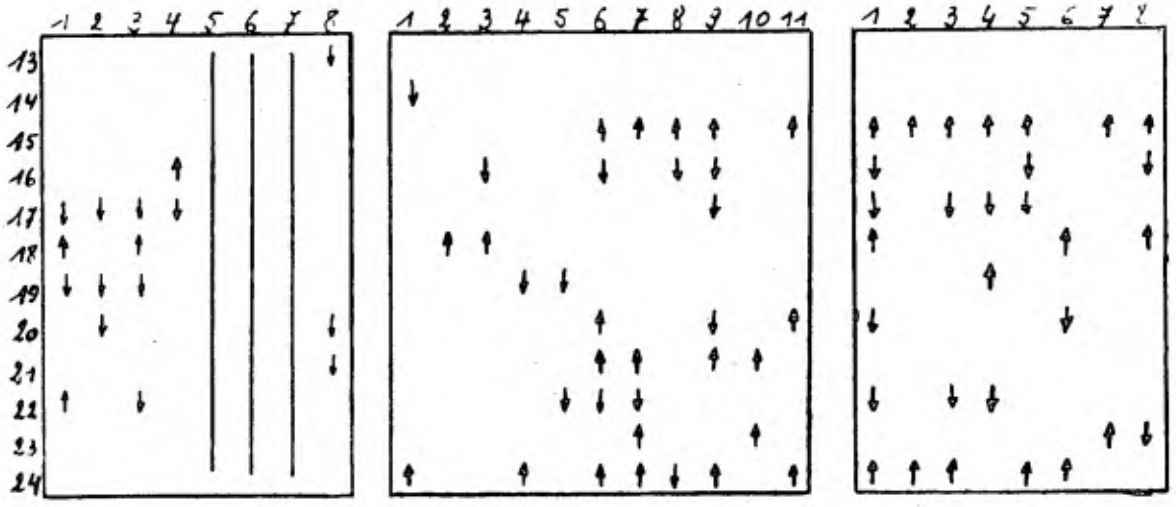
3e subtest (faisant intervenir un référentiel nouveau présentant un grand nombre de graphiques)
Groupe A : 27 rejets négatifs et 3 rejets positifs /96
Groupe B : 8 rejets négatifs et 28 rejets positifs /96
Groupe C : 21 rejets négatifs et 8 rejets positifs /96
Groupe D : 3 rejets négatifs et 38 rejets positifs /96
Ici aussi, l'opposition est également répartie entre les items.
Les groupes B et D sont également compétents, mais le support des exercices (sous forme de graphiques à exploiter) est sans doute enseigné dans les premières années du secondaire.

2. Homogénéité du groupe A.



Le groupe présente une assez grande homogénéité.
Des rejets systématiques apparaissent pour la classe 7 (11/19) dans le sens positif, et pour la classe 8 (8/19) dans le sens négatif, confirmant les observations faites en cours d'expérimentation.

3. Homogénéité du groupe B



Nous observons ici des effets systématiques d'apprentissage pour certaines classes :

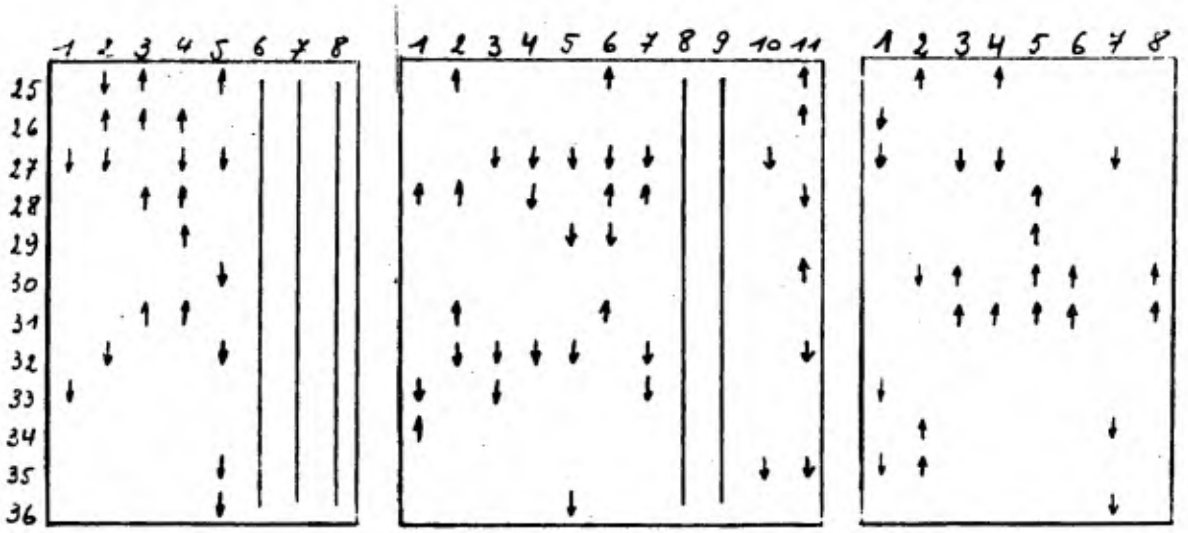
15 : 12/19] pour les épreuves d'entraînement (2 et 3).
24 : 11/19]

Inversément, des carences sont mises en évidence :

16 : 7/19] pour les épreuves d'entraînement (2 et 3).
22 : 6/19]

17 : 8/13 : pour les épreuves avec références variées (1 et 3).

4. Homogénéité du groupe C.



Les effets systématiques sont assez importants dans ce troisième groupe.

Effets positifs :

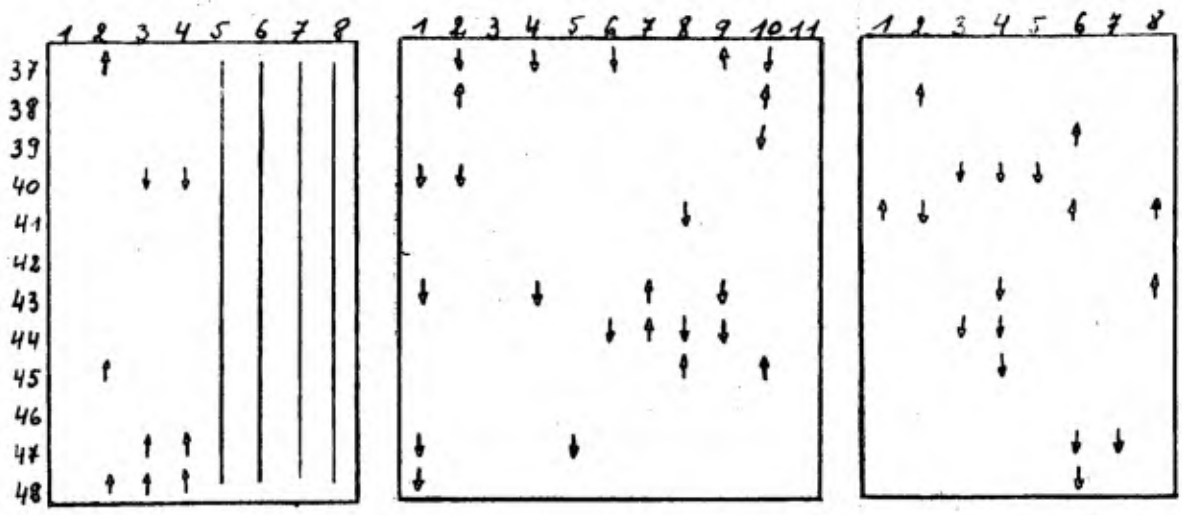
25 : 7/22
28 : 7/22 } pour l'ensemble des épreuves,
31 : 9/22

26 : pour l'épreuve prétestée (3/5).

Effets négatifs :

27 : 14/22 : pour l'ensemble des épreuves
32 : 8/14 : pour les deux premières épreuves.

5. Homogénéité du groupe D.



Bien que ces classes soient prélevées dans des réseaux d'enseignement très différents(écoles secondaires "rénovées" et "traditionnelles"), l'hétérogénéité du groupe est bien répartie sur les classes.

Seule la classe 40 est légèrement plus faible (7 rejets/23).

On observe 19% de rejets, pour 26% au groupe A, 28% au groupe B et 30% au groupe C!

6. Evaluation hiérarchique.

Cette évaluation s'effectue au travers des 27 items qui constituent les posttests.

a) Matrices de concordance.

Groupe A : aucune concordance significative ;

Groupe B : quatre couples interconcordants :
9-10/10-14/14-15/15-18.

Groupe C : un couple : 10 - 23

Groupe D : un couple : 15 - 18

b) Etude des hiérarchies.

C'est bien entendu le groupe B qui a retenu notre attention.

Couple 9-10 : les items 6 et 8se hiérarchisent.

Cependant, leur très faible pourcentage de réussite rend difficile leur interprétation.

Couples 10-14 et 14-15 : Les items 7 et 11se hiérarchisent.

L'item 7 a un faible pourcentage de réussite (4%).

Par contre, l'item 11 a trait aux mêmes activités que celles du critère.

Couple 15-18 : Ces deux items hiérarchisent 12 items autour d'eux. Les items 5,6,7,8 ont un pourcentage de réussite très bas.

Mais 6 de ces items présentent la même particularité que le critère : sans une difficulté spécifique au niveau de la recherche des références, ils ont tous la forme d'un choix multiple (2,12,13,16,17,25,27), la plupart du temps exigeant plusieurs choix pour être corrects Seuls les items 14 et 17 sont ouverts.

Tous deux atteignent un niveau de difficulté élevé (R5 et R6) comme le critère.

3° RECHERCHES D'AMEGAN

Nous reprenons ici l'essentiel des travaux effectués par Michel DEPASSE (1) dans le cadre de son mémoire de licence en Sciences psychologiques et pédagogiques.

A. Objectifs poursuivis par Samuel AMEGAN.

Samuel AMEGAN a été chargé par les autorités du Québec de mettre en évidence quelles pourraient être les différences entre deux structures d'enseignement primaire.

La première présente une didactique plus rigide, plus"ancienne". Il la nomme "conventionnelle".

La seconde est plus récente, plus souple avec les élèves, moins contraignante. Il la qualifie d'"active".

Les questions centrales que se pose Amégan concernent les caractéristiques des unes et des autres.

Il se demande également quelles sont les influences que le type d'enseignement peut avoir sur les élèves. Les capacités des élèves des classes "conventionnelles" ont été comparées à celles des élèves des classes "actives", avant et après une intervention pédagogique susceptible d'entraîner certains comportements créatifs.

---

(1) DEPASSE, Michel
Traitement des données relatives aux effets de deux systèmes éducatifs.
Utilisation de l'évaluation informative.
Mémoire non publié- U.L.B. Juin 1977

B. Tâches choisies et schéma expérimental.

1. Echantillon.

L'échantillon expérimental est constitué de six classes canadiennes équivalant à des cinquièmes primaires de Belgique, choisies selon un critère d'accessibilité.

Trois de ces classes sont "actives", les trois autres étant "conventionnelles".

Un tirage au sort a déterminé dans chaque groupe la classe expérimentale et les deux classes témoins.

a) Les classes expérimentales (2 et 6) reçoivent l'ensemble de la procédure, c'est-à-dire :
- le prétest
- le traitement
- le posttest

b) Les classes témoins 1 et 4 subissent le prétest et le posttest.

c) Les classes témoins 3 et 5 ne subissent que le posttest.

2. Tâches.

1) le prétest se compose de deux épreuves réalisées par TORRANCE (1).

---

(1) TORRANCE E.P. Torrance Tests of Creative Thinking. Princeton, New Jersey, Publication personnelle, 1966

a) Le Torrance figuratif forme A, où l'on demande à l'élève de "montrer son imagination à l'aide de dessins". Le travail de l'élève consiste à achever des dessins constitués de lignes courbes sans signification, ou à réaliser des dessins différents avec, comme point de départ, chaque fois deux verticales...

Cette épreuve comprend trois parties (appelées "jeux"):

- le jeu 1 donne lieu à deux notes, l'une sanctionne l'originalité de la réponse, l'autre l'élaboration de la réponse ;
- les jeux 2 et 3 donnent lieu chacun à 4 notes séparées :
  - la fluidité, c'est-à-dire le nombre de réponses à la question ;
  - la flexibilité, c'est-à-dire la variété dans les réponses ;
  - l'originalité, c'est-à-dire le fait pour une épreuve de ne pas être banale ;
  - l'élaboration, c'est-à-dire le nombre de détails pertinents intervenant dans la réponse (par exemple décoration, ombrage, couleur, ...).

b) Le "Torrance verbal forme A", où l'on demande à l'élève de "montrer son imagination à l'aide de mots". Les tâches consistent à poser le plus grand nombre de questions à propos d'un dessin, à donner le plus grand nombre de raisons qui ont pu amener la scène représentée sur le dessin et à imaginer le plus grand nombre d' usages possibles de boîtes de carton vides...

Amégan a utilisé cinq "jeux" de cette épreuve :
les "jeux" 1,2,3,5,7. Chaque jeu donne lieu à 3 notes : l'une pour la fluidité, la deuxième pour la flexibilité et la troisième pour l'originalité.

2) Le traitement consiste, selon Amégan, en une "sensibilisation des élèves à la production intellectuellle divergente à partir de données académiques et non académiques."
Les questions posées, de type "Guilford production divergente"(1), invitent explicitement les élèves à être divergents et à "sortir des sentiers battus".

3) Le posttest se compose de trois épreuves :

a) Le "Torrance figuratif forme A", comme au prétest.

b) Le "Torrance verbal forme A", auquel on avait adjoint le jeu 4. Les items concernant ce jeu n'ont cependant pas été pris en considération pour pouvoir comparer le prétest au posttest.

c) Le "Guilford production divergente", comporte 21 items de nature plus scolaire. Voici, à titre d'exemple, un des items.
- Donne plusieurs écritures du verbe venir lorsqu'il est conjugué. Fais bien attention à l'orthographe.

---

(1) GUILFORD, J.-P.
The nature of human intelligence. New York.
Mac Graw Hill - 1967

3. Tableau résumé du schéma expérimental.

	Prétest	Traitement	Posttest
	Torrance figuratif Torrance verbal	items type Guilford Pr. diverg.	Torrance figuratif Torrance verbal Guilford pr. divergente
classe 6 active expérimentale	O_1	X	O_2
classe 1 active témoin	O_3		O_4
classe 5 active témoin			O_6
classe 2 convention. expérimentale	O'_1	X	O'_2
classe 4 convention. témoin	O'_3		O'_4
classe 3 convention. témoin			O'_6

O représente une observation.

X représente l'administration du traitement.

C. Objet du mémoire de M. DEPASSE par rapport à la recherche de Samuel AMEGAN.

Après avoir administré les épreuves et récolté les résultats, il fallait les analyser et les interpréter. Il fut convenu qu'Amégan utiliserait les outils statistiques de son choix, tandis que Depasse profiterait des données de l'expérimentation pour mettre à l'épreuve les techniques d'Evaluation Informative.

Il ne s'agissait pas d'opposer les techniques, mais bien de voir dans quelle mesure l'évaluation informative mettrait en évidence des effets collectifs d'apprentissage.

D. Résultats observés par M. DEPASSE.

1. Dichotomisation des données.

La technique d'évaluation informative exige la présentation des données sous une forme dichotomique (1 ou 0).
Aucune des épreuves retenues n'était sous cette forme.
Il a donc fallu effectuer une transformation de la note d'origine.

a) Principes retenus :

- Pour chaque item, déterminer une note qui servira de seuil. Les notes supérieures à ce seuil sont assimilées à une réussite et les notes inférieures à un échec.
- Le seuil choisi pour un item sert pour l'ensemble des classes ayant répondu à cet item.

b) Choix du seuil :

C'est la note moyenne de l'ensemble des notes observées lors de la première passation du test qui a été choisie.

c) Conséquence de ce choix :

La moyenne n'est qu'exceptionnellement un nombre entier.
Elle constitue un paramètre idéal pour séparer deux groupes de notes.
La probabilité pour que les notes dichotomisées présentent un pourcentage de résultats positifs proche de 50% à la première passation est très grande.
Nous observerons donc le plus souvent, soit de la différenciation, soit de l'interaction.
L'évolution lors des passations ultérieures pourra être étudiée.
Remarquons que la dichotomisation fait perdre une partie non négligeable de l'information. Des améliorations ou des détériorations, parfois sensibles, ne sont pas décelées lorsqu'elles se situent de part et d'autre du seuil sans le franchir. Cependant, comme la technique met en évidence des effets collectifs de la didactique et non des progrès individuels, que des effets collectifs se traduisent par une évolution de la moyenne, cet inconvénient n'est pas de nature à affaiblir la pertinence des observations.

2) Evaluation informative.

L'analyse se situant au niveau des classes, c'est la technique proposée au professeur (programme METHRØ9)qui a été employée. Depasse a imaginé de comparer les résultats obtenus aux items dans chaque classe aux prétests et aux posttests en retenant les symboles suivants :

Enquête supérieure : ↗

Différenciation : dif.

Interaction : int.

Enquête inférieure : ↘

a) Etude des résultats globaux

Résultats globaux 1,2,4,6 au Torrance figuratif

	↗	dif.	int.	↘
Prétest	3	13	19	5
Posttest	15	8	15	1

Résultats globaux 1,2,4,6 au Torrance verbal

	↗	dif.	int.	↘
Prétest	4	14	37	5
Posttest	21	14	25	0

Conclusion :

L'amélioration de la performance se marque tant par l'augmentation du nombre d'items d'enquête supérieure que par la diminution du nombre d'items d'enquête inférieure au posttest.

b) Comparaison des classes "actives" et "conventionnelles".

Résultats au Torrance figuratif.

	classes actives 6 + 1	classes convent. 2 + 4	classes actives 6 + 1	classes convent 2 + 4
pré-test	1	2	1	4
posttest	8	7	0	1
différences	7	5	-1	-3
Amélioration observée	17,5%	12,5%	2,5%	7,5%

Résultats au Torrance verbal.

	actives 6 + 1	convent. 2 + 4	actives 6 + 1	convent. 2 + 4
pré-test	3	1	3	2
posttest	14	7	0	0
différences	11	6	-3	-2
amélioration observée	18,33%	10%	5%	3,33%

Conclusions.

L'amélioration des "classes "actives" est légèrement plus importante que celle des classes "conventionnelles".

On remarque cependant qu'au " Torrance figuratif" les classes "actives" n'ont qu'un seul résultat inférieur au prétest contre 4 pour les classes "conventionnelles", ce qui leur enlève des possibilités d'amélioration et met en cause l'égalité des capacités de départ pour ces deux types de classes.

c) Etude des résultats au Guilford production divergente.

Ce test n'ayant pas été administré au prétest, il ne peut faire l'objet d'une comparaison.

Mais il est possible d'opposer les résultats des trois classes "actives" aux résultats des trois classes "conventionnelles" et de les transformer en pourcentages pour l'ensemble du test.

	classes actives 6+1+5	classes convent. 2+4+3	classes actives 6+1+5	classes convent. 2+4+3
notes	14	3	14	20
% de succès ou d'échec	11,11%	2,38%	11,11%	15,87%

Conclusion :

La part de succès des classes "actives"est plus importante et leur part d'échec moindre.

d) Attribution des effets constatés.

Le plan expérimental de S. Amégan présente une carence importante. Il manque, pour rejoindre le dispositif de SOLOMON, un groupe expérimental, celui qui n'aurait reçu que le traitement et le posttest. Ceci a pour conséquence qu'il devient très difficile de neutraliser l'hypothèse rivale de la contamination de l'effet prétest sur le traitement.

La comparaison des résultats d'enquête supérieure aux prétests et aux posttests entre les trois groupes de classes montre que cette crainte est fondée :

Torrance figuratif.

	Classes exp. 6 + 2	classes témoins 1 + 4	classes témoins 3 + 5
Notes au pré-test	0	3	-
Notes au posttest	8	7	0
Différence	8	4	-
%age de progrès	20%	10%	-

Torrance verbal.

	classes exp. 6 + 2	classes témoins 1 + 4	classes témoins 3 + 5
Notes au pré-test	1	3	-
Notes au posttest	13	8	4
différences	12	5	-
progrès en %	20%	8,33%	-

Guilford production divergente.

	classes exp. 6 + 2	classes témoins 1 + 4	classes témoins 3 + 5
nombre de succès	13	2	2

Conclusion.

Le fait d'avoir subi un traitement améliore la performance pour les trois épreuves.

Le fait de subir l'épreuve en prétest améliore aussi les résultats.

La maturation et /ou l'expérience n'est que peu efficace pour le Torrance verbal et inefficace pour le Torrance figuratif.

e) Comportement des items aux tests de Torrance.

Afin de déterminer les aspects des épreuves qui ont donné lieu à un apprentissage, DEPASSE a examiné chaque classe ayant présenté à la fois le prétest et le posttest.

Item par item, il a établi un bilan des améliorations après avoir eu soin de grouper les items selon les rubriques : fluidité, flexibilité, originalité et élaboration.

1. Interprétations proposées.

DEPASSE estime qu'il y a amélioration lorsqu'un item varie entre le prétest et le posttest de la manière suivante :

- enquête inférieure → enquête supérieure,
- enquête inférieure → différenciation,
- différenciation → enquête supérieure,
- interaction → enquête supérieure.

L'évolution d'un item de l'enquête inférieure vers l'interaction n'est pas retenue, car elle n'est pas imputable à la didactique, mais à des facteurs que l'on ne peut mettre en évidence.

Dans les autres cas, il n'y a pas d'amélioration ou il y a détériorisation.

2. Tableau récapitulatif des observations.

La colonnne A indique le pourcentage d'items améliorés pour chaque classe, la ligne B de même, mais pour chacune des rubriques.

	Fluidité							Flexibilité							Originalité									Elaboration			
	figural		verbal					figural		verbal					figural			verbal						figural			
	1	2	1	2	3	4	5	3	4	6	7	8	9	10	5	6	7	11	12	13	14	15		8	9	10	A
6	+	+	+	+	+		+	+			+	+		+	+		+	+		+							56%
1	+	+	+		+	+						+			+			+									32%
2	+	+	+				+		+					+								+				+	32%
4		+	+				+		+																		16%
B	64%							29%							22%									8%			

3. Conclusions.

La fluidité, qui concerne essentiellement l'aspect quantitatif de la production divergente, donne lieu à plus d'amélioration que les autres rubriques. L'aspect le moins entraînable serait l'élaboration, qui touche beaucoup plus la façon d'exprimer une idée et non l'idée elle-même.

La classe 6 (expérimentale et active) s'améliore le plus.

La classe 4 (ni expérimentale, ni active) s'améliore le moins.

Les classes 1 et 2, respectivement expérimentale conventionnelle et témoin active, présentent une amélioration semblable et intermédiaire.

CONCLUSIONS

Comparées aux techniques classiques utilisées dans les recherches que nous avons évoquées, les méthodes d'évaluation informative produisent des résultats très nuancés. Elles autorisent une interprétation plus fine de la qualité du matériel d'évaluation employé et mettent en évidence des effets particuliers liés à des circonstances locales.

Le contexte dans lequel se poursuivaient ces recherches, ainsi que les objectifs visés par les expérimentateurs, ont nécessité l'adaptation de nos techniques, surtout quant à la présentation des résultats enregistrés. Les tableaux qui en résultent n'ont d'autre objet que de faciliter l'interprétation des informations contenues dans les données.

Nous nous sommes limité ici à des commentaires généraux sur les circonstances expérimentales, mais nous nous sommes davantage attaché à l'analyse du matériel supportant l'évaluation, en espérant ne pas avoir trahi ces chercheurs qui ont accepté de nous fournir leurs données.

TABLE DES MATIERES

UTILISATION DES METHODES D'EVALUATION INFORMATIVE DANS DES RECHERCHES EXPERIMENTALES

	Pages
INTRODUCTION	1
1° RECHERCHES DE BAMBA	2
A. Etude de l'homogénéité du groupe expérimental	3
1. Technique XHI^2	3
a) Vérification de l'homogénéité de l'ensemble de la population	3
b) Etude de l'homogénéité des sous-groupes	4
1. Sous-groupe lingalaphone	4
2. Sous-groupe francophone	4
2. Technique arc sinus	5
3. Technique XHI^2 classique	5
B. Etude des épreuves expérimentales	6
1. Contenu des épreuves	6
a) épreuves de vocabulaire (1 et 2)	6
b) épreuve de lecture (3)	6
c) épreuve de calcul (4)	7
d) épreuve de langage oral (5, 6, 7)	7
e) épreuve de compréhension verbale (8)	7
2. Analyse des résultats	7
a) épreuves de vocabulaire	7
- Liste des items du test 1	7
- Liste des items du test 2	8
- Comportement de l'ensemble de la population et des différents sous-groupes	9
- Comportement des classes séparées	10
- Etude des résultats par la XHI^2 classique	11

Pages

b) épreuve de lecture 11
- Liste des items du test 3 11
- Comportement de l'ensemble de la population et des différents sous-groupes 12
- Comportement des classes séparées 12
- Etude des résultats par la XHI^2 classique 13

c) épreuve de calcul 13
- Liste des items du test 4 13
- Comportement de l'ensemble de la population et des différents sous-groupes 13
- Comportement des classes séparées 14
- Etude des résultats par la XHI^2 classique 14

d) épreuve de langage oral 15
- Liste des items du test 5 (production verbale) 15
- Comportement de l'ensemble de la population et des différents sous-groupes 15
- Comportement des classes séparées 16
- Etude des résultats par la XHI^2 classique 17
- Présentation des items du test 6 (correction syntaxique) 17
- Comportement de l'ensemble de la population et des différents sous-groupes 17
- Comportement des classes séparées 17
- Etude des résultats par la XHI^2 classique 18
- Présentation des items du test 7 (maniement oral de la langue) 18
- Comportement de l'ensemble de la population et des différents sous-groupes 18
- Comportement des classes séparées 19

Pages

e) épreuve de compréhension verbale 19
- Liste des items du test 8 20
- Comportement de l'ensemble de la population et des différents sous-groupes 20
- Comportement des classes séparées 22
- Etude des résultats par la XHI^2 classique 23

2° RECHERCHES DE FRYDMAN 24

1) Objectifs de la recherche 24
2) Dispositif expérimental 24
- Représentation schématique du dispositif 25
3) Traitement 26
4) Posttests 27
5) Résultats 27

A. Constitution des échantillons expérimentaux 28
1. Analyse du groupe total 28
- Tableau synthétique des observations sur l'ensemble de la population 29
2. Analyse des sous-groupes 30
a) Etude de la cohérence interne des sous-groupes 30
- Groupe A 30
- Groupe B 30
- Groupe C 31
- Conclusions 31
b) Applications d'autres techniques 31

B. Analyse des résultats du traitement 32
1) Première épreuve 33
2) Deuxième épreuve 34
3) Troisième épreuve 35
4) Quatrième épreuve 36
5) Cinquième épreuve 37
6) Sixième épreuve 38

Pages

7) Evaluation hiérarchique 39
a) Matrices de concordance 39
Groupe A 39
Groupe B 40
b) Etude des hiérarchies 41
8) Application d'autres techniques 42
9) Conclusions 42
C. Analyse des résultats aux posttests 43
1. Homogénéité du groupe total 43
- Interprétation 44
- Comparaison pour chacun des subtests 44
1er subtest 44
2e subtest 45
3e subtest 45
2. Homogénéité du groupe A 46
3. Homogénéité du groupe B 47
4. Homogénéité du groupe C 48
5. Homogénéité du groupe D 49
6. Evaluation hiérarchique 50
a) Matrice de concordance 50
b) Etude des hiérarchies 50

3° RECHERCHES D'AMEGAN 51

A. Objectifs poursuivis par Samuel Amégan 51
B. Tâches choisies et schéma expérimental 52
1. Echantillon 52
2. Tâches 52
1) Le prétest 52
a) Torrance figuratif forme A 53
b) Torrance verbal forme A 53
2) Le traitement 54