

NUMÉRO
SPÉCIAL



BULLETIN INTER IREM N°12

sommaire

EVA

Un colloque sur l'évaluation a eu lieu à Strasbourg du 12 au 14 mars 1976. Il a débuté par une conférence et s'est terminé par une table ronde. Entre temps se sont réunis un certain nombre de groupes de travail.

- PROBLEMES MORAUX ET POLITIQUES DE L'EVALUATION PEDAGOGIQUE 5
- TABLE RONDE 12
- GROUPES DE TRAVAIL 16
 - Docimologie 16
 - Attitudes et opinions 19
 - Evaluation des acquis à l'Ecole Elémentaire 22
 - Réflexion sur des items pour des évaluations centrées sur des critères 27
 - Noyaux 30
 - Objectifs 35
 - Textes 39
 - Signification des comportements de réponse 46
 - Confection et utilisation de grilles d'observation en classe .. 48
- LISTE DES PARTICIPANTS 50

QUATRIEME ANNEE — Juillet 1976
Edité par l'I.R.E.M. de LYON



ADRESSES DES I.R.E.M. EN 1976

Directeurs

GREMILLARD

BESANCON
Faculté des Sciences, La Bouloie -- 25300 BESANCON CEDEX
Tél. : (81) 80.43.60

BORDEAUX
351, Cours de la Libération -- 33405 TALENCE
Tél. : (56) 80.74.42

BREST
Université de Bretagne occidentale, avenue V le Gorgeu -- 29200 BREST - Tél. : (98)
Tél. : (98) 80.16.94 (poste 690)

CAEN OU BASSE-NORMANDIE
I.U.T., Boulevard Maréchal Juin -- 14000 CAEN
Tél. : (31) 81.13.62 (poste 627)

CLERMONT-FERRAND
Complexe Scientifique des Cézeaux, B.P. 45 -- 63170 AUBIERE
Tél. : (73) 92.22.26 (poste 33.97)

DIJON
Faculté des Sciences -- Mirande B.P. 138 -- 21004 DIJON CEDEX
Tél. : (80) 41.00.34 (poste 647.648.691)

GRENOBLE
Tour de Mathématique, Domaine Universitaire, B.P. 41 -- 38401 SAINT MARTIN
D'HERES
Tél. : (76) 54.61.45 (poste 502)

LILLE
Faculté des Sciences et Techniques -- B.P. 36 -- 59650 VILLENEUVE D'ASCO
Tél. : (20) 56.92.00

LIMOGES
U.E.R. Sciences Exactes et de la nature, 123, rue Albert Thomas -- 87100 LIMOGES
Tél. : (55) 77.57.15 (poste 336) - (55) 79.24.12

LYON
Université Claude Bernard -- LYON I -- 43, bld du 11 Novembre 1918 -- 69621
VILLEURBANNE
Tél. : (78) 52.03.37 ou 52.07.04 (poste 37 - 24)

MARSEILLE
U.E.R. de Marseille Luminy, 70 Route Léon Lachamp -- 13288 MARSEILLE CEDEX 2
Tél. : (91) 41.39.40 - (91) 41.15.40 (poste 32 - 10)

MONTPELLIER
Faculté des Sciences, Place Bataillon -- 34000 MONTPELLIER CEDEX
Tél. : (67) 63.42.14 ou (67) 63.91.44 (poste 383)

NANCY-METZ OU LORRAINE
Boulevard Charlemagne, Locaux de l'I.U.T. -- 54001 NANCY
Tél. : (28) 27.38.99

OBJECTIFS DU BULLETIN INTER -- IREM

POURQUOI CE BULLETIN ?

- Disposer d'un outil à la disposition de chacun contenant tous les renseignements pratiques concernant les IREM (adresses, activités, qui ?, où ?, quand ?, etc...).
- Faire le point régulièrement sur *un thème*, entre autre à l'occasion d'une rencontre inter-IREM.
- Etre un lien de critiques, d'échanges et de rencontres.

CE BULLETIN EST DESTINE A QUI ?

- Aux personnes travaillant dans les IREM (animateurs, stagiaires, directeurs, étudiants, vacataires, ...).
- A tous ceux qui en France ou à l'étranger veulent être au courant des activités des IREM.

CE BULLETIN EST DISPONIBLE OU ?

- Dans les IREM (s'adresser au correspondant de l'IREM de votre Académie).
- Dans les centres de documentation qui le demandent.

ECRIVEZ !! On vous publiera.
Envoyer vos articles à l'adresse suivante:

Brigitte SENECHAL
Directeur de l'IREM de CAEN
IREM -- IUT
Boulevard Maréchal Juin
14000 CAEN

NANTES U.E.R. de Mathématiques — Université de Nantes, Boulevard Michelet, B.P. 1044 — 44000 NANTES CEDEX Tél. : (40) 74.50.70 (poste 295)	DHOMBRES
NICE Département de Mathématique, Parc Valrose — 06034 NICE CEDEX Tél. : (93) 84.60.29 (poste 372)	MERIGOT
ORLEANS Université — Domaine Universitaire de la Source — 45045 ORLEANS CEDEX Tél. : (38) 63.22.16 (postes 624 - 638)	ROUCHIER
PARIS SUD Université Paris VII — 2, place Jussieu, Tour 56 — 75005 PARIS Tél. : (1) 336.25.25 (postes 53-83 à 53-86)	REVUZ
PARIS NORD C.S.P. Avenue Jean Baptiste Clément — 93430 VILLETANEUSE Tél. : (1) 820.61.70 (poste 4271)	ROUGEE
PICARDIE U.E.R. de Saint Quentin, 48 Rue Raspail — 02100 SAINT QUENTIN Tél. : (23) 62.54.46	LABESSE
POITIERS Faculté des Sciences, 40, avenue du Recteur Pineau — 86022 POITIERS CEDEX Tél. : (49) 46.26.30 (poste 687) - (49) 46.27.35	BOROWCZYK
REIMS Faculté des Sciences — B.P. 347 — 51062 REIMS CEDEX Tél. : (26) 47.03.13	DAVID
RENNES Université — B.P. 25 A — 35031 RENNES CEDEX Tél. : (99) 36.48.15	GABORIEAU
ROUEN B.P. 27 — 76310 MONT SAINT AIGNAN Tél. : (35) 70.42.73	BUISSON
STRASBOURG Département de Mathématiques, 10, rue du Général Zimmer — 67084 STRASBOURG Tél. : (88) 61.48.20	MARTINET
TOULOUSE Université Sabatier, 118, route de Narbonne, 31400 TOULOUSE Tél. : (61) 52.14.14	DESO

Problèmes moraux et politiques de l'évaluation pédagogique

Conférence de G. GLAESER,
Directeur de l'I.R.E.M. de Strasbourg

En ouvrant ce colloque Inter-I.R.E.M. consacré à l'évaluation en didactique des mathématiques, je formulerais un souhait : que le niveau des débats s'élève au-dessus d'une contestation banale !

La mise en cause de notre système éducatif est un thème inépuisable. La critique des pratiques pédagogiques inefficaces, nocives et injustes est si aisée qu'elle sert d'alibi à tous ceux qui récusent toute recherche didactique, grâce à quelques maîtres-mots : "sélection", "élitisme", etc..., etc... Or l'évaluation est une technique basée sur des recherches et non un thème de polémique littéraire.

Lorsque le contestataire se donne la peine d'argumenter, il exprime souvent une grande part de vérité. Mais la tentation de l'accusation confuse et excessive est très forte. Les arguments contradictoires se mêlent à des contre-vérités. Il est vraiment dommage que le ton s'élève rarement au-dessus de la polémique, sans que l'on se donne les moyens scientifiques de valider ou de nuancer ses accusations.

A titre d'exemple, je citerai le livre de Michel Tort sur "Le quotient intellectuel" dont j'approuve les trois quarts des conclusions. Mais j'y trouve à chaque page des à peu près, des détails bâclés, des sophismes, des refus d'approfondir.

Il faut que ce colloque adopte une méthode qui lui donne les moyens de travailler efficacement sur le détail sans que les débats généraux soient étouffés.

Pour y parvenir, nous proposons l'organisation suivante en trois temps :

1° Je vais essayer de dresser maintenant un inventaire, une classification et une analyse des arguments qui me semblent les plus forts dans la contestation de l'évaluation (et à plus forte raison du système éducatif basé sur la "bachomanie"). Cet exposé contiendra par la force des choses une forte dose des banalités que je prétends dénoncer !

2° Nous demandons aux participants d'éviter d'aborder ces thèmes pendant la seconde partie du colloque consacrée à des problèmes techniques. Ils pourront prendre note des points qui leur semblent importants.

Cependant, s'il est des participants qui préfèrent approfondir les thèmes abordés dans mon exposé, ou préparer un contre-rapport, il pourront constituer une commission improvisée pendant la seconde partie.

3° Une table ronde est prévue dimanche de 10 h à 12 h pour discuter sur les thèmes de mon exposé, sur le contre-rapport (s'il existe) ou sur tous les problèmes annexes que les participants auraient notés pendant la période de "trêve".

I. Nécessité de l'Evaluation

1° Dans une société évoluée, il est des responsabilités techniques qu'on ne peut confier à n'importe qui sans se préoccuper de sa compétence dûment contrôlée.

Je ne puis m'empêcher de frémir en songeant à la "société sans écoles" d'Ivan Illich, où, pour soigner une rage de dents on s'adressera à un volontaire autodidacte qui sera convaincu de sa qualification en art dentaire.

Certes, Illich généralise à partir de situations marginales ; il est plausible que pour alphabétiser telle société d'Amérique Centrale le système en vigueur ne soit pas le plus efficace. Mais de là à confier le pilotage d'un avion long courrier, la construction d'un tunnel ou l'enseignement de la géométrie en quatrième au premier autodidacte venu ... Les arguments d'Illich perdent toute valeur car ils sont excessifs et prétendent à la généralité.

2° On ne peut espérer de progrès pédagogique si l'on ne dispose d'aucun outil objectif pour trancher les controverses. Nous sommes las des bavardages de dilettantes où Monsieur X prétend que "les élèves" ne comprennent pas la géométrie, Monsieur Y affirme le contraire, sans qu'ils ne proposent aucun moyen pour jeter quelques lumières dans le débat.

II. Les critères ne doivent pas être choisis empiriquement

Ayant admis que la compétence n'est pas toujours un mythe, on voudrait choisir rationnellement des critères de compétence. Ce n'est pas toujours le cas.

Exemples :

Il était admis naguère que la connaissance du latin était un préalable à toute étude de médecine.

Sous l'ancien régime, on pouvait être instituteur de village sans savoir lire et écrire ; le critère décisif était la recommandation du curé.

De nos jours, sous les régimes les plus variés, on attribue de nombreuses responsabilités techniques sur des critères d'orthodoxie idéologique, et on les refuse sur des critères racistes ou sociaux.

Actuellement, en France, on prétend que pour enseigner les mathématiques dans le Premier Cycle du second degré il suffit d'un baccalauréat (éventuellement littéraire) et de quelques années d'études supérieures (dans n'importe quelle spécialité !).

Notre ministre vient de décider qu'un inspecteur d'Académie n'a pas besoin d'avoir enseigné !

On peut rejeter toutes ces absurdités au nom du bon sens et constater que ces lubies ne sont pas innocentes. Combien de vocations n'a-t-on pas irrémédiablement découragées à l'aide de barrières artificielles, alors que toutes les véritables conditions du succès étaient réunies.

L'établissement de "profils" correspondants à chaque qualification est un problème que l'on peut aborder scientifiquement.

III. La "bachomanie" et les "épreuves-sans-critères"

Notre système éducatif est subordonné aux examens, aux concours. Les usages empiriques qui en découlent imprègnent toute la pédagogie motivée par l'obtention des diplômes.

L'incohérence, l'irrationnalité, l'injustice et l'"infiabilité" de nos institutions, dans ce domaine, ont été abondamment prouvées depuis 50 ans. Sans ressasser à nouveau la liste des aspects kafkaïens de notre distribution de peaux-d'âne, signalons simplement que les diplômes explicitent rarement ce qu'ils prétendent prouver. Exemple d'exception : tel brevet de natation assure que le lauréat a parcouru 500 mètres dans des conditions explicitées. C'est une information restreinte mais précise.

Mais, que savons-nous — en toute certitude — d'un bachelier ? Assurément il sait lire, écrire et compter. Je vous laisse le soin de trouver s'il est possible d'en dire plus.

Que peut-on affirmer à propos d'un agrégé de mathématiques ? Il a certainement poursuivi pendant plusieurs années un certain type d'études. Il sait faire une démonstration et il a résolu beaucoup de problèmes non triviaux.

Mais cela, on le savait déjà lors de la candidature à l'agrégation (sous réserve qu'il ne s'agisse pas d'une candidature fictive).

De plus, on est certain que l'agrégé a beaucoup d'autres connaissances, d'aptitudes, etc... Mais lesquelles ?

La disproportion entre la lourdeur de ces institutions et la pauvreté des renseignements qu'ils fournissent est aveuglante : le coût de cette maigre information est exorbitant.

On ne peut plus tolérer que l'on continue à soumettre des masses de candidats à des épreuves sans que l'on ait des idées claires de ce que l'on prétend mettre en évidence.

Comment se fait-il que l'institution des examens et concours continue à inspirer tant de confiance à quelques "bons esprits" ?

C'est peut-être à cause du curieux phénomène suivant qu'il convient d'analyser.

"On constate que quelques candidats brillants réussissent simultanément à tous les concours auxquels ils se présentent, même dans les cas où les épreuves ont été notoirement défectueuses. N'est-ce pas une preuve de la fiabilité du système ?"

L'explication est simple : un candidat qui possède beaucoup de compétences ou d'aptitudes sera reçu à l'Ecole Polytechnique parce que les épreuves exigeaient certaines qualités (inconnues des organisateurs du concours) et à l'Ecole Normale parce qu'il fallait d'autres qualités (tout aussi mystérieuses).

La masse des candidats pourvus d'un profil plus éclectique joue son avenir à la "roulette russe".

Que peut-on dire d'un candidat qui a brillamment réussi au concours d'entrée à l'Ecole des Mines et échoué aux Ponts et Chaussées ?

Osera-t-on prétendre que cela prouve qu'il est plus apte à prospecter des sous-sols qu'à gérer le réseau routier ?

IV. Les tests empiriques

Mais si le fonctionnement des examens et concours est notoirement défectueux, beaucoup de tests à prétention scientifique ne sont pas à l'abri des critiques.

Il était inévitable, qu'à l'origine, on établit des batteries de tests au "pifomètre", et qu'on essaya de les étalonner a posteriori.

Ainsi, la boutade d'Alfred Binet (avant 1910) : "L'intelligence c'est ce que mes tests mesurent" était à la rigueur une réponse d'attente légitime.

Mais depuis les travaux décisifs de Spearman et du développement de l'analyse factorielle on dispose d'outils puissants pour analyser des tests. Et d'ailleurs cette tâche est systématiquement menée à bien dans de nombreux laboratoires.

Il n'en est que plus inadmissible que des dilettantes fassent joujou avec des tests. C'est aussi dangereux que de jouer avec des médicaments. Tout chercheur qui aborde l'évaluation a pour devoir d'assimiler la critique de Spearman et d'apprendre les moyens d'éviter les défauts catalogués depuis plus de cinquante ans.

La méthode des tests n'est légitime qu'avec des épreuves soigneusement étalonnées, construites pour mettre en jeu des critères parfaitement connus.

V. La sélection sociale et l'autre

Conçues pour tester certaines aptitudes, les épreuves reflètent généralement tout autre chose : elles défavorisent systématiquement quelques couches de la population qui sont culturellement sous-équipées ou qui sont réfractaires à des modes intellectuels discriminatoires.

L'obligation d'étudier le latin a longtemps joué le rôle du suffrage censitaire et barré l'accès des trésors culturels au prolétariat. De nos jours, on essaie de promouvoir sous l'appellation abusive de "Mathématiques Modernes" une espèce de charabia ésotérique qui remplit la même fonction.

Dans l'analyse sociologique des mécanismes de cette sélection les inégalités économiques n'interviennent qu'au second degré. Les enfants privilégiés sont avant tout ceux qui trouvent des livres, des disques, des revues dans le milieu familial. Ce sont ceux qui ont

beaucoup parlé avec leurs parents avant l'âge de 2 ans. L'échelle de ces privilégiés n'est pas exactement la même que celle des classes sociales.

Dans la classe ouvrière, on ne peut confondre l'environnement culturel des manoeuvres légers avec celui d'un horloger, d'un typographe ou d'un instituteur.

Les enfants d'un boutiquier cossu sont souvent très défavorisés. Par contre en tête des privilégiés culturels on trouve les enfants issus de l'intelligentia universitaire, et non les futurs capitalistes. La discrimination se fait non seulement en fonction du milieu familial, mais aussi selon l'équipement socio-culturel de l'habitat.

Pressé d'étayer la thèse — largement justifiée — d'une discrimination sociale scandaleuse dans l'utilisation du Q.I., Michel Tort ne se donne pas la peine d'examiner ces nuances ...

Selon les besoins de l'interprétation de ses statistiques, il annexe les petits épiciers à la haute bourgeoisie ou aux travailleurs. Il lui est facile de mettre en parallèle une école "bourgeoise" du XVIème arrondissement et un baraquement scolaire en bordure d'un bidonville. Mais il ne se donne pas la peine d'examiner les variations que l'on peut enregistrer au sein d'un même milieu socio-culturel. Les différences considérables de performances entre divers enfants de travailleurs émigrés ne s'expliquent pas uniquement par des facteurs sociaux.

Notre système d'examen privilégie un type de culture "de luxe" au détriment de toutes les acquisitions culturelles que l'on développe au contact du travail intelligent. S'il n'en était ainsi, les enfants des oisifs seraient désavantagés par rapport à ceux qui vivent dans l'observation de l'exercice d'une profession.

Ainsi la sélection sociale met souvent en évidence des aptitudes artificielles, et elles défavorisent d'autres aptitudes très importantes dont les privilégiés sont naturellement dépourvus. Ainsi s'établit le portrait-robot du "bon élève" dans le cadre de notre système scolaire. Il ne correspond pas toujours à la plus forte personnalité et comporte une forte dose de conformisme et de docilité.

A partir de cette critique — partiellement légitime — la contestation déraile à pieds joints vers l'excès contraire ! Le bon élève n'existerait pas !

Le talent, l'intelligence, l'imagination, la curiosité seraient des mythes, et prétendre le contraire c'est sombrer dans l'élitisme ! C'est ainsi qu'on préconise un égalitarisme médiocre et une éducation à base de "bandes dessinées, de téléche, de Guy Lux et de Mireille Mathieu". On nous invite à promouvoir un enseignement pour béotien, qui ne doivent jamais s'élever au-dessus des marguerites. Lorsqu'on aura constaté que les enfants vomissent ce type d'activité scolaire, il ne restera plus qu'à prêcher la "société sans école".

Il est pourtant clair que l'école et l'Université sont des institutions très ambiguës. Elles sont à la fois des instruments de sélection et de promotion, et c'est le milieu idéal pour former des contestataires. C'est des rangs de l'Université (et non dans l'armée) que sont sortis les Karl Marx, les Lénine, et plus près de nous les Baudelot et les Establet !

Monsieur Prudhomme dirait que l'Université sert à défendre nos institutions et au besoin à les combattre !

Face à la sélection sociale ou raciale, hautement condamnable, il existe une sélection par le mérite, le talent, l'originalité. Si c'est vraiment le meilleur qui gagne, je n'y trouve rien à redire. Il existe une pédagogie de l'effort où l'on demande à chaque élève de dépasser constamment ses possibilités du moment. Cette pédagogie s'adresse à tous, aux élèves doués et aux moins doués. Elle leur propose des tâches adaptées à leur niveau. L'évaluation correcte des résultats obtenus est un puissant stimulant qui permet à chacun de donner le meilleur de lui-même. A ce titre, la véritable mathématique est un puissant outil de promotion : elle permet à chacun de donner la mesure de ses capacités intellectuelles.

Dans cette optique, nous devons fournir des outils d'auto-évaluation à nos élèves pour leur permettre de suivre leurs progrès, pour encourager leurs efforts sans démagogie ni complaisance naïve. A la limite, cet outil devrait être suffisamment crédible pour décharger l'enseignant de l'obligation de rendre des oracles !

Il faut apprendre à nos élèves à fournir des réponses correctes et non des réponses qui font plaisir au maître.

VI. Déontologie de l'évaluation en pédagogie et en recherche

Chaque enseignant doit être conscient de l'impact de son évaluation. Mettre des notes, corriger des devoirs, interpréter les résultats chiffrés du travail d'une année n'est pas un acte innocent,

et il peut engager tout l'avenir des élèves. Dans le système actuel, il en est de même de la correction des examens et des concours.

Il est particulièrement grave que la préparation des épreuves d'examen soit généralement confiée à des personnes qui n'ont pas reçu de formation spéciale pour cette activité et que des raisons d'opportunité administrative tendent à laisser étouffer toute critique sur la valeur des sujets ; il serait souhaitable que les I.R.E.M. passent outre et organisent une évaluation systématique des sujets d'examens et de concours, ce qui conduira peut-être à redresser la situation.

Chaque chercheur doit être très exigeant devant ses conclusions et se livrer au doute méthodique au sujet de ses propres travaux. Il faudrait en finir avec la pratique qui tolère l'emploi de méthodes que l'on sait être critiquables, mais que l'on persiste à utiliser par paresse intellectuelle. Nous savons depuis 50 ans que la description des aptitudes ne peut se faire selon un modèle totalement ordonné, que deux individus classés ex aequo dans un concours ne sont pas interchangeables, qu'il est malhonnête de chiffrer les notes au quart de point près, que, si l'on ne prend pas certaines précautions draconiennes les résultats des corrections ne sont pas fiables, etc... etc...

Puisque nous savons tout cela, il faudrait bien que nous cessions de nous prêter à cette mascarade ridicule en jouant avec l'avenir des jeunes, en nous prenant au sérieux. C'est pour préparer l'abolition de ce jeu dangereux qu'il faut que nous préparions des outils éprouvés, et c'est pour contribuer à cette tâche que nous essayons de travailler dans ce colloque.

Table ronde

Rapport établi par F. PLUVINAGE

La séance plénière de fin de colloque avait été initialement intitulée "table ronde". Nous lui avons conservé cette dénomination, bien qu'en fait elle en ait différé par sa forme.

Le présent rapport est articulé autour des thèmes abordés, et ne respecte donc pas complètement l'ordre chronologique des interventions. Une ligne en pointillés sépare des interventions qui avaient été entrecoupées, en séance, par des interventions sur un autre thème.

L'utilisation de l'évaluation

Il y a danger d'utilisation des recherches en évaluation. Il n'y a pas réellement de décalage temporel entre recherche et utilisation : dès le départ, le chercheur doit être conscient des risques d'utilisation (Latreille).

Il se pose un problème de récupération au niveau des enseignements ; et l'utilisation est impossible à empêcher (Buisson).

Quelques exemples de mauvaises utilisations :

- Questionnaires utilisés par des personnes ignorantes des conditions de la recherche dont ces questionnaires étaient un instrument (Robin).
- Utilisation de recherche de Piaget à des fins d'évaluation (J. Méténier).

.....

En fait, les qualités d'une évaluation n'ont *pas* toujours une *grande importance* vis-à-vis de son utilisation voulue :

les instances dirigeantes se préoccupent peu des travaux étudiés à ce colloque (Schaeffer).

.....

La pression économique joue avant tout : lorsqu'il y a plus de gens que de postes on a besoin d'un critère pour choisir parmi des candidatures a priori toutes valables. Mais le moyen de sélection est sans importance, il faut simplement pouvoir dire : "Les résultats de telle épreuve conduisent à retenir les trois candidatures suivantes et à écarter les autres" (Buisson).

L'évaluation et l'évalué. Aspect négatif (exemples) :

- Elève doué en dessin *collé* à cause des maths (Schaeffer).
 - Jeune fille n'ayant pu devenir puéricultrice à cause de son ignorance du volume du cylindre (Méténier).
-

Rôle ambigu des mathématiques

D'une part *sélection répressive*, d'autre part *planche de salut* (Glaeser).

Est-ce que "planche de salut" est à entendre comme dans ce titre de la presse : "Le fils de la concierge entre à Polytechnique" ? (Barbançon).

Parmi les candidats brillants aux olympiades, certains sont issus de milieux "modestes" (Glaeser). Il y a des possibilités de *promotion par les maths*. (Glaeser).

Les "qualités" de la sélection par les maths

"La sélection par les maths est à la fois équitable et efficace" (Citation d'un recteur par Mme Méténier).

Pour ce qui est des maths, on joue sur la conviction répandue qu'il est possible d'attribuer à chacun l'une des étiquettes Bon ou Mauvais (Mme Noël).

En définitive, on n'évalue jamais *que* le niveau de connaissances en math.

Et compte tenu des travaux actuels, on peut facilement récuser le caractère de "pureté" attribué à la sélection par les maths (Latreille).

.....
Il existe des aptitudes en mathématiques (Glaeser).

L'évaluation et les enseignants

L'un des principaux résultats, pour les enseignants, de la pratique d'une évaluation sérieusement conduite est *l'autoformation* (Robin).

Tout enseignant doit-il être formé à la recherche (en évaluation) ? (Méténier).

Il y a lieu de distinguer *formation* et *sensibilisation* à la recherche. La nécessité est celle d'une compréhension par les enseignants des travaux de recherche (Bonniol).

Un enseignant doit être capable de procéder à certaines *observations* (Robin).

Une distinction peut être faite entre l'utilisation et l'élaboration des outils d'observation (Schaeffer).

Une activité à promouvoir est la pratique par les enseignants d'une attitude de recherche : le questionnement sur ce que l'on fait soi-même (Barthel).

L'attitude d'observation, de questionnement est cataloguée comme recherche pour les uns, mais pas pour les autres (Méténier).

"Je suis frappé par l'inertie, le découragement de beaucoup d'enseignants : "On ne possède pas des vérités qui existent et que les spécialistes connaissent". Nous devrions faire comprendre aux enseignants qu'ils peuvent nous poser des problèmes" (Tourneur).

Evaluation et recherche

Le terme de "recherche" est à préciser, car il peut faire illusion. Penser à la pédagogie comme à quelque chose d'analogue à l'étude du pétrole ou d'un métal, c'est méconnaître le *caractère "réactif" de l'objet d'observation*. Le cadre de pensée doit s'inspirer de la psychologie clinique, et le rapprochement à faire pour la recherche pédagogique est avec la recherche opérationnelle (Cardinet).

Il y a une distinction à faire entre pédagogie et évaluation : de manière constante, la pédagogie prend en compte d'une part des connaissances mais aussi des options personnelles de type idéologique ou politique. Tandis qu'il ne m'est pas nécessaire, par exemple, d'adhérer aux opinions de Bonniol pour reconnaître ses résultats (Pluinage).

.....
"Je ne suis pas convaincu que l'évaluation est une science" (Buisson).

La science, c'est la docimologie. L'évaluation est :

- une *technique* au mieux,
- un *art* parfois,
- une *saloperie* au pire" (Bonniol).

Global ou local ?

Dans le groupe "objectifs", une question a été : "Qu'évalue-t-on ?"

En fait on évalue toujours *un tout*. Et il importe de se préoccuper de l'étude des fonctionnements effectifs d'une évaluation (Latreille).

Si par exemple, on se propose une évaluation d'un enseignant, les élèves y tiennent une place : il est impossible d'évaluer l'efficacité d'un enseignant sans l'observation des réactions des élèves. On peut se centrer sur une analyse globale, tout en gardant en vue un objectif terminal à poursuivre (De Ketele).

.....
On peut dissocier *l'observation* et *l'expérimentation*.

L'observation est globale, commande ses méthodes et aboutit à des hypothèses. L'expérimentation n'a de valeur que locale, est commandée par les méthodes et aboutit à des conclusions (sous réserve du contrôle de certains paramètres).

Une évaluation déterminée doit d'abord être située dans l'une de ces deux catégories pour qu'il soit possible de se prononcer sur ses vertus (Pluvinage).

Objectifs d'évaluation et projets des évaluateurs

De Ketele a parlé de la poursuite d'*objectifs* d'évaluation. Mais n'est-ce pas le *projet* de l'évaluateur qui tient le plus grand rôle ?

Or ce projet (qui peut aller jusqu'à s'opposer aux objectifs exprimés) ne se laisse pas aisément mettre à jour (Duval).

Cette mise au jour devrait être faite (De Ketele).

Seule l'observation des *phénomènes de surface* est possible. Et ensuite se pose tout le problème de l'interprétation (Duval).

Pour moi, la question essentielle est de définir les objectifs de la recherche. On ne mettra jamais à jour toutes les variables mais l'opérationnalisation des objectifs constitue une contribution fort importante (De Ketele).

Les évaluations procèdent par sommation de comportements individuels. Or l'enseignement fait intervenir des phénomènes de groupe. Une de ses caractéristiques essentielle est l'*interaction* professeur-élèves. *Comment contrôler* une intervention à ce niveau par des évaluations d'objectifs ? (Duval).

Ce qui est fondamental est de *communiquer* les objectifs, de manière à ce que *la critique soit possible* (De Ketele).

Docimologie

Rapport de C. DUPUIS et J. METENIER

Travaux présentés et discutés

Marie-Claire Dauvisis (I.R.E.M. de Toulouse) a décrit une célèbre expérience de multicorrection faite à Montauban sur un sujet style B.E.P.C. Les 20 évaluateurs, stagiaires d'un groupe I.R.E.M. ont composé le sujet, discuté d'un barème et corrigé 24 copies tirées au hasard parmi beaucoup. Pourtant les écarts entre

les notes d'une même copie peuvent être énormes et aucune procédure de normalisation n'a permis de dégager une certaine cohérence entre les correcteurs.

Référence :

La rigueur du professeur de Mathématiques et la notation par J. Cransac et M.-C. Dauvisis — Bulletin de l'Association des Professeurs de Maths. de l'Enseignement Public n° 300, Septembre 75 (A.P.M.E.P., 29, rue d'Ulm, Paris 5ème).

A l'I.R.E.M. de Strasbourg, l'expérience de Toulouse a été reprise en essayant dans un premier temps de raffiner le barème avant une nouvelle expérience de multicorrection : ce procédé n'a pas augmenté la fiabilité de la correction. Dans un deuxième temps, une tentative de codage des comportements de réponse a conduit à un échec ce qui a amené à analyser plus soigneusement le sujet proposé ...

Référence :

A propos d'une expérience de docimologie par C. Dupuis et F. Pluvinage. Bulletin de l'Association des Professeurs de Mathématiques de l'Enseignement Public n° 303 (1976).

J.-J. Bonniol a présenté au groupe les résultats d'une recherche menée par le GROUPE de RECHERCHE en DOCIMOLOGIE D'AIX sur les évaluations des examinateurs au baccalauréat.

Manipulant deux variables expérimentales à deux modalités :

- la connaissance ou la non-connaissance du livret scolaire des candidats par les examinateurs lors des épreuves orales,
- la connaissance ou la non-connaissance par les mêmes examinateurs de la note qu'ils avaient attribuée aux candidats lors de l'écrit,

il met en évidence que les corrélations entre notes de l'année, notes d'écrit et notes d'oral sont bonnes,

- qu'elles ne sont pas meilleures lorsque l'examineur connaît le livret scolaire, ce qui suggérerait que le livret scolaire n'est guère pris en compte en réalité,
- qu'elles sont meilleures lorsque l'examineur connaît la note qu'il a lui-même attribuée à l'écrit ce qui suggère une tendance à ne pas se déjuger lors de l'oral quand il possède l'information sur sa première évaluation.

Référence :

Les résultats de ces travaux seront publiés ultérieurement.

Discussion

1° *La vraie note et la normalité*

On est à peu près sûr que l'hypothèse selon laquelle, quelle que soit la copie, les notes attribuées par plusieurs correcteurs sont distribuées suivant une loi de Laplace-Gauss est déraisonnable, et que, par ailleurs, il ne faut pas espérer que la dispersion observée diminue par la seule multiplication des correcteurs.

Par contre (hypothèse à vérifier) on peut espérer que cette dispersion diminuera si on observe un certain nombre de précautions dans la préparation de l'épreuve :

- préciser les objectifs de chaque question,
- multiplier les questions portant sur une même notion en prévoyant des recoupements.

2° *Les critères de l'évaluateur*

L'interrogation directe de l'évaluateur ne donne pas d'indications fidèles sur ses méthodes et ses critères d'évaluation. Diverses procédures ont été proposées :

- ne pas donner la copie de l'élève à l'évaluateur que l'on veut observer mais à une autre personne. L'évaluateur questionne cette dernière pour parvenir à noter la copie
- construire des copies fictives de façon à mieux contrôler les variables intervenant
- éliminer des facteurs tels que les fautes d'orthographe ou l'écriture en dactylographiant les copies.

Par qui ces expériences peuvent-elles être pratiquées ?

Si une prise de conscience des divergences de notation est indispensable pour motiver les professeurs à la réflexion sur les problèmes posés par la notation, de telles expériences peuvent provoquer des réactions affectives profondes. Il faut donc veiller à leurs modalités et prendre en charge les problèmes liés à de telles remises en cause.

Prolongements

Les travaux présentés sont poursuivis et développés par leurs auteurs.

Attitudes et opinions

par Régis GRAS (Rennes) et Françoise KITTEL (Strasbourg)

Rapport de B. GILLET et D. BOISNARD

Exposé de R. Gras

Témoignages d'analyses d'enquêtes faites à l'entrée et à la sortie de classes de quatrième. Il s'agissait de connaître les opinions des élèves vis-à-vis des mathématiques et de mesurer les glissements qui pouvaient s'opérer au cours de l'année.

Le traitement de ces analyses s'est fait :

- par calcul de pourcentage et établissement de corrélations ;
- par une analyse factorielle de correspondances et établissement d'une classification hiérarchique.

I. Contenu du questionnaire

1° *Entrée en quatrième*

L'enquête est du type : "Aimez-vous les mathématiques ?". Elle porte sur 200 élèves.

Le genre de questions est du type :

- "Es-tu convaincu par :
- . une figure soigneuse ?
 - . une affirmation par un élève ou une personne qui te semble plus douée ?
 - . une démonstration rigoureuse ?
 - . des connaissances antérieures ? "

"Tes résultats en cinquième étaient : très bons - bons - ..."

2° *Fin quatrième*

Le questionnaire est destiné à faire une analyse comparative entre les réponses à l'entrée et à la fin de la quatrième.

Il porte essentiellement sur :

- . l'intérêt porté par les élèves à la matière ;
- . les facteurs de réussite ;
- . les méthodes pédagogiques.

3° Discussion

Les remarques des participants à la séance ont concerné :

- la formulation des questions (forme positive ou forme négative. Cf. enquête à Strasbourg) pour obtenir des résultats cohérents ;
- le vocabulaire (un psychologue aidait les élèves à comprendre les questions) ;
- l'établissement du questionnaire (il est nécessaire de poser de nombreuses questions pour en isoler quelques unes, regrouper d'autres ou fournir de nouvelles variables). Cf. les travaux de nos collègues suisses qui ont d'abord testé des questions avant de faire un questionnaire ;
- il manque aussi des questions essentielles du genre : quand un élève éprouve-t-il le besoin d'une démonstration ?

II. Les résultats

Ils portent sur la deuxième enquête car ils sont plus importants pour l'étude des glissements qui se sont opérés en cours d'année. Les résultats ont été calculés en pourcentages et ont permis les remarques suivantes :

- Fixation plus nette de l'attrait, du goût *pour ou contre* les mathématiques observées dans les classes expérimentales.
Exemple : "Aimez-vous les mathématiques ?" Parmi ceux qui avaient répondu "un peu", on a noté un glissement vers "beaucoup" ou vers "pas du tout".
- Les réponses sont en "attente" dans les classes témoins, c'est-à-dire les classes traditionnelles : préoccupation d'examen, de passage dans les classes supérieures.
- "Qu'attendez-vous des mathématiques ?" Les classes témoins réaffirment leurs préoccupations d'examens, de passage, alors que les classes expérimentales voient l'utilité, les connaissances, les métiers.
- La part du professeur semble plus importante dans les classes expérimentales que dans les classes traditionnelles.
- "Que reprochez-vous aux mathématiques". Le nombre des élèves des classes expérimentales qui avaient répondu "rien" est resté au même niveau (22 - 21 %). Par contre, le nombre des élèves des classes témoins qui avaient répondu "rien", est descendu de 31 à 18 %.

III. Classification hiérarchique pour obtenir une typologie des élèves

Deux grandes classes se dégagent :

- Les "Bons élèves". Ils attendent beaucoup des adultes, du professeur, des parents, mais ils sont en général immatures. Ils se partagent en cinq catégories :
 - ceux qui ne pensent qu'au professeur et au travail mais qui ont peu d'amour pour les maths,
 - ceux qui sont écrasés par le professeur, qui n'apprécient pas les maths mais aiment l'éducation physique et les travaux manuels,
 - ceux qui sont dominés par leurs parents et dont le but est la réussite,
 - ceux qui se disent bons élèves et visent la seconde de leur choix,
 - les "fades", qui font des maths comme n'importe quelle autre matière, n'ont ni envie, ni goût particulier.
 - Les "Contestataires". Ils refusent l'adulte mais se raccrochent aux groupes, aux camarades ou, au contraire, s'en détachent par une personnalité beaucoup plus affirmée (en "bien" ou en "mal").
- Quelques sous-catégories :
- les fatalistes, n'aimant ni les maths ni les professeurs mais appréciant les groupes et les camarades,
 - ceux qui aiment les maths mais refusent les professeurs en général. Ils attendent des mathématiques une promotion,
 - les littéraires,
 - les réfractaires, les rébarbatifs, ceux qui en général ont beaucoup de difficultés un peu partout.

Analyse d'opinions par la technique d'entretien

Exposé de F. Kittel

Un entretien est une conversation sérieuse ayant un but bien déterminé. Cet entretien et les résultats fournis par un questionnaire devraient se compléter.

- Ses fonctions sont :
- de recueillir des données,

- d'informer,
- d'agir sur les sentiments, les comportements,
- de motiver.

On peut distinguer deux sortes d'entretiens :

- l'entretien clinique, visant à un diagnostic (l'interviewé est au centre de l'entretien),
- l'entretien psychosocial, visant à connaître des opinions (le chercheur ou un thème est au centre de l'entretien).

Buts d'un entretien

- connaître ce que l'on veut étudier, un sujet ou une population,
- permettre la possibilité de rédiger un questionnaire, d'en définir les formulations,
- permettre d'atteindre ou d'essayer d'atteindre un autre niveau d'analyse. En particulier atteindre le *conscient* ou l'*inconscient*, ce qui laisse apparaître des oppositions, par exemple :
 - souhait avoué de participation mais passivité inconsciente,
 - faire des études pour leur utilité mais en réalité en éprouver du plaisir (culpabilité du plaisir).

Evaluation des acquis à l'Ecole Elémentaire

Rapport de M. N. AUDIGIER

I. Résumé des travaux présentés

Section élémentaire de l'I.N.R.D.P. (J. F. Richard, J. Colomb, M. N. Audigier)

Esquisse de la problématique et de la méthodologie de la recherche qui a pour titre "étude du rôle des moyens d'expression dans l'apprentissage mathématique".

Cette recherche a débuté en 71 et a plusieurs points d'application aux divers niveaux de l'école élémentaire : par exemple au CP une étude sur les représentations ensemblistes.

Cette évaluation concerne environ 50 classes à chaque niveau du CP au CM2.

Du point de vue méthodologique un certain nombre de variables sont prises en compte, notamment la progression (ce mot recouvre ici l'enchaînement des notions mathématiques et les activités pédagogiques qui y conduisent, très précisément décrites).

Il faut noter que cette progression est appliquée dans les classes une année avant l'expérimentation : on peut ainsi espérer que la progression est à peu près assimilée par les maîtres. Il a été précisé qu'il ne s'agissait en rien de trouver la "meilleure" progression. Définir une progression apparaît simplement comme une contrainte expérimentale permettant d'observer des choses qu'on ne pourrait pas observer sans cela.

Cette évaluation présente donc la caractéristique suivante : elle se place dans un contexte pédagogique qu'elle se donne les moyens de connaître.

I.R.E.M. de Lyon (R. Charnay)

Mise en place cette année d'une évaluation des acquis en fin de CM2, dans la perspective d'une liaison CM2-sixième aux deux niveaux (élèves - enseignants).

Plus précisément, à partir du programme 70, une liste d'attente a été définie en termes de savoir-faire (exemple : savoir faire une addition, trouver le suivant d'un nombre). Une série d'items est construite pour permettre de mesurer le degré de réalisation de ces attentes.

Ces mêmes items seront passés en début de sixième et quatrième.

Recherche Inter-I.R.E.M. minicalculateur (E. Cauzinille)

Il s'agit d'une évaluation, au niveau de la quatrième, de l'apport des calculateurs programmables dans l'enseignement de l'algèbre.

Cette évaluation concerne 50 classes et a débuté depuis deux ou trois ans. A partir du programme d'algèbre, une nouvelle progression a été élaborée. Cette progression a été mise en oeuvre dans 25 classes utilisant un minicalculateur et dans 25 classes ne l'utilisant pas.

La comparaison de population a été faite par l'intermédiaire de tests mensuels portant sur :

- le contenu mathématique de façon globale,
- plus précisément : la notion de variable et constante.

Institut de Recherche Pédagogique de Genève (M. Hutin)

Evaluation qui se situe dans le cadre d'une réforme de l'enseignement portant sur 30 000 élèves (1 000 enseignants) du canton de Genève et programmée sur 68-78.

Cette réforme a conduit à l'élaboration de trois types d'instruments :

- progression mathématique (au sens enchaînement des notions mathématiques) ;
- moyens d'enseignement ;
- instruments de formation des maîtres.

L'évaluation a lieu à cinq niveaux :

- épreuves de contrôle élèves ;
- impact de l'enseignement ;
- étude approfondie de type Piagetien sur 40 ou 50 sujets ;
- efficacité du recyclage des maîtres ;
- en commun avec les autres cantons : évaluation de l'introduction du nouvel enseignement par le biais :
 - . d'un questionnaire aux enseignants,
 - . d'épreuves mesurant les effets de l'enseignement.

I.R.E.M. de Strasbourg (R. Duval)

Deux sortes de travaux sont menés :

1° Evaluation du développement du raisonnement entre 11 et 16 ans en particulier étude des décalages entre le raisonnement logique et le raisonnement mis en oeuvre dans des situations où interviennent des mécanismes physiques.

2° Observation d'une cohorte d'élèves sur une durée de 4 ans (sixième à quatrième) sur un certain nombre d'activités (en marge du programme) ; le choix des activités est lié au fait
. que ces activités intéressent les élèves,
. que ces activités permettent aux élèves de s'organiser et communiquer l'information.

II. Discussion

Un certain nombre de questions ont été soulevées :

- Envisager l'évaluation comme outil pédagogique (moyen pour l'enseignant d'évaluer son travail).
- Comment évaluer une formation de type professionnel (par exemple, celle qui est donnée dans les Ecoles Normales) ?
- Qu'appelle-t-on évaluation des acquis ? Est-ce qu'il ne faut évaluer que les acquis à l'école élémentaire ?
- Qu'est-ce une progression ?
- Que peut-on évaluer ? Pourquoi ? Comment ?

Seuls un certain nombre de points ont pu faire l'objet d'un réel échange :

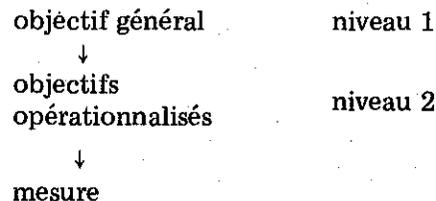
1° Qu'est-ce qu'une progression ? Pourquoi la fixer ou non, selon le type d'évaluation ?

Il semble que l'on puisse se placer dans deux cas :

- On cherche à faire un constat de ce qui se passe à l'instant t ; alors il suffit d'avoir fixé l'enchaînement des notions mathématiques, le choix des activités y conduisant peut être laissé aux maîtres.
- On cherche à faire un constat à l'instant t , mais en plus on veut tenter d'expliquer les résultats, les différences obtenues. Alors, il semble qu'il faille non seulement fixer l'enchaînement des notions mathématiques, mais encore le type d'activités et leur ordre d'introduction.

2° Notion d'opérationnalisation des objectifs

Le schéma classique suivant a été mis assez radicalement en question :



Ce schéma conduit à conclure que l'on a évalué l'objectif de niveau 1. Or les résultats que l'on obtient portent sur les objectifs de niveau 2. On ne peut donc conclure qu'à ce niveau : on a évalué les objectifs opérationnalisés.

Pour aller plus loin, il faudrait être sûr que les objectifs opérationnalisés rendent compte de l'objectif général — et précisément ce point n'est jamais assuré.

Exemple : Si on se fixe comme objectif de niveau 1 l'apprentissage de la multiplication. On peut traduire cet objectif en terme de comportements observables (objectifs opérationnalisés) en décidant qu'il faut qu'un élève connaisse la technique, sache faire des problèmes relevant d'une situation multiplicative, etc... Mais peut-on en conclure, s'il a les comportements attendus, qu'il possède la notion de multiplication ? Certainement pas. On peut simplement dire qu'il sait faire une multiplication de x chiffres par y chiffres — qu'il sait résoudre tel type de problème.

Il découle de cette perspective qu'on ne peut pas évaluer des objectifs très généraux (on ne peut pas évaluer une notion mathématique, la relation maître-élève, etc...).

3° Qu'est-ce qu'il est intéressant d'évaluer ? Il semble qu'on puisse distinguer deux types d'évaluation :

- 1 Les évaluations centrées sur l'élève : elles ont pour résultat de distribuer les élèves sur une échelle (de notes par exemple). Ce sont des évaluations ayant un aspect normatif.
- 2 Les évaluations centrées sur le système pédagogique, les conditions pédagogiques.
Il s'agit alors de donner des informations sur les conséquences de telle pratique pédagogique.

Les participants du groupe ont rejeté la perspective 1 comme inintéressante et posant beaucoup de problèmes.

Par contre, il leur a semblé intéressant de parler d'évaluation dans le cas 2 et ils se sont donc placés dans cette perspective pour introduire à nouveau la distinction suivante :

a) On définit un comportement attendu et on mesure la distance entre le comportement attendu et le comportement observé. Mais il a été souligné que la notion de comportement attendu présente certaines caractéristiques qui sont apparues à certains comme négatives.

— Elle suppose la définition d'une *norme* qui est fixée *a priori* par l'évaluateur.

— Elle appauvrit l'observation dans la mesure où il y a focalisation de l'observation et de l'analyse sur ce comportement attendu.

— Elle pose le problème de l'adéquation de l'épreuve à ce qu'elle prétend mesurer.

b) On se propose de comparer des comportements observés dans des situations pédagogiques différentes. Dans ce cas, on ne définit pas d'ordre sur les comportements, il n'y a pas d'aspect normatif, l'évaluateur n'a pas à choisir un comportement optimal.

A l'issue de telles comparaisons, on pourra dire, quand on propose aux élèves telle ou telle activité, qu'ils réussissent à telle épreuve, échouent à telle autre, etc... L'évaluation est alors un outil de description de l'éventail des comportements obtenus dans certaines conditions pédagogiques.

Il se peut *qu'a posteriori* on puisse hiérarchiser ces comportements.

Il faut souligner que ces deux optiques a) et b) ont des conséquences méthodologiques très importantes.

Il est apparu à certains participants que la voie b) pouvait représenter une nouvelle façon intéressante d'envisager l'évaluation en matière pédagogique ; c'est un type d'évaluation qui dispense l'évaluateur de définir a priori des normes de comportement.

Bibliographie

- Compte rendu du Colloque sur la Didactique — I.R.E.M. de Bordeaux. Article : Problématique et méthodologie de la Recherche.
- Numéro de "Recherche Pédagogique" à paraître en avril 76 au SEVPEN Mathématique au Cours Préparatoire : compte rendu d'une expérimentation.
- Revue Française de Pédagogie à paraître : article concernant la recherche "Etude du rôle des moyens d'expression dans l'apprentissage mathématique".
- Publications du Service de la Recherche Pédagogique de Genève notamment : Thèse de Monsieur Hutin.

Réflexion sur des items pour des évaluations centrées sur des critères

Rapport de P. MOUCHE et J. BOROWCZYK

Travaux présentés

- 1° Expériences d'évaluation continue de mathématiques dans les 9 classes de seconde du Lycée Expérimental de Saint-Quentin par M. Misery.
- 2° Cube de J. Cardinet (Institut Romand de Recherche et de Documentation Pédagogique, Neuchâtel).

Résultats des travaux

1° Dans les 9 classes de seconde (constituées après épreuves de classement), des épreuves normalisées ont lieu toutes les six

semaines en mathématiques pour évaluer le profil des élèves (connaissances ; savoir-faire) et lui permettre de se situer par rapport à lui-même.

Chaque épreuve comporte 12 items de 8 points chacun ; pour chaque item et pour l'ensemble de l'épreuve, on définit :

- a) Le coefficient de réussite d'un ensemble d'élèves, c'est-à-dire le pourcentage d'élèves ayant atteint l'objectif fixe évalué par l'item.
- b) La moyenne des notes obtenues.
- c) L'écart-type σ .
- d) Le coefficient de sélectivité : différence des coefficients de réussite des ensembles d'élèves ayant une note respectivement supérieure à $m + \frac{\sigma}{2}$ et inférieure à $m - \frac{\sigma}{2}$.

On observe que plus le coefficient de réussite à un item est faible plus le coefficient de sélectivité est élevé, ce qui pénalise les élèves faibles.

Un étude plus précise des modalités de questionnement a permis de diminuer l'écart entre les deux coefficients (introduction de questions d'échauffement non cotées influençant les réponses ultérieures).

Il a semblé que deux finalités étaient mêlées dans cette étude :

- savoir ce qu'une classe ou un groupe d'élèves sait,
- orienter l'élève à partir des connaissances acquises et de sa capacité d'apprendre. Or, il est difficile d'utiliser un même item pour deux objectifs différents.

2° Jean Cardinet à l'Institut Romand de Recherche et de Documentation Pédagogique Neuchâtel présente alors un CUBE de données permettant une classification de l'Univers des items en fonction de trois buts d'évaluation :

- différencier les classes : connaissance des acquis de la classe (pour fixer les points d'ancrage pédagogiques) ;
- différencier les élèves : classement des élèves (groupes de niveau ou individuellement) ;
- différencier les stades d'apprentissage (seuils franchis pour mesurer les progrès accomplis dans le temps).

Un item ne peut desservir qu'un de ces objectifs.

Pour différencier les *classes* (gestion pédagogique), on doit disposer de questions conduisant à des réussites (ou échecs) par tous les élèves d'une même classe, ce qui permet de savoir si une notion est bien assimilée ou non.

Pour différencier les *élèves*, les questions devront avoir un fort pouvoir discriminatoire.

Enfin d'autres tests seront surtout sensibles aux *stades d'apprentissage* afin de marquer les progrès accomplis dans le temps.

Les tests de Q.I. par exemple comportant surtout des questions d'intelligence générale sont insensibles aux progrès accomplis lors de l'apprentissage.

3° Enfin, on a signalé des travaux en cours :

à l'I.R.E.M. de Nancy, QCM ouvert pour classes de seconde et terminale ;

à Mons, *Dépoer* étudie l'influence de différents aspects de l'environnement de la question : données superflues, ordre de présentation, action présentée au passé ou au présent, grandeur des nombres, longueur de l'énoncé ...

Références bibliographiques

- 1 — Rapport sur les activités d'évaluation 74-75 (Lycée Henri Martin, 02100 St-Quentin).
- 2 — Vers une évaluation formative en Mathématiques (I.R.E.M. de Picardie, 18 p.).
- 3 — Premières conclusions tirées de l'expérience d'évaluation continue en mathématiques dans le cadre des secondes différenciées (Lycée Henri Marin, LHM 1.76.1, 20 p.).
- 4 — Jean CARDINET : L'élargissement de l'évaluation, Institut Romand de R.D.P. 75, 02 avril 75.
- 5 — J. BERBAUM : L'action pédagogique dans l'enseignement du second degré. Nathan. 1971.
- 6 — GILBERT : Mathetic.
- 7 — LEE, J. CRONBACH, GOLDINE, C. GLEESER, MARINDER, NANDA ... The dependability of behavioral measurements theory of generalizability for scores and profiles (Wiley).

8 — DEPOVER, Thèse Université de Mons (à paraître). Etude de différents aspects pour poser des questions.

Principales questions soulevées

Les questions ne peuvent mesurer que certains aspects de l'éducation.

Par exemple : classement d'élèves *ou* notions acquises ou non dans une classe *ou* stade de l'apprentissage ; la question elle-même devient sujet d'étude.

En mathématiques, franchissement de *seuils* nécessitant un matériel d'évaluation adapté.

Noyaux

Rapport de Madame PEYRIERAS

Les travaux de trois équipes de recherche ont été présentés :

1 *Mlle Chelly* de l'I.N.R.D.P. et *Mlle Malcuit*, professeur enseignant, ont parlé des travaux d'évaluation entrepris dans le cadre de la recherche des C.E.S. expérimentaux.

Les hypothèses de base sont :

- La prise en charge de tous les élèves y compris ceux de type III.
- La recherche d'une meilleure individualisation de l'enseignement ; la structure en groupes de niveau étant un des moyens par lesquels on espère y parvenir.

Des conditions de travail ont été prises :

- Travail en équipe de professeurs avec élaboration des progressions
élaboration des modes d'approche
élaboration des contrôles.
- Formation de groupes sur critères multiples et mobilité des élèves entre les groupes.
- Pédagogie différenciée adaptée aux groupes et pédagogie de soutien.

On a été amené à évaluer les résultats de l'apprentissage

- A partir de contrôles faits par les professeurs dans l'exercice habituel de la classe.
- En analysant ces épreuves à l'aide de grilles sur lesquelles n'apparaissent que des codes :
 - . un code qui rend compte du contenu. (On se réfère à une analyse détaillée du programme. Ce programme est un impératif institutionnel) ;
 - . un code qui rend compte du langage (transcodage) ;
 - . un code qui rend compte du comportement ;(dans ces deux derniers cas, on se réfère au dictionnaire des objectifs généraux).

En ce qui concerne les enseignements qui se dégagent de ces travaux,

- . le fait d'avoir intégré les élèves de type III n'a pas fait baisser le niveau de la population concernée,
- . ce qui serait un programme minimum commun serait constitué par les mécanismes, ce qui est décevant.

Interventions et questions

— M. Glaeser souhaite qu'il y ait une large circulation de l'information afin d'éviter aux nouvelles équipes de recherche de repasser par les mêmes tâtonnements.

— Les groupes de niveau-matière sont-ils devenus des groupes-classes ? Non — l'unité de base est de 72 ou 96 élèves répartis en sous-groupes mobiles de niveau de 24 qui ne sont pas nécessairement les mêmes en mathématiques, français, langues. Les élèves travaillent en groupes hétérogènes ou semi-hétérogènes pour les autres activités.

— Vous avez défini des objectifs propres à la matière et des objectifs généraux. Seuls les objectifs propres à la matière ont pu être atteints.

— A-t-on déterminé des seuils ? Envisage-t-on de définir le noyau ? Lequel ?

On a noté une grande convergence pour les seuils de blocage par exemple :

- . échec des élèves faibles au plan de la formulation verbale ;
- . échec également dans l'intégration de deux notions apparemment maîtrisées séparément.

— Est-ce qu'il ne s'est pas posé de problèmes psychologiques lors de changements de groupes ?

Mlle Malcuit a donné toute assurance sur ce point ; cela se passe bien dans 90 % des cas.

— Mme Pelnard a apporté une information à propos des mécanismes, qui ressort de la comparaison des résultats de deux échantillons d'élèves de CM2 : l'un extrait des classes primaires dans lesquelles on a introduit les mathématiques dites modernes, l'autre extrait des classes soumises au programme rénové. Des élèves brillants du secteur expérimental ont été mis en difficultés faute de posséder certains mécanismes que les élèves du traditionnel maîtrisaient parfaitement.

Il a été fait remarquer que ce qui est parfois appelé mécanisme recouvre des niveaux de comportement très divers.

Autres questions

Interdisciplinarité ?

Des projets de structure en cours d'élaboration devraient permettre un travail interdisciplinaire.

Bibliographie

1° Protocole de St Quentin.

2° Numéros 41 et 58 de "Recherche Pédagogique" et un numéro sous presse : les mathématiques en sixième et cinquième dans les C.E.S. expérimentaux.

2 *Monsieur Denis*, professeur à Saint-Etienne, Animateur à l'I.R.E.M. de Lyon, avec une équipe de pédagogues et de psychologues a proposé un modèle d'établissement qui va être concrétisé à la rentrée 77 : C.E.S. expérimental Jules Vallès de la Ricamarie (banlieue de St-Etienne), C.E.S. 600 associé à une S.E.S.

La population de la région est :

- . ouvrière à 70 %,
- . menacée de chômage donc soumise à la reconversion,
- . composée de nombreux immigrés.

La localité n'a pas de C.E.S. Le projet est un cours d'élaboration depuis 10 ans.

Les objectifs premiers sont :

- . ne pas léser les élèves,
- . leur donner, en plus, un apprentissage de responsabilité et d'autonomie ;

c'est un projet total qui intègre :

- les structures,
- le fonctionnement interne,
- l'environnement.

Ce C.E.S. n'est qu'une partie d'un ensemble qui inclut le C.E.C. (Centre éducatif et culturel). Au point de vue des structures pédagogiques, on se refuse tout cloisonnement. Les élèves sont distribués de façon aléatoire en unités de 120 élèves pris en charge par huit adultes professeurs de disciplines différentes, chacun responsable d'un groupe de 15 élèves appelé groupe de vie.

A l'entrée dans l'établissement sera fait un bilan :

- . bilan en terme de connaissances,
- . bilan psychomoteur,
- . bilan de sociabilité.

On constituera des groupes d'intérêt avec mobilité dans une même unité. Des groupes de niveau-matière apparaîtront suivant les besoins.

On ne gardera pas le cheminement linéaire dans le programme actuel.

L'élève prend en charge son processus d'apprentissage, il avance à son rythme et selon ses motivations et on lui fournit les moyens de s'autoévaluer.

La base du C.E.S. est la créativité.

M. Denis applique ses principes pédagogiques dans sa classe de troisième dont les élèves sont recrutés selon un mode traditionnel.

Question

— N'y a-t-il pas un problème au niveau de la motivation ?

Réponse

— Elle doit partir des disciplines dites d'éveil.

Autre question

— Au niveau de l'individualisme et de la sociabilité ?

Réponse

— Il faut exploiter la dynamique de groupe.

Pris par le temps nous n'avons pas pu développer la discussion.

Bibliographie

- Avant-projet du C.E.S. Expérimental Jules Vallès de la Ricamarie. Publication de l'OCCE - Loire - Ecole des Ovides 1, Rue des Ovides, 42100 St-Etienne, sur les processus d'apprentissage du C.E.S. Expérimental de la Ricamarie par A. Denis — Document I.R.E.M. Lyon.
- Tout document sur Carnet d'Observation, test, etc... à demander à OCCE, Loire.

3 Mme Pelnard-Considère, chef de travaux au service de recherche de l'I.N.O.P., associé au C.N.A.M., travaille au laboratoire de psychologie différentielle à Paris dirigé par M. Reuchlin. Elle poursuit une recherche en liaison avec l'I.N.R.D.P.

Il est difficile de résumer le contenu de la recherche de Mme Pelnard. Elle a élaboré, avec une équipe d'enseignants et de psychologues, des épreuves qui vont de la fin de la dixième à la fin de la quatrième. (Épreuves standardisées de raisonnement mathématique).

Mme Pelnard a suivi en gros la progression du programme. Elle a essayé de donner des formes aussi proches que possible, aux différents niveaux, aux exercices correspondant à une même notion. De plus, on a inséré la même question à des épreuves destinées à deux ou trois degrés différents. Une unité de construction se retrouve dans l'ensemble des épreuves. Des analyses factorielles faites, il ressort avec évidence *qu'un seul facteur général* explique la part commune de variance des notes dans chaque groupe. Dans la mesure où les contenus sont mathématiques, ce facteur prend une double signification de facteur de raisonnement et d'intégration des connaissances mathématiques. Il peut être vu comme mesurant la "pensée mathématique naturelle"; la pensée naturelle se définit comme une pensée logique qui utilise des "blocs d'information" provisoirement indissociés; ce processus en lui-même nécessaire au développement intellectuel de l'enfant présente dans l'apprentissage des mathématiques un aspect bénéfique par les supports conceptuels qu'il apporte, mais doit être constamment remis en question et reanalysé au cours de l'apprentissage. C'est cet aspect qui constitue une des causes majeures de difficultés pour certains enfants. La construction du raisonnement mathématique ne peut se faire que sur des contenus eux-mêmes structurés et précisés dans leur conceptualisation par

un apprentissage qui tient à la fois compte du niveau de développement logique et de ce qu'il est convenu d'appeler les mécanismes de base. Il est impératif de ne proposer à chaque enfant que ce que son niveau de maturation lui permet d'assimiler, ce qui a pour corollaire la nécessité d'adapter l'enseignement sans pour autant introduire des systèmes de ségrégation.

Le temps imparti n'a permis aucune discussion sur ce sujet.

La bibliographie est portée sur le document :

"Le développement de la pensée mathématique de la dixième à la quatrième", J. Pelnard-Considère ; J. Levasseur.

(Document ronéoté, Laboratoire de psychologie différentielle, 41, rue Gay-Lussac, 75005 Paris).

Objectifs

Rapport de A. DENIS

1 Les premiers travaux relatés dans ce groupe furent ceux que nos collègues belges ont réalisés depuis 1970. (MM. NOEL, TOURNEUR et Mme NOEL).

L'analyse porte sur des objectifs exclusivement d'ordre cognitif, issus des éléments du programme. Le but poursuivi est d'amener les professeurs à définir clairement les objectifs ou l'objectif visé par l'instrument de mesure classique que sont les questions aux élèves. En termes de contenus, un objectif terminal est choisi pour définir *une unité didactique*. A l'intérieur de cette unité des objectifs intermédiaires conduisant à la compétence terminale sont définis et organisés à l'aide d'un structurogramme non forcément linéaire. Au départ de chaque unité sont situés les items prérequis pour accéder aux connaissances contenues dans cette unité. Toujours situé dans le cadre d'une évaluation d'apprentissage se trouve un *test diagnostique* avec grille de correction pour autocorrection permettant une utilisation plus individualisée du *livret d'exercices correctifs* contenant des exercices programmés à utiliser en séances de rattrapage. Terminant la séquence d'étude d'une unité se trouve un *bilan corrigé*, lui, par le professeur et ne portant presque exclusivement que sur les items de l'unité étudiée nécessaires à l'étude de la séquence suivante. Tous ces documents peuvent être consultés dans les deux livres suivants.

1° "Recherche en éducation"
"Effets des objectifs dans l'apprentissage"

2° "Pédagogie et Recherche"
que l'on peut se procurer auprès de la Direction Générale des Etudes, 26, Boulevard de Berlaimont, 1000 Bruxelles.

Un débat suivit la présentation de ces travaux.

2 M. De Ketele nous présenta ensuite *une typologie* de l'évaluation s'inspirant des travaux de M. Cardinet.

Pour définir une évaluation, quatre dimensions sont choisies dans l'ordre :

QUOI ?

Objet de l'évaluation.

POURQUOI ?

Quel est le but visé par l'évaluation ? En vue de prendre quel genre de décision ?

QUI EVALUE ?

Une personne différente de l'objet ou de la personne évaluée ou non ?

COMMENT ?

Quelques grands types de "méthodes".

Chaque dimension contient un certain nombre de modalités non indépendantes l'une de l'autre (cf. Annexe).

Cette typologie permet de concrétiser la définition suivante :

Une évaluation est un processus permettant d'identifier, de délimiter, d'obtenir et de communiquer toute information permettant un jugement, un choix et une décision parmi les alternatives diverses en relation avec les objectifs poursuivis.

Une remarque de M. Tourneur a permis d'insister sur le fait qu'à chaque finalité correspond un type de questions posées ou d'informations à recueillir d'un même cursus.

M. De Ketele a insisté sur le fait d'avoir recours à plusieurs types d'évaluation.

3 Monsieur Hutin du Service de Recherche Pédagogique de Genève a exposé ensuite les travaux qu'il a été amené à effectuer dans le cadre de la rénovation de l'enseignement des mathématiques.

Les travaux ont porté sur l'école élémentaire. Les objectifs à atteindre et pour lesquels ont été construits des instruments d'évaluation ont été dans l'ordre :

- 1° Raisonnement logicomathématique (analyse de type piagetien).
- 2° Construction d'un langage mathématique.
- 3° Développement et acquisition de stratégies.
- 4° Invention.
- 5° Autonomie.
- 6° Connaissance mathématique.

L'apprentissage par conflit — à partir de l'analyse des erreurs commises — a été mis en avant.

L'exemple de test à 12 questions auxquelles il ne faut répondre qu'à 10 d'entre elles a été mentionné pour évaluer la formation délivrée par certains types d'apprentissage.

Annexe

Typologie de l'évaluation — Modalités définies sur chaque dimension. (M. De Ketele)

I Quoi ?

1. Elèves
2. Enseignants
3. Programme ou Curricula (contenant aussi les démarches mentales)
4. Techniques
5. Matériel
6. Méthodes
7. Système

II Pourquoi ?

1. Classer à l'intérieur d'une population (Evaluation normative)
2. Classer en sous-groupes pour différenciation des stratégies
3. Trouver des variables critères permettant de prédire la réussite
4. Sélection (fonction du projet initial)
5. Déceler les facteurs d'apprentissage
6. Hiérarchisation

7. Inventorier les objectifs de perfectionnement en vue d'une maximalisation de l'apprentissage de type individualisé (tendant vers une courbe en J) ou global (courbe de Gauss)
8. Bilan de maîtrise des objectifs
9. Diagnostiquer les causes d'erreurs pour promouvoir des stratégies de rattrapage
10. Décision d'orientation s'inscrivant dans un processus continu.

III Qui évalue ?

Cinq modalités de type hétéroévaluations :

1. Autorités externes
2. Personnes indépendantes
3. Autorités internes
4. Par des pairs
5. Par des subordonnés

et une modalité n° 6 : Autoévaluation à considérer comme un objectif fondamental de l'apprentissage.

IV Comment ?

1° *Evaluation ponctuelle classique empirique sans objectif explicite*

1. Examen écrit
2. Interview libre
3. Mise en situation libre

2° *Evaluation ponctuelle centrée sur des objectifs explicites*

4. Tests centrés sur des objectifs
5. Interview centré sur des objectifs
6. Mise en situation

3° 7. Analyse des contenus (a priori et a posteriori)

4° *Evaluation en cours de processus*

8. Observation libre occasionnelle
9. Observation systématique
10. Observation provoquée

Textes

Rapport de F. PLUVINAGE

Deux sujets ont été abordés dans ce groupe :

1. L'analyse de sujets d'examen, essentiellement le bac.
2. La lisibilité de textes mathématiques destinés à des élèves.

Concernant le premier sujet, un travail est en voie d'achèvement par un groupe de l'I.R.E.M. de Strasbourg. Il pourrait donner lieu prochainement à un article dans le Bulletin A.P.M. Par ailleurs, les échanges écrits suscités par la question et centralisés par Daniel Reisz pour l'A.P.M. ont été signalés au groupe.

Le deuxième sujet a fait l'objet d'un article dans le Bulletin A.P.M. n° 301 (déc. 75). Des compléments bibliographiques ont été indiqués et les prolongements possibles examinés sommairement.

Références bibliographiques

1. Sujets d'examen

Classification NLSMA

. exposée par Wilson J.W. dans : Bloom-Hastings—Madaus "Handbook on formative and summative evaluation of student learning" Mc Grax Hill Book, 1971,
. résumée par Tourneur Y. dans Math & Paedagogia, n° 56 et 57 (1972).

G. de Landsheere, "Evaluation continue et examens — Précis de docimologie", 3e édition, Labor, Fernand Nathan (1974).

B. Arnaud, "L'épreuve de mathématiques du Bac. C' C, Lyon 1975", Bulletin A.P.M. n° 301, (1975), p. 730.

A. Tortrat, "Sur le problème de T.C., Paris, Juin 1975", Bulletin A.P.M. n° 302 (1976), p. 115.

Régionale parisienne de l'A.P.M.E.P., "Rapport de la commission "Examens et docimologie"", Bulletin A.P.M. n° 302 (1976), p. 121.

2. Lisibilité

Robert Kane dans The Math. Teacher (1970).

Colette Laborde, "Un langage de prononcé de formules en math.",
Revue Française de Pédagogie (1975).

R. Duval et F. Pluvinaige, "Une grille d'analyse et son application à
deux manuels scolaires", Bulletin A.P.M. n° 301 (1975),
p. 709.

Compléments sur la lisibilité et prolongements éventuels

Un indice de lisibilité a surtout une valeur *comparative et relative*.

Exemples :

— Prenons un texte quelconque. Calculons son indice de lisibilité. Amputons ensuite ce texte d'une phrase sur deux au hasard. Le nouveau texte aura un indice de lisibilité voisin de celui du texte initial, mais aura toute chance de n'être pas compréhensible.

— En amputant chaque phrase d'un texte de son adjectif ou de son adverbe le plus long, on améliore son indice de lisibilité. Facilite-t-on par là-même sa lecture ?

Les travaux actuels ne nous permettent pas de dire que, si un texte obtient tel indice de lisibilité, il est accessible à tel pourcentage d'élèves d'un niveau donné.

Mais entre deux textes traitant des mêmes sujets et usant d'un vocabulaire similaire, on peut dire que celui dont la lecture est la moins perturbée par des phénomènes du type "bruit de fond" est celui qui a le meilleur indice de lisibilité.

Et si l'on désire faciliter la lecture d'un texte en le remaniant, il est intéressant de connaître cette notion de lisibilité. On peut ainsi savoir sur quels paramètres jouer.

Exemple :

a) *Texte initial*

"Cet aspect d'image (de la géométrie) n'est aucunement inné, ainsi que le montrent les développements algébriques formels auxquels se prêtent certains élèves du second cycle, et doit être soigneusement entretenu en faisant bien la distinction : une figure est (expérimentalement) convenable pour représenter une situation mathématique". (Bulletin A.P.M. n° 301, p. 774).

b) *Texte remanié*

"Cet aspect d'image de la géométrie n'est pas inné : il faut voir à quels développements algébriques formels se prêtent certains élèves du second cycle. On doit donc l'entretenir soigneusement. Et ce faisant, il faut bien distinguer :

- la situation mathématique,
- la figure (expérimentalement) convenable pour représenter cette situation".

(Note : le rapporteur a pris la liberté de choisir cet exemple pour illustration. En fait, aucun exemple n'avait été indiqué lors de la séance. A titre d'essai, le lecteur peut se livrer au même exercice sur le paragraphe qui suit, dans le Bulletin A.P.M., le paragraphe cité. Il se rendra compte qu'un essai d'amélioration de la lisibilité peut conduire à mettre d'autres défauts en évidence).

En plus de l'indice de Flesch (utilisé dans l'article cité du Bulletin A.P.M.), il existe d'autres indices comme celui de Shale. Cet indice est contesté par R. Kane (dans l'article indiqué en bibliographie) ; la pondération des termes utilisés dans une phrase paraît en effet étrangère à l'idée de "bruit de fond". Et c'est bien celui-ci que l'indice de Flesch permet d'apprécier (comparativement).

En mathématique, l'utilisation de formules symboliques pose une question sur l'emploi de l'indice de lisibilité. En effet, il apparaît qu'une formule peut revêtir des aspects différents selon sa familiarité pour le lecteur et son rôle dans le texte. Si elle est déchiffrée, il faut lui attribuer l'indice correspondant à sa lecture à haute voix. Mais elle peut aussi être l'objet d'une perception globale instantanée, un peu comme une figure qui est immédiatement identifiée. Seule l'observation de réactions individuelles peut permettre d'obtenir des éléments autorisant à conclure.

Prolongements possibles

Ce qui paraît actuellement le plus intéressant, c'est d'entreprendre une étude expérimentale sur la lisibilité. Celle-ci pourra prendre en compte avant tout les réactions des élèves. Selon le niveau auquel une telle recherche est menée, elle peut être entreprise par un enseignant dans sa classe ou par une équipe spécialisée (avec tous les cas intermédiaires possibles).

A propos des sujets de mathématique au baccalauréat

Une équipe de l'I.R.E.M. de Strasbourg a entrepris un travail d'analyse de sujets de bac. Ce travail porte sur les sujets de math. des séries C et D, proposés ces quatre dernières années dans l'académie de Strasbourg.

La procédure d'analyse utilisée a été la suivante :

- 1° Repérer chaque question effectivement posée au candidat (il n'y a pas toujours coïncidence avec la numérotation de l'énoncé en "questions" : une "question" de l'énoncé peut comporter plusieurs questions effectives).
- 2° Repérer les *faits spécifiques*, autrement dit les connaissances atomiques, qui interviennent dans la résolution de chaque question.
- 3° A partir de la détection des faits spécifiques, repérer la couverture du programme scolaire par le sujet : quels sont les paragraphes de ce programme concernés par les différentes questions.
- 4° Toujours à partir de la détection préalable des faits spécifiques, déterminer le niveau NLSMA de chaque question.

L'un ou l'autre exemple de détermination sera indiqué dans l'article relatant ce travail. Il est cependant bon d'insister ici sur le point de départ constitué par les faits spécifiques. C'est lui qui permet une certaine fiabilité dans la détermination du niveau : on peut voir par exemple si l'énoncé de la question mentionne ou non lesquels utiliser, ce qui correspond à la discrimination des catégories A et B d'une part, C et D d'autre part.

5° Un regroupement de certains niveaux NLSMA permet de dresser des histogrammes parlants. S'agissant d'examen, les regroupements suivants ont été effectués :

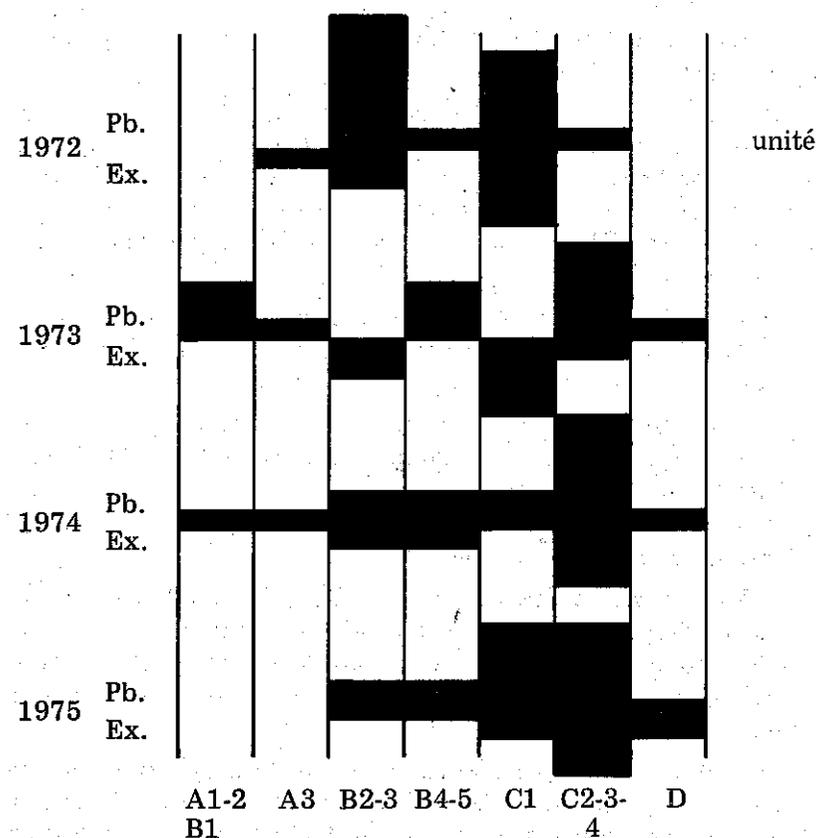
- . A1-2, B1 connaissances élémentaires,
- . A3 calcul n'utilisant que des connaissances élémentaires,
- . B2-3 connaissance de structure,
- . B4-5 changement de formulation,
- . C1 applications routinières,
- . C2-3-4 interprétation, comparaison, reconnaissance de formes,
- . D découverte.

On obtient ainsi sept classes bien hiérarchisées.

A titre d'exemple, nous donnons ici, les histogrammes obtenus pour les sujets de la série C, de 1972 à 1975. L'unité verticale est simplement la question. Pour chaque sujet, on trouve les exercices en-dessous de l'axe horizontal et le problème au-dessus. Les principales conclusions se voient sur cet histogramme comparatif :

- 1° Augmentation régulière du niveau de complexité au fil des années,
- 2° Difficulté relative des exercices par rapport au problème,

BACC. C Strasbourg — Niveaux NLSMA



Histogrammes par questions

3° Les exercices font double emploi avec le problème, hormis le fait qu'ils complètent la couverture du programme (sauf en 1973, où les exercices sollicitent des niveaux non pris en compte par le problème).

Par ailleurs, la couverture des programmes par les sujets de bac. évolue différemment dans les séries C et D :

— pour les sujets de série C étudiés, la couverture des programmes est en augmentation au fil des années ;

— pour les sujets de série D, il y a au contraire une bonne stabilité. On pourrait presque parler d'une standardisation des sujets.

D. Reisz apporte les précisions suivantes, issues des échanges A.P.M. :

Le courrier A.P.M. met en évidence une inquiétude manifeste des collègues.

L'augmentation au fil des années, de la complexité cognitive est un phénomène général. Il y a des cas beaucoup plus frappants que celui de l'académie de Strasbourg. Ainsi, un sujet de Poitiers reproduit la première question du concours d'entrée à l'Ecole Normale Supérieure d'il y a deux ans. Voir aussi la longueur du dernier sujet de série C à Paris.

En revanche, la couverture du programme est très diversifiée, et l'augmentation observée à Strasbourg pour la série C n'est pas générale.

La discussion suscitée par cette question dans le groupe fait essentiellement apparaître des préoccupations quant à l'ABSENCE DE CRITERES clairement précisées :

- pour l'élaboration des sujets,
- pour la correction des copies.

N'est-il pas inquiétant, par exemple, que sur 30 propositions de sujets, 20 soient immédiatement rejetés comme non conformes (exemple communiqué par D. Reisz, sur l'élaboration des sujets de Dijon). Or ces sujets sont proposés par des collègues, enseignant en Terminale.

La connaissance des raisons de ces rejets serait utile. Malheureusement, la proposition de rendre publics tous les sujets présentés, au bout de deux ans, a été rejetée. Il apparaît d'ailleurs que les raisons de ce refus de l'Inspection générale ne sont pas toutes très claires. Ainsi l'édition scolaire n'y a-t-elle pas un rôle ?

Une critique faite à un sujet : "Il manque d'originalité" apporte peut-être quelque lumière, mais soulève des interrogations. Que veut-on tester chez le candidat bachelier ? Le bac. ne serait-il pas un examen, mais plutôt un *rite initiatique* (et alors l'absence de critères précisés, le secret des commissions, bref "l'ambiance de loge maçonnique", se comprend très bien) ?

Historiquement, cette notion de rite a pu être justifiée dans le passé : parmi la population limitée des lycéens, il s'agissait de déterminer ceux que l'on jugeait dignes d'accéder au premier grade universitaire. Leur culture générale était en principe achevée, ils devenaient de jeunes adultes prêts à se spécialiser dans une branche de leur choix. Or la situation actuelle diffère trop de cette situation passée pour qu'une réflexion sur le bac. soit constamment différée dans le monde enseignant. Et, au travers de l'analyse de sujets proposée par l'A.P.M., c'est finalement un point de départ pour une telle réflexion qui est mis en évidence.

Signalons pour terminer, quelques détails techniques apparus lors de la discussion, et qui pourraient être utiles à des groupes s'attelant à une telle tâche d'analyse :

— La difficulté n'est pas entièrement déterminée par la complexité cognitive.

Exemple : dans un sujet de Strasbourg, intervenait une intégrale

$\int_1^x \frac{1}{t} dt$. Très peu de candidats ont pensé à envisager le cas $x < 1$. Malgré cela, la complexité cognitive de cette question n'est pas considérable. Demander de dériver $F(x) = \int_1^{1+x^2} f(t)dt$ est une complexité cognitive plus grande, même avec l'indication d'utiliser la dérivation des fonctions composées. Et pourtant, la réussite sera sans doute meilleure auprès des élèves.

— Dans certains cas, l'attribution d'un niveau NLSMA à une question est sujette à caution : il peut y avoir plusieurs procédures de résolution ; l'entraînement des élèves peut différer suivant les classes. C'est pourquoi l'article du groupe I.R.E.M. de Strasbourg indiquera ses attributions, sujet par sujet, ce qui en permettra la critique.

— Un histogramme tenant compte du barème de correction différerait peut-être sensiblement de l'histogramme présenté (où l'unité est la question). Une étude dans ce sens sera faite sur deux

des sujets étudiés, pour lesquels il est possible de disposer du barème (barème du Bas-Rhin, celui du Haut-Rhin n'étant pas le même (!)).

Prolongements possibles

Des travaux analogues d'analyse peuvent facilement être entrepris par des petits groupes d'enseignants. Il est nécessaire de procéder à une ou deux séances préliminaires pour mettre au point l'attribution de niveaux aux questions. L'expérience montre que la procédure indiquée ci-dessus permet de se mettre rapidement d'accord et d'obtenir une fiabilité acceptable.

Un autre type de travail consiste à explorer divers critères pour aboutir à la confection de sujets en accord avec ces critères. Donnons un exemple volontairement naïf, mais parlant : on peut tenter, à partir de la même idée centrale, de bâtir d'une part un sujet pour lequel les échecs soient significatifs, et au contraire un sujet pour lequel les réussites soient significatives. (Note : l'exemple est naïf dans la mesure où il ne réfère pas encore à des critères bien fixés. Mais c'est peut-être en essayant d'affirmer concrètement des idées, au départ trop grossières, que l'on risque d'être à même de préciser des critères opérationnels).

SIGNIFICATION DES COMPORTEMENTS DE REPONSE

Rapport de R. DUVAL

Un document, "Démarches de réponse en mathématique" a été remis aux participants et succinctement présenté. Il s'agit des résultats d'une enquête faite auprès des élèves en fin de cinquième et destinée à éclaircir certains problèmes soulevés dans des enquêtes préalables faites dans le cadre des I.R.E.M. de Strasbourg et de Rouen. Dans la présentation, ce sont les aspects méthodologiques qui ont été soulignés et non les résultats. Deux exemples ont été pris pour indiquer comment à partir des réponses fournies à une séquence d'items il est possible de déterminer les différentes stratégies de réponses adoptées et d'avoir ainsi une information soit sur la façon dont les élèves utilisent une notion soit sur la façon dont ils ont perçu la question posée. Un troisième exemple a été regardé pour montrer l'intérêt de plusieurs modalités dans la présentation d'un questionnaire : elles permettent d'avoir les différentes réactions pour des variations minimales d'une question, alors qu'il n'est pas possible de les présenter toutes aux mêmes individus.

Mais les questions des participants de ce groupe ont porté sur l'utilisation de l'analyse des correspondances pour l'interprétation des données. Elles ont donné lieu à d'assez longues explications. Nous n'en retiendrons ici que l'idée principale.

Tout d'abord : quelle est l'utilité d'un codage des données en disjonctif total ? Le codage en disjonctif total est un codage en 0,1 des différentes données d'une réponse avec la condition que la note 1 apparaisse une fois et une seule dans la suite des notes ayant trait aux différentes données retenues pour une réponse. L'intérêt d'un tel codage tient au fait que l'analyse du tableau $I \times R$ pour en rechercher les tendances donne, en disjonctif total le même résultat que l'analyse du tableau $R \times R$ (I étant l'ensemble des indices attachés aux individus et R l'ensemble des indices affectés aux types de réponses).

La deuxième question soulevée a été : "comment interpréter les axes qui sortent ?" En général, l'intérêt de l'analyse des correspondances est de partir de données amorphes, non structurées, pour essayer de voir s'il n'y a pas des lignes de force qui organisent ce magma apparent. Si des axes se dégagent le problème est donc de déterminer à quoi ils correspondent. Or, ici, l'utilisation de l'analyse des correspondances est un peu différente. Avec les différentes modalités de questionnaire (pour l'utilisation de l'analyse des correspondances il apparaît plus commode de recourir à trois modalités) on a une organisation de principe par rapport à laquelle il peut y avoir des déviations. Le travail d'interprétation concerne essentiellement le sens et la valeur de ces déviations qui sont observées.

Inventaire des données et des modalités
des critères

Confection et utilisation de grilles d'observation en classe

Présentation d'une étude
sur la communication effectuée par
un groupe de l'I.R.E.M. de Strasbourg

Plan de l'exposé

Evolution du groupe de 1972-76
Conditions d'observation

Le groupe a essayé de faire varier un maximum de paramètres pour essayer de déterminer dans quelle mesure il pouvait exister des constantes indépendantes. Aussi des observations ont été faites :

- dans différents établissements
- dans différentes classes
- dans différentes matières (math., français, sciences...)
- par différents observateurs.

Codage utilisé

Un observateur "enregistre" la conversation en utilisant les signes suivants :

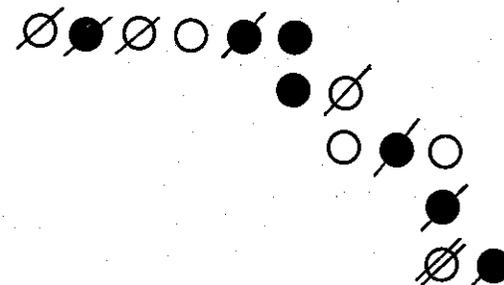
	Intervention du professeur	Intervention de l'élève
Question	○	●
Affirmation	⊘	●
Ordre	⊘	●
Jugement de valeur	⊘	●

Il importe de noter que chaque signe correspond à *une phrase*.

L'enchaînement se fait soit horizontalement, soit verticalement suivant les critères :

- enchaînement horizontal si la proposition (phrase) clarifie ou résume ce qui précède ;
- enchaînement vertical si la proposition contient un élément nouveau ou déterminant pour la progression.

Exemple d'un relevé :



Ces relevés ont été appelés "chapelets".

Les chapelets sont établis sur toute la durée du cours.

Grilles

Lorsque le chapelet est établi, on comptabilise le nombre d'interventions de chaque nature par tranche de 5 minutes. La transcription graphique des résultats (en abscisse le temps, en ordonnée le nombre de phrases) donne un aperçu de l'évolution des échanges entre professeur et élèves pendant le cours.

Exemple d'analyse des résultats

L'exemple traité est trop long pour être rapporté ici. Cependant il est intéressant de noter que :

- 70 % des phrases dites durant le cours le sont par le professeur ;
- sur l'ensemble de toutes les phrases, les questions des élèves dépassent rarement 2 % et se regroupent en début et fin d'heure ;
- les affirmations des élèves correspondent quantitativement aux questions des professeurs. (Les élèves dialoguent très rarement entre eux).

Perspectives

- Poursuivre les observations.
- Etudier en détail les chapelets pour découvrir des "séquences types".
- Et bien d'autres encore ...

Références

Dialogue de l'éducation : Leroy
Grilles de Bayer et Landsheere

Interventions et questions

Elles concernent essentiellement :

- La méthode de travail.
- Le codage immédiat et l'étalonnage des observateurs.

- La sensibilité et la fiabilité de l'outil.
- Le comportement de l'enseignant.
- Les objectifs du groupe.
- L'évolution interne du groupe : composé d'une quarantaine de stagiaires la première année (72), le nombre s'est sensiblement réduit pour rester stationnaire autour de 5-6 depuis 74. Depuis la rentrée 75-76 le groupe travaille d'une manière autonome (l'animateur ne participe pas aux séances de travail).

Participants au Colloque évaluation

STRASBOURG — du 12 au 14 mars 1976

ALLAIN Yvon, I.R.E.M. de Brest
 AMIGUE René, Université de Provence Centre d'Aix
 AUDIGIER M. Noelle, I.N.R.D.P.
 BACHMANN Etienne, I.R.E.M. de Strasbourg
 BARBANCON Gérard, I.R.E.M. de Strasbourg
 BARDY Michel, I.R.E.M. de Nancy
 BARTHEL M. Thérèse, I.R.E.M. de Nancy
 BOISNARD Danièle, I.R.E.M. de Rennes
 BONNIOL J. Jacques, Université de Provence Centre d'Aix
 BOPP Nicole, I.R.E.M. de Strasbourg
 BOROWCZYK Jacques, I.R.E.M. de Poitiers
 BUISSON Pierre, I.R.E.M. de Rouen
 CARDINET Jean, I.N.R.D.P. Neuchâtel
 CARREAU Gilbert, I.R.E.M. d'Orléans
 CAUZINILLE Evelyne, C.N.R.S. Rue de la Tourelle, Paris
 CHARNAY Roland, E.N. de Bourg-en-Bresse
 CHELLY Clémence, I.N.R.D.P. Paris
 CHENEDE Monique, I.R.E.M. de Rennes
 COLOMB Jacques, I.N.R.D.P. Paris
 CREPIN Roger, I.R.E.M. de Limoges
 DAUVISIS M. Claire, I.R.E.M. de Toulouse
 DE COINTET Michel, I.R.E.M. de Strasbourg
 DE KETELE J. Marie, Université de Louvain, Belgique
 DENIS Alain, I.R.E.M. de Lyon
 DESRUES Claudine, I.R.E.M. de Brest
 DIMIER René, I.R.E.M. de Lyon
 DUMONT Bernard, E.S.A.M. Bourges
 DUPUIS Claire, I.R.E.M. de Strasbourg
 DUVAL Raymond, I.R.E.M. de Strasbourg
 GABORIEAU J. Pierre, I.R.E.M. de Rennes
 GILLET Bernard, I.R.E.M. de Strasbourg
 GLAESER Georges, I.R.E.M. de Strasbourg
 GRAS Régis, I.R.E.M. de Rennes

GREMILLARD Jean, I.R.E.M. de Besançon
 GUINARD Colette, Université de Provence Centre d'Aix
 HUTIN Raymond, Recherche Pédagogique, Genève
 KAHN Claudine, I.R.E.M. de Strasbourg
 KITTEL Françoise, I.R.E.M. de Strasbourg
 KOZLIK Frédéric, I.R.E.M. de Strasbourg
 LAMBINET Paulette, I.R.E.M. de Strasbourg
 LASSAVE Claude, I.R.E.M. de Toulouse
 LAURENT André, I.R.E.M. de Marseille
 LEGRAND Louis, I.N.R.D.P. Paris
 LE NEZET Georges, I.R.E.M. de Rennes
 MALCUIT M. Jeanne, C.E.S. Berlioz Colmar
 METENIER Jacqueline, I.R.E.M. de Rouen
 MEYER Etienne, I.R.E.M. de Strasbourg
 MEYER J. Marie, I.R.E.M. de Strasbourg
 MICHAU Geneviève, I.R.E.M. de Besançon
 MILHAUD Nadine, I.R.E.M. de Toulouse
 MISERY Michel, I.R.E.M. de Picardie
 MONFRONT (Mme), 22 rue Marceau, 92170 Vanves
 MOUCHE Paul, I.R.E.M. de Grenoble
 NOEL Guy, Université d'Etat, Mons, Belgique
 NOEL (Mme), Université d'Etat, Mons
 PELE Colette, OFRATÉME
 PELNARD-CONSIDERE Jacqueline, I.N.O.P. Paris
 PERRIER Yvonne, I.R.E.M. de Rennes
 PEYRIERAS Jeanne, I.R.E.M. de Limoges
 PLUVINAGE François, I.R.E.M. de Strasbourg
 QUEVAREC Jean, I.R.E.M. de Rennes
 RAULIN Laurence, I.R.E.M. de Marseille
 REISZ Daniel, I.R.E.M. de Dijon
 RICHARD J. François, Université Paris 8
 ROBIN Daniel, I.N.R.D.P. Paris
 ROLAND J. Pierre, I.R.E.M. d'Orléans
 ROUSSEL Yves, I.R.E.M. de Picardie
 SAVALLE Albert, I.R.E.M. de Rouen
 SCHAEFFER Michel, I.R.E.M. de Strasbourg
 SURLAIVE Claude, I.N.R.D.P.
 TOURNEUR Yvan, Université d'Etat, Mons, Belgique
 VERDIER Jacques, I.R.E.M. de Nancy

ISBN - 0338 - 7135

Bulletin Inter-I.R.E.M. - Villeurbanne

