

L'évaluation EVAPM

Aspects méthodologiques de l'évaluation

Antoine BODIN, François COUTURIER et Régis GRAS

Cette analyse est centrée sur l'exploitation des résultats de l'Observatoire EVAPM.

Au cours des sept dernières années les enquêtes EVAPM ont produit des données importantes, en nombre comme en qualité, sur les acquisitions des élèves dans le domaine mathématique, et cela de la classe de Sixième à la classe de Première incluse.

Les brochures EVAPM réalisées en grande partie grâce au travail militant des membres de l'équipe EVAPM, présentent des résultats commentés et des analyses qualitatives que la plupart des observateurs s'accordent à trouver intéressants et susceptibles d'aider les collègues dans leur réflexion comme dans leur action.

L'un des principes d'EVAPM étant de mettre les résultats des évaluations à la disposition des enseignants aussi rapidement que possible, ces analyses n'utilisent qu'une partie de l'information disponible. De plus ces analyses, si elles répondent à certaines questions en posent de nouvelles.

ticulier l'analyse implicative développée par Régis GRAS et ses étudiants, ainsi que la théorie des Réponse aux Items (IRT) encore peu connue dans notre pays.

Ces méthodes seront rapidement présentées sur des exemples, mais le but de l'atelier est essentiellement d'amener les participants à se poser de nouvelles questions et à envisager différentes façons d'essayer d'y répondre.

Des fichiers de résultats statistiques seront disponibles, prêts à être traités en fonction des questions qui se dégageront. Le traitement pourra être fait sur place par les animateurs mais les participants qui voudraient aller plus loin pourront demander une copie des fichiers informatiques. L'atelier s'appuiera sur les nombreux documents EVAPM. Des texte permettant de rentrer dans la problématique de recherche (si ce n'est déjà fait) seront remis aux participants.

De nombreux exemples pourront être donnés, mais les animateurs de l'atelier comptent largement sur les participants pour faire émerger des questions :

Que souhaite-t-on savoir de plus ?

Que peut-on espérer mieux comprendre, relativement aux comportements observés chez des élèves en situation d'évaluation ?

Exemple de question analysée

Inégalité triangulaire

Si l'on avait une feuille assez grande, pourrait-on construire :

1°) Un triangle dont les côtés mesurent : 30 cm ; 18 cm ; 45 cm ?

☐ OUI ☐ NON

Coche la bonne réponse

Explique ta réponse

Oui et explication correcte
EVAPM 4/91 : 36 %

Non réponse
15 %

2°) Un triangle dont les côtés mesurent : 28 cm ; 14 cm ; 44 cm ?

☐ OUI ☐ NON

Coche la bonne réponse

Explique ta réponse

Oui et explication correcte
EVAPM 4/91 : 37 %

Non réponse
17 %

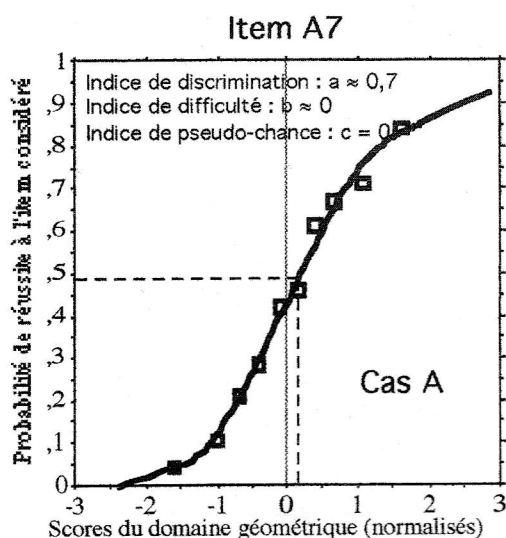
Réussite conjointe
EVAPM 4/91 : 30 %

EVAPM 4/91 Question A16-17

Le type de données recueillies justifie et autorise l'utilisation de méthodes d'analyse statistiques complémentaires et en par-

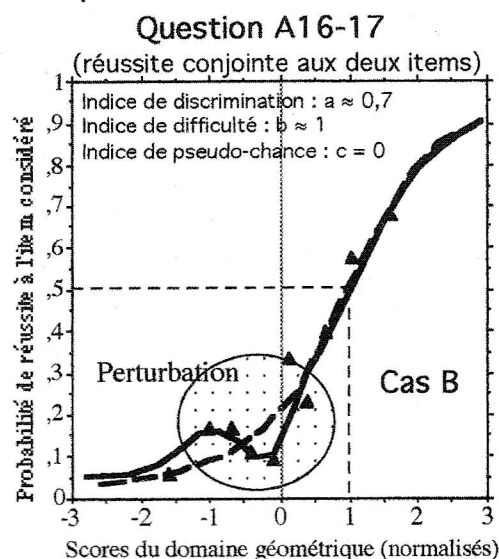
L'utilisation des données de masse telles que celles fournies par les enquêtes EVAPM et en utilisant des méthodes modernes d'analyse de données ?

Exemples de courbes de réponse



Cas A : ajustement ogival légitime

Cas B : ajustement ogival compromis et contredit par l'analyse didactique qui permet d'expliquer la perturbation observée



Données EVAPM 4/91

Les courbes de réponse du type de celles produites ci-dessus sont très utilisées dans le cadre de la théorie des réponses aux items, plus connue sous le nom d'IRT «Item response theory» (étude des qualités psychométriques et informatives des questions, constructions d'échelles.... A un niveau plus modeste (et moins controversé), elles sont très utiles pour étudier la validité didactique des questions par rapport à des domaines ou à des champs précisés (exemples lors de l'atelier et explication des divers indices utilisés).

Documents distribués lors de l'atelier et que vous pouvez demander à votre régionale de l'APMEP

- 1 - Une présentation générale de l'Observatoire EVAPM
- 2 - Une note méthodologique relative aux enquêtes EVAPM
- 3 - Un document «Mise en relief» présentant quelques résultats qui pourront être utilisés pendant l'atelier
- 4 - Les épreuves «A» de Quatrième 1991, «WA et WB» de Troisième 1992
- 5 - Un texte présentant une analyse implicite de l'épreuve WA de Troisième 1992 (méthode Régis GRAS)

14 Evaluation du fonctionnement du système d'enseignement des mathématiques en France

Antoine BODIN et Jean-Pierre SICRE, EVAPM

L'Observatoire EVAPM est un Groupe de Travail de l'Apmp, Equipe de recherche associée à l'INRP.

En premier lieu, nous chercherons à montrer pourquoi et de quelle façon il est possible de considérer que, au moins en France, l'ensemble des actions ayant pour finalité la transmission et le développement des connaissances et des savoirs du domaine mathématique peut être structuré en système.

Nous mettons ensuite en évidence l'existence et le positionnement d'un système partiellement isomorphe au premier : le sys-

tème de régulation de ce système de formation, avec en particulier ses actions de contrôle-évaluation (sans que la distinction soit toujours très claire).

En particulier, nous étudierons les relations existant entre les divers agents, institutions et catalyseurs de ce système de régulation :

Éléments strictement institutionnels :

Direction de l'Évaluation et de la Prospective, Inspection Générale et Régionale,

Direction des Lycées et Collèges,...

Éléments professionnels :

L'Association des Professeurs de Mathématiques et l'observatoire EVAPM, les IREM, la Société Mathématique de France, etc...

Éléments liés à la Recherche :

INRP, Recherche «fondamentale» en Didactique des Mathématiques et Groupe de Recherche du CNRS,

et aussi : Études internationales, média, organisations syndicales, parentales, etc.....

En fait il ne s'agira, dans ce qui précède, que de broser le décor. L'essentiel devrait consister à se demander quels sont les faits, pouvant plus ou moins relever de l'évaluation, qui ont des effets sur le fonctionnement du système, et quels sont ces effets.

Chaque fois que possible, nous nous appuierons sur des exemples concrets et précis, c'est à dire à contenus mathématiques explicites.

La question qui nous intéresse peut encore s'énoncer d'une autre façon : comment se construisent les décisions qui gouvernent l'enseignement des mathématiques dans notre pays ? (Il serait en effet facile de montrer que, dans ce domaine, les choses varient considérablement d'un système éducatif à un autre).

La recherche n'est pas assez avancée pour que l'on puisse répondre avec suffisamment d'assurance à la plupart des questions qui peuvent se poser, et l'analyse systémique envisagée demande à être affinée et supporterait des mises en questions et des concours multiples.

Une première collaboration, amorcée avec des sociologues de l'éducation, pourrait contribuer à cet affinement. La prise en compte des connaissances et des représentations des enseignants est tout aussi essentielle. D'autre part, il est clair que notre sujet est directement lié à celui de la transposition didactique, même s'il le déborde sans avoir la prétention de l'épuiser ; nous ne développerons pas ce point ici.

Pour simplifier, nous serons amenés à distinguer deux types d'évaluation participant à la régulation du système :

- Évaluations du premier type : jugements d'évaluation fondés sur l'expérience ou l'opinion sans qu'une méthodologie puisse être mise en évidence et communiquée à autrui.
- Évaluation du deuxième type : actions

d'évaluation explicites et assorties d'une méthodologie communicable et susceptible d'être mise en examen (critique de la méthode).

Bien sûr, la plupart des jugements d'évaluation participent des deux types. Les définitions ci-dessus montrent que les deux types d'évaluations se caractérisent par des démarches différentes sans qu'il soit possible de postuler que les évaluations du premier type seraient de moins bonne qualité que celles du second type (dans notre cas, qu'elles conduiraient systématiquement à prendre de plus mauvaises décisions).

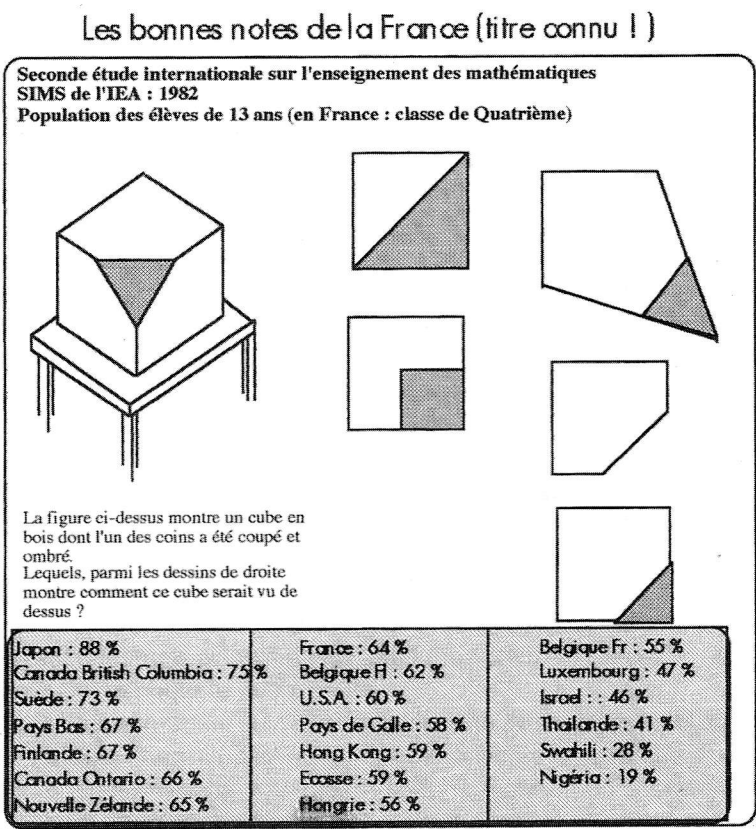
Notre hypothèse est que, jusqu'aux années 80, les évaluations du premier type dominaient largement dans notre système. Les idées et théories relatives aux évaluations plus instrumentées ne se sont développées qu'à partir des années 60 et ne sont que récemment parvenues à maturité (du moins dans notre pays)..

L'apport des études et évaluations internationales ?

Dans le système éducatif français, la Direction de l'Évaluation et de la Prospective (la DEP) occupe une place dominante en ce qui concerne l'évaluation institutionnelle, et même l'évaluation tout court (comme en témoigne sa forte implication dans des évaluations qui se présentent comme purement pédagogiques).

Le mot d'ordre du Directeur actuel de la DEP est bien connu : «*développer la culture de l'évaluation dans le système éducatif*» et il est certain que la DEP a fourni un travail très important depuis une dizaine d'années, travail qui a accru la qualité des évaluations réalisées, des indicateurs fournis et leur impact sur le système.

D'autres institutions ou organismes tels que l'INRP, les IREM, les MAFPEN, et...l'APMEP ont aussi contribué à développer l'évaluation dans le système éducatif. En particulier, l'Observatoire EVAPM, développé depuis 1986 dans le cadre de l'APMEP avec des concours ou appuis divers est généralement pris au sérieux aussi bien par les enseignants de mathématiques que par les chercheurs ou par les «*décideurs*».



L'intérêt d'un tel palmarès est très limité !
Ce n'est qu'une occasion pour signaler que sur les 175 questions de l'enquête, deux seulement concernent des objets de l'espace. Encore s'agit-il de questions de nature quasiment perceptive.
En fait, la situation des divers pays par rapport à l'enseignement de la géométrie de l'espace est très variable et il est quasiment impossible de trouver une question qui obtiennent l'agrément des pays engagés dans une telle étude.

Quelles évaluations ont des effets ?

Quelles évaluations ont des effets positifs sur le fonctionnement du système ?
Guy BROUSSEAU a depuis longtemps, avec d'autres, dénoncé les effets négatifs de certaines évaluations (1979 à CAMPINAS - texte cité). Nous voyons d'ailleurs facilement les effets réducteurs d'évaluations effectuées dans tel ou tel pays (mais peut-être aussi chez nous): «écrasement de l'apprentissage sur des objectifs insignifiants»... dressage et apprentissages uniquement procéduriers,... perte du sens des connaissances...

Toutefois il peut aussi y avoir des effets positifs (encore que définir cette expression ne serait pas facile, et il vaudrait mieux dire «qui semblent positifs»).

Ainsi, après les évaluations du SPRESE et de l'IEA en 1984, il était clair que les enfants de notre pays avaient perdu, en géométrie, des compétences que beaucoup considéraient comme importantes. La modification des programmes de collège qui s'ensuivit en 1985 a modifié la situation de façon importante et des effets massifs

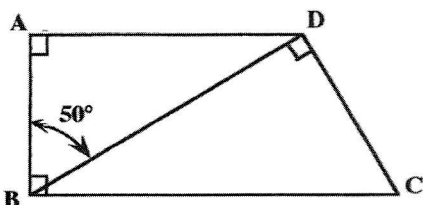
sont apparus en quelques points. La question présentée ci-dessus n'est qu'un exemple de questions qui semblent montrer que le niveau de compétence des élèves peut doubler en quelques années dans certains sous-domaines.

Les choses ne se passent pas toujours aussi bien!
Par exemple, l'observation de compétences jugées insuffisantes dans le domaine de la proportionnalité et les mesures prises (en définissant des capacités exigibles) ne semblent pas avoir suffi pour que suffisamment d'élèves parviennent à maîtriser la notion de pourcentage.

N.B.: Les réflexions et modélisations proposées doivent beaucoup à un travail collectif mené en 1990 et 1991 dans le cadre du Groupement de Recherche «Didactique et acquisitions des connaissances scientifiques» du CNRS (Sous la responsabilité de Gérard VERGNAUD, Guy BROUSSEAU et Antoine BODIN).

ABCD est un trapèze rectangle en A et B, l'angle $\widehat{BDC}$ est droit et $\widehat{ABD} = 50^\circ$

Indique la mesure des angles suivants, sans donner le détail des calculs:



Réussite conjointe :
EVAPM3/90 : 66 %

$\widehat{DBC}$

EVAPM3/90 : 82 %
DEP3/90 : 82,5 %
EVAPM6/87 : 27 %
SPRESE 3/84 : 48 %

$\widehat{ADB}$

EVAPM3/90 : 83 %
DEP3/90 : 82,7 %
EVAPM6/87 : 23 %
SPRESE 3/84 : 44 %

$\widehat{ADC}$

EVAPM3/90 : 73 %
DEP3/90 : 74,7 %
EVAPM6/87 : 17 %
SPRESE 3/84 : 36 %

Non réponse :
EVAPM3/90 : 05 %

Question d'origine SPRESE, utilisée dans l'évaluation fin de troisième 1990 organisée par la Direction de l'Évaluation et de la Prospective (D.E.P.), et utilisée simultanément dans EVAPM 3/90 avec l'autorisation de Monsieur le Directeur de la DEP.

Après une augmentation de 40%, un objet vaut 84 F.

Combien valait-il avant cette augmentation ?

Expliquez ce que vous faites pour trouver la réponse

Résultat exact

Quatrième 1989 : 05 %

Troisième 1990 : 22 %

Elèves orientés en Seconde LEG : 24 %

DEUG lettres et S.H. Paris 7 1985 : 30 %

Première 1993 :

Série S : 80 %

Elèves orientés en TC : 71 %

Série B : 59 %

Série A1 : 49 %

Série F : 53 %

Question non traitée
Première 1993 :

Série S : 04 %

Série B : 10 %

Série A1 : 11 %

Série F : 11 %

Introduction directe du
coefficient 1,4

Première 1993 :

Série S : 19 %

Elèves orientés en TC : 21 %

Série B : 17 %

Série A1 : 21 %

Série F : 19 %

Question Université PARIS 7

Pour aller plus loin

- **BODIN A. (1992)** - Les mathématiques en fin de Troisième générale - évolution des compétences observées chez les élèves au cours des années 80. In Rapport à Monsieur le Ministre de l'Éducation Nationale établi par Monsieur Claude THELOT, Directeur de la DEP (1992). Publié dans le dossier Éducation et formations n°17 d'octobre 1992 (DEP) et dans le bulletin de l'APMEP Février 1993.
- **BROUSSEAU G. (1979)** : Evaluation et théories de l'apprentissage en situations scolaires. Conférence faite à la rencontre CIEAEM de CAMPINAS (Polycopié).
- **Educational Testing Service (1991)** : Learning Mathematics - Princeton - USA.
Il s'agit du rapport international de l'étude largement rapportée en France par la presse en 1992 et qui classait la France 6ème sur 20 (ce qui n'a bien sûr aucune signification).
- **ROBITAILLE D.F. and all. (1993)** - Curriculum Frameworks for Mathematics and Science. TIMSS Monograph n° 1. Traduction française de la seconde partie (BODIN A) : TIMSS - une étude de l'IEA Présentation et traduction française de la grille de classification (framework for mathematics) utilisée dans le cadre de la Troisième Étude Internationale sur l'Enseignement des Mathématiques et des Sciences de l'IEA (disponible à l'IREM de BESANÇON).
- **ROBIN, D.; BARRIER, E. (1985)** : Enquête internationale sur l'enseignement des mathématiques Le cas Français. - Institut National de Recherches Pédagogiques (Paris)
- **THÉLOT C. (1994)** - L'évaluation du système éducatif français - Revue Française de pédagogie N° 107 - INRP. PARIS.

Les brochures EVAPM de l'APMEP

Ces brochures contiennent les épreuves, les résultats et les analyses des évaluations des programmes de mathématiques menées par l'APMEP (BODIN. A., et All)

- EVAPM6/87 Évaluation fin de Sixième 1987 (Paris 1987)
- EVAPM5/88 Évaluation fin de Cinquième 1988 (Paris 1988)
- EVAPM4/89 Évaluation fin de Quatrième 1989 (Paris 1989)
- EVAPM3/90 Évaluation fin de Troisième 1990 (Paris 1990)
- EVAPM6/89-5/90 Compléments 1991 des évaluations Sixième et Cinquième
- EVAPM2/91 Évaluation du programme de Seconde 1991 (Paris 1991)
- EVAPM1/93 Evaluation fin de Première 1993

Le fascicule 1 d'EVAPM1/93 (168 pages) est constitué des documents utilisés pour l'évaluation (dossier professeur).

Le fascicule 2 (192 pages), publié en 1994, contient les résultats et les analyses.

