PROBLEMES DE L'EVALUATION DES SAVOIRS MATHEMATIQUES

I.R.E.M. de Besançon

Antoine BODIN
Collège d'Ornans

I – L'EVALUATION A L'I.R.E.M. DE BESANCON : 1979-1983.[1]

Au cours de la période 1979-1983, le groupe évaluation de l'I.R.E.M. de Besançon, réunissant régulièrement des collègues de tous les niveaux d'enseignement (sauf élémentaire), et s'appuyant sur un réseau de professeurs correspondants, a produit un référentiel[2] d'évaluation couvrant l'ensemble du premier cycle secondaire.

Le but de ce travail était d'aider les collègues de collège dans l'amélioration de leurs pratiques évaluatives, de favoriser la reconnaissance des objectifs poursuivis de façon plus ou moins implicite, ainsi que l'harmonisation des critères et la communicabilité des résultats.

Au cours de ce travail, nous avons eu le souci de prendre en compte les habitudes et les attentes des enseignants de façon à construire un outil d'évaluation qui soit recevable. Nous avons bien sûr cherché à élargir le champ de l'évaluation et à améliorer les modes de questionnement, mais en restant dans des limites acceptables.

La démarche suivie a été tout à fait empirique : reconnaissance d'une classe d'objectifs, reconnaissance et formulation des objectifs, traduction en termes opérationnels, construction des épreuves. L'essentiel de notre travail a porté sur des savoirs qualifiés de minimums, mais, peu à peu nous nous sommes penchés en amont sur les prérequis et en aval sur les objectifs d'approfondissement.

Le référentiel obtenu se compose d'environ 300 micro-objectifs opérationnalisés et regroupés en classes. Chaque classe correspondant à un domaine conceptuel (la proportionnalité...) ou à un domaine de formation (les puissances...). Pour chaque

(1) Ce texte a été présenté au Séminaire National de Didactique des Mathématiques à Paris le 26 mai 84.

(2) Ensemble des objectifs décrits en termes de capacités et de savoir-faire. Il décrit des actions observables : identifier, prouver, tracer... Il précise en outre dans quel type d'activités et dans quelles conditions de réalisation les savoirs et savoir-faire seront acquis. Voir l'article d'**Odile Backscheider** «le contrôle continu en L.E.P.». Bulletin APMEP n°345.

«petit x» n° 7 pp. 5 à 28. 1985

classe nous avons élaboré un test de validation. Chacun des 30 tests a été passé par un nombre important d'élèves (plusieurs milliers pour la plupart) et de nombreux résultats ont été retournés à l'I.R.E.M., ce qui a permis de construire des étalonnages et de connaître les taux de réussite question par question.

Les conditions de passation, peu standardisées, font que les résultats obtenus ne peuvent rendre compte que de ce que les élèves savent faire lorsque les enseignants estiment qu'ils sont prêts à le faire. La volonté d'efficacité immédiate et d'exhaustivité ont amené à reléguer au second plan l'aspect recherche. Après coup il nous vient bien des regrets : certes l'outil (Bodin et al. 1982) fonctionne et est largement utilisé (il fonctionne peut-être même trop bien et pourrait de ce fait éviter qu'un certain nombre de questions ne soient posées), mais, faute de plan d'expériences et d'hypothèses préalables, il ne fournit pas toujours les renseignements que l'on pourrait en attendre. Malgré ces réserves, au jour le jour, nous avons découvert ou sans doute redécouvert une multitude de petits faits difficiles à coordonner et qui ont sensiblement modifié notre conception de l'évaluation et la construction des épreuves.

Voici d'abord quelques observations.

1) Les attentes des enseignants dépendent du degré de réussite des élèves.

Plus précisément, lorsqu'une épreuve est bien réussie par les élèves, elle est jugée trop facile et est, à terme, remplacée par une autre plus difficile. Il n'y a pas symétrie, lorsqu'une épreuve est mal réussie, ce sont les élèves qui sont faibles et rarement l'épreuve qui est jugée trop difficile.

Exemple 1.

Au début de notre travail nous avions vraiment les savoirs minimums à l'esprit. Le test 6A (annexe 1) sur les opérations que nous avons fait à ce moment était bien réussi (réussite : R = 63%). Ce test a été jugé trop facile et l'année suivante nous l'avons remplacé par un test qui au sens de nos critères (la réussite signifie que les 2/3 des questions sont réussies) n'est plus réussi que par 19% des élèves. Le score moyen est de 9,6 sur 20, la distribution des résultats est gaussienne. Ce test donne satisfaction aux enseignants.

Exemple 2.

Le test 6D (annexe 3) sur le parenthésage et les priorités a été réussi par 61% des élèves. Jugé trop facile, il a pratiquement été retiré de la circulation.

Exemple 3.

Après analyse des prérequis nécessaires à l'abord d'une question nouvelle, un groupe d'enseignants construit un test de prérequis. Miracle ! l'épreuve est bien réussie. **La conclusion aurait pu être** que les élèves étaient prêts pour l'étape suivante.

En fait il a simplement été décrété que l'épreuve était trop facile et qu'il fallait la modifier.

On peut donc estimer qu'un bon test pour les enseignants est un test qui met en échec la moitié des élèves.

Selon nos observation, et sous réserve de vérification, on peut estimer que pour une question donnée, on obtiendrait :

$R \geqslant 75\%$	: question jugée trop facile
$75\% > R \geqslant 50\%$	: question facile
$50\% > R \geqslant 25\%$	: question correcte
$25\% > R \geqslant 10\%$	: question difficile
$10\% > R$	: question trop difficile.

2) Les enseignants évaluent pour mettre des notes et non pour connaître les représentations ou les processus utilisés par les élèves.

Nous avons adopté chaque fois que nécessaire des questions associées (Pluvinage 1978).

Exemple.

Entoure les multiples de 4	121	136	100	64	74

Cet item traduit l'objectif : reconnaître les multiples de 4. Il convient certainement de coter 0 s'il y a une erreur. En effet certains élèves croient qu'il convient de considérer la somme des chiffres, d'autres que tout nombre dont le chiffre des unités est 4 est divisible par 4. Les enseignants éprouvent une certaine répugnance à noter comme il est indiqué et préfèrent souvent compter 1/5 de point par réponse juste, au risque de conforter l'élève dans son erreur.

Des soucis de justice amènent souvent à attribuer la moitié des points à des réponses qui traduisent manifestement une représentation erronée favorisent ainsi la perpétuation de l'erreur.

3) La réussite d'un élève à une question dépend :

- de la formulation de la question,
- de la place de la question dans l'épreuve,
- de l'ensemble de l'épreuve,
- de l'apprentissage,
- de la proximité de l'apprentissage,
- de la proximité d'autres apprentissages,
- ...

peut-être aussi :

- de la fonction assignée à l'évaluation (telle qu'elle est reconnue par l'élève),
- de l'heure de l'évaluation,
- et si l'on en croit certains, du menu du dernier repas («des frites à la cantine» signifie un désastre pour l'évaluation...).

Voici quelques exemples.

a) Les deux questions suivantes traduisent sensiblement le même objectif. Bien sûr, une analyse de la tâche permet de distinguer ces questions, mais dans les deux cas il s'agit de savoir utiliser le compas pour placer un point à une distance donnée de deux autres. Comme on le voit le taux de réussite en classe de sixième passe de 89% à 42% selon le cas.

En utilisant le compas, marquer un point M situé à 3cm du point A et à 4 cm du point B. (Attention : les traits de compas doivent rester visibles). A • B • TEST 6 F* R = 89%	Trouve à l'aide du compas le centre L d'un cercle passant par A et B et de rayon 4 cm. A • B • TEST 6F R = 42%

b) L'opération :

247,52 X 93,03

est réussie par 68% des élèves dans le cas du test 6A (trop facile !) où cette question est la huitième d'un test en dix questions. Cette opération n'est réussie que par 36% des élèves dans le cas du test 6A* (annexe 2) où elle apparait en troisième position dans un test en 20 questions recto-verso. (Bodin et al. 1982, fascicule 2).

c) Les questions de calcul avec parenthèses, bien réussies dans le test 6D (encore trop facile !) ne le sont plus lorsqu'elles sont placées dans le test 6A*.

3 X [7 - (12 - 10)] passe ainsi de 82% à 63%.

[35 - (27 - 19)] - [(17 - 4) - (3 - 5)] passe de 64% à 44%.

Si en plus on ajoute une difficulté, les effets s'accumulent :

4 X 11 - 3 X 7 + 2 est réussi à 65% dans 6D

mais 1 + 4 X 11 - 3 X 7 + 2 n'est plus réussi qu'à 23% dans 6A*.

Remarquons que dans tous les cas présentés, l'homogénéité des populations est attestée par la réussite égale à des tests non modifiés.

d) Dans le test 6D*, l'item :

(+ 7) + (+ 4) = ... ; (+ 9) + (- 3) = ... ; (- 7) + (- 8) = ... ; (- 4) + (- 8) = ...

est réussi par 71% des élèves (sixième).

Dans le test 5D*, l'item :

8 + 9 = ... ; - 3 + 8 = ... ; 6 - 15 = ... ; - 3 - 12 = ...

n'est plus réussi que par 41% des élèves (cinquième).

La différence ne semble pas s'expliquer par la suppression des parenthèses, mais plutôt par l'introduction, en cinquième, de la multiplication des relatifs.

Toujours dans 5D* :

(+ 4) X (- 9) = ... ; (+ 5) X (+ 7) = ... ; (- 5) X (+ 8) = ... ; (- 9) X (- 3) = ...

est réussi par 64% des élèves.

On trouve aussi des stabilités étonnantes :

(+ 73) + 0 = ... ; 0 - (+ 15) = ... ; 0 - (- 5) = ...

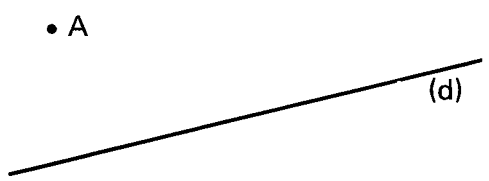
est réussi en sixième par 44% des élèves,

- 13 + 0 + 4 = ... ; (- 18) X 0 X (+ 3)

est aussi réussi par 44% des élèves en cinquième.

4) Il n'y a pas, a priori, de questions faciles, mais il en est de difficiles.

Malgré notre volonté de rester au niveau des savoirs minimums, seule une vingtaine de questions parmi les 600 que nous avons posées sont réussies par 80% ou plus des élèves. Cela ne signifie pas pour autant que les élèves ne savent rien faire, voici par exemple un item réussi à 80% en quatrième.



Tracer la perpendiculaire à (d) passant par A.
Tracer la parallèle à (d) passant par A.
Tracer une autre droite de façon à obtenir un carré.

Pourquoi décréter que cette question est très facile ? Si ce n'est justement parce qu'elle est très bien réussie. Par contre beaucoup de questions sont mal réussies, qui a priori ne mettent pas en jeu des processus plus complexes :

Test 4A :

Ecrire sous forme simplifiée

$$\frac{3}{5} - \frac{7}{4} + \frac{8}{9}$$

réussi par : 24%

Test 5F*

On a répandu uniformément 48 m^3 de gravier dans une cour rectangulaire de 50 m de long sur 32 mètres de large.
Quelle est l'épaisseur de la couche de gravier ? Réponse . . . .
Quels calculs as-tu fait ? . . . .

réussi par : 15%

Voici par exemple une question «très difficile» en sixième, mais aussi en troisième :

tracer un carré dont la diagonale mesure 10 cm.

5) Des questions du savoir dit minimum sont loin d'être maitrisées et ne sont pas pour autant objet d'apprentissages ultérieurs.

Exemples : la question :

construire un triangle ABC tel que : $\widehat{ABC} = 35°$
$\widehat{BAC} = 55°$
$AB = 3{,}8$ cm

est réussie par 30% des élèves de sixième et par 31% des élèves de troisième.

Des questions concernant l'ordre des nombres :

ranger dans l'ordre croissant 7 nombres :
— décimaux **en sixième**
— décimaux relatifs **en cinquième**
— rationnels **en quatrième**

sont toutes réussies par environ 40% des élèves.

6) Il est difficile de formuler les objectifs et de vérifier les acquisitions des élèves.

Il ne suffit pas d'énoncer : «l'élève de sixième saura calculer l'aire d'un triangle»,

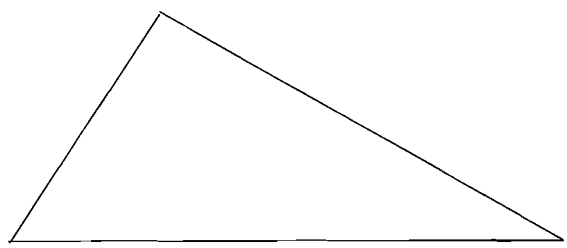
en effet, selon l'opérationnalisation retenue on pourra aussi bien affirmer que cet objectif est atteint par presque tous les élèves ou par quasiment personne.

En effet, après apprentissage.

a) Calculer l'aire d'un triangle dont un côté mesure 10 cm et la hauteur correspondante mesure 7 cm.

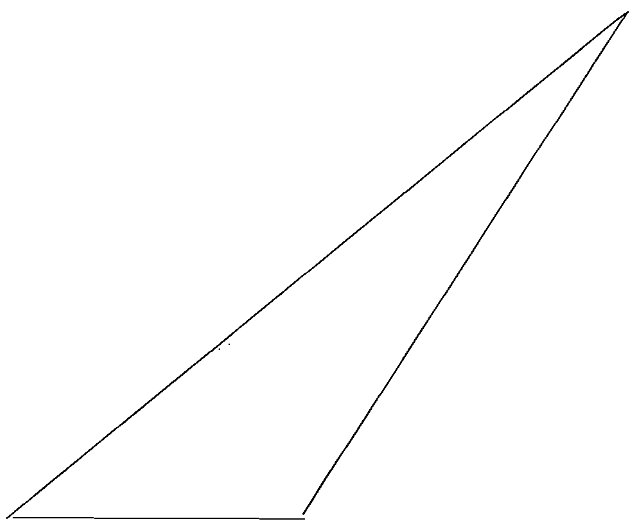
Cette question est réussie par environ 70% des élèves.

b) Effectuer les mesures nécessaires et calculer une valeur approchée de l'aire du triangle dessiné ci-dessous :



Cette question n'est plus réussie que par 46% des élèves.

c) Même question qu'en b), mais avec la figure suivante :



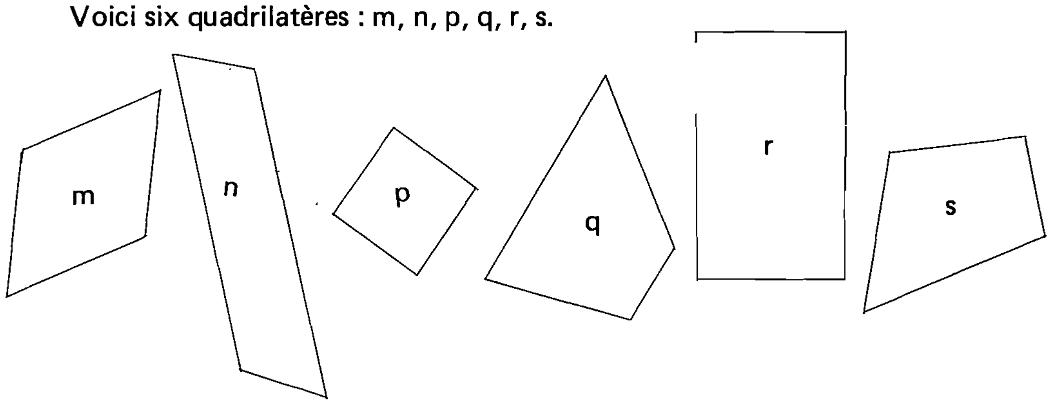
le pourcentage de réussite tombe dans ce cas à

R = 20%

Toujours en classe de 6ème (test 6E* voir en annexe 4) l'objectif «reconnaitre et distinguer parallélogramme et trapèze (non parallélogramme)», peut être aussi opérationnalisé sous les formes suivantes :

a)

Voici six quadrilatères : m, n, p, q, r, s.



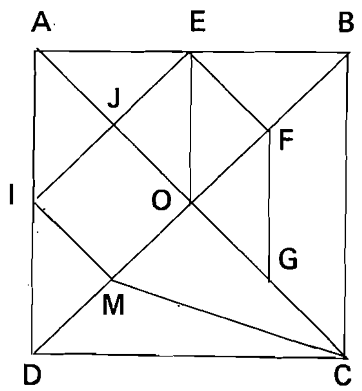
Complète les phrases ci-dessous par l'un des mots ou expressions suivants : RECTANGLE, LOSANGE, CARRE, TRAPEZE, PARALLELOGRAMME, QUADRILATERE QUELCONQUE.

choisis le mot ou l'expression qui convient le mieux, chacun d'eux ne doit être utilisé qu'une seule fois.

m est un	n est un....................
p est un.....................	r est un ....
q est un.....................	s est un ....

60% des élèves reconnaissent ici, en n un parallélogramme et en q un trapèze.

b)



Utilise les mots qui te semblent convenir le mieux pour compléter les phrases suivantes :

JEFO est un .... IEFM est un ....
EFGO est un.................... IJCM est un....................
COM est un.................... BEO est un ....

Dans ce cas, il n'y a plus que 23% des élèves qui reconnaissent en IJCM un trapèze et en EFGO un parallélogramme.

Toujours à propos de la formulation des objectifs il convient de remarquer que les questions qui les traduisent peuvent faire volontairement ou non l'objet d'un détournement de consigne de la part des élèves. Ainsi, l'exactitude d'un résultat n'est pas toujours le garant de l'atteinte d'un objectif.

Par exemple :

— calcul mental transformé en calcul écrit,

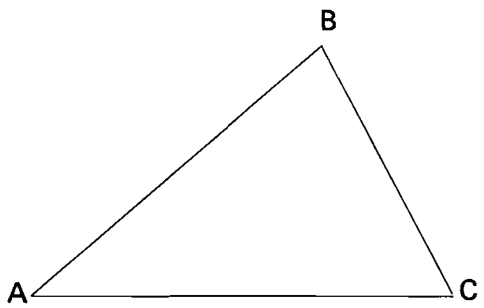
— calcul d'ordre de grandeur qui devient un calcul exact d'où l'élève déduit l'ordre de grandeur demandé,

— tâtonnements divers qui expriment sans doute une certaine habileté de l'élève mais qui peuvent être sans rapport avec l'objectif contrôlé.

7) Toutes les questions ne sont pas recevables par les enseignants.

Lorsqu'une question s'éloigne trop du questionnement habituel des enseignants, elle est souvent rejetée. C'est en particulier le cas pour des questions faisant davantage appel à l'activité de l'élève qu'à des connaissances spécifiques.

En classe de sixième, la question



Place un point S sur [AC] tel que les triangles ABS et BCS aient même périmètre.

est difficilement acceptée. En réalité elle ne correspond pas à un apprentissage et n'est réussie que par environ 5% des élèves.

Lorsque la question ne s'éloigne pas trop des habitudes elle peut être acceptée et même adoptée. Elle est alors intégrée au système d'évaluation des enseignants. L'expérience prouve qu'elle fait alors l'objet d'apprentissages systématiques et le taux de réussite des élèves augmente d'une année à l'autre.

C'est sans doute le cas pour la question suivante :

Tracer un quadrilatère dont les diagonales sont perpendiculaires et qui ne soit pas un losange.

qui en classe de 4ème est peu réussie dans nos premiers essais et qui est maintenant réussie par 55% des élèves.

Toujours en quatrième (test 4D), la question ci-dessous semble être à la limite de l'acceptable. Elle est réussie par 19% des élèves.

> Deux carrés sont tels que leurs côtés se coupent en DEUX points (c'est-à-dire que deux des points de l'un sont aussi des points de l'autre).
> TRACER trois cas de figures différentes.

Tout se passe comme si les collègues craignaient, au premier cycle, de trahir leurs élèves en leur posant des questions inhabituelles et comme si ces derniers ne manifestaient que rarement des capacités de transfert. Il y a peut-être une relation de cause à effet à rechercher dans cette observation. Quoi qu'il en soit, il semble que les questions demandant un transfert de connaissances soient beaucoup plus nombreuses dès le début du second cycle.

II – VERS UNE CLARIFICATION DES FONCTIONS DE L'EVALUATION.

Les observations qui précèdent sont ponctuellement intéressantes et riches d'enseignement pour qui doit construire des épreuves d'évaluation, mais d'une certaine façon, elles sont paralysantes. Comment repérer le niveau réel d'un élève dans un domaine déterminé ? Ce que l'on pourra observer va dépendre de tant de paramètres, la moindre différence dans la conception de l'épreuve pouvant entraîner une importante différence dans les résultats ! A moins de décider que le niveau de connaissance de telle partie des mathématiques serait par définition ce que mesurerait tel test connu de tous. On ne voit pas bien comment parvenir à une certaine harmonisation des objectifs et à la communicabilité de résultats dépourvus d'ambiguïté

Dès le début de notre travail, nous avons souhaité mettre au point un instrument d'**évaluation formative**. Il s'agissait de permettre à l'enseignant d'ajuster son action aux savoirs des élèves et aux élèves de savoir sur quels points ils devaient faire porter leurs efforts. Contrairement à l'**évaluation sommative**, destinée à faire le bilan d'une formation, cette évaluation s'inscrit à l'intérieur de la formation dont elle constitue le régulateur et, loin d'être une fin en soi, elle est un moyen au service de la construction des savoirs des élèves. Cette démarche n'est possible que si les enseignants évaluent pour connaître les savoirs des élèves, pour savoir où ils en sont dans leurs apprentissages et non pour satisfaire à telle demande de notes de l'administration. Dans la pratique, nos épreuves ont souvent été utilisées à des fins sommatives, ce que nous ne pouvons que regretter. Certes, il est nécessaire de faire des bilans, mais il conviendrait dans ce cas d'utiliser des épreuves conçues à cet effet. Tout au plus est-il acceptable d'utiliser pour l'évaluation sommative une part de l'information recueillie au cours des évaluations formatives. A condition de s'assurer que cette information reste encore valable au moment du bilan.

La nécessité ressentie par tous d'une certaine harmonisation concernant tant les objectifs poursuivis que les exigences relatives aux savoirs maîtrisés par les élèves à un palier donné, devrait conduire à la mise à la disposition des collègues, pour les besoins de l'évaluation sommative, d'épreuves référencées, étalonnées et présentant des qualités certaines de validité et de fidélité. A l'I.R.E.M. de Besançon, nous songeons à proposer de telles épreuves, mais celles que nous avons construites jusqu'à présent ne peuvent remplir cette fonction. Dans le cas de l'évaluation formative, les qualités recherchées ne sont pas essentiellement celles mentionnées ci-dessus. En fait, une bonne épreuve d'évaluation formative devrait permettre de diagnostiquer les sources d'erreurs, de remonter aux causes de l'erreur pour amener sa suppression ultérieure. Dans cet esprit, parler d'évaluation formative, c'est nécessairement parler d'évaluation individualisée. Là encore, cela ne sera possible que si les enseignants disposent d'une banque de questions correspondant à des objectifs identifiés et accompagnées des analyses des procédures habituellement suivies par les élèves, de leurs significations par rapport à l'état du savoir des élèves et des moyens de remédier aux erreurs éventuelles. La tâche est immense et il est clair que nous n'avons fait que l'ébaucher.

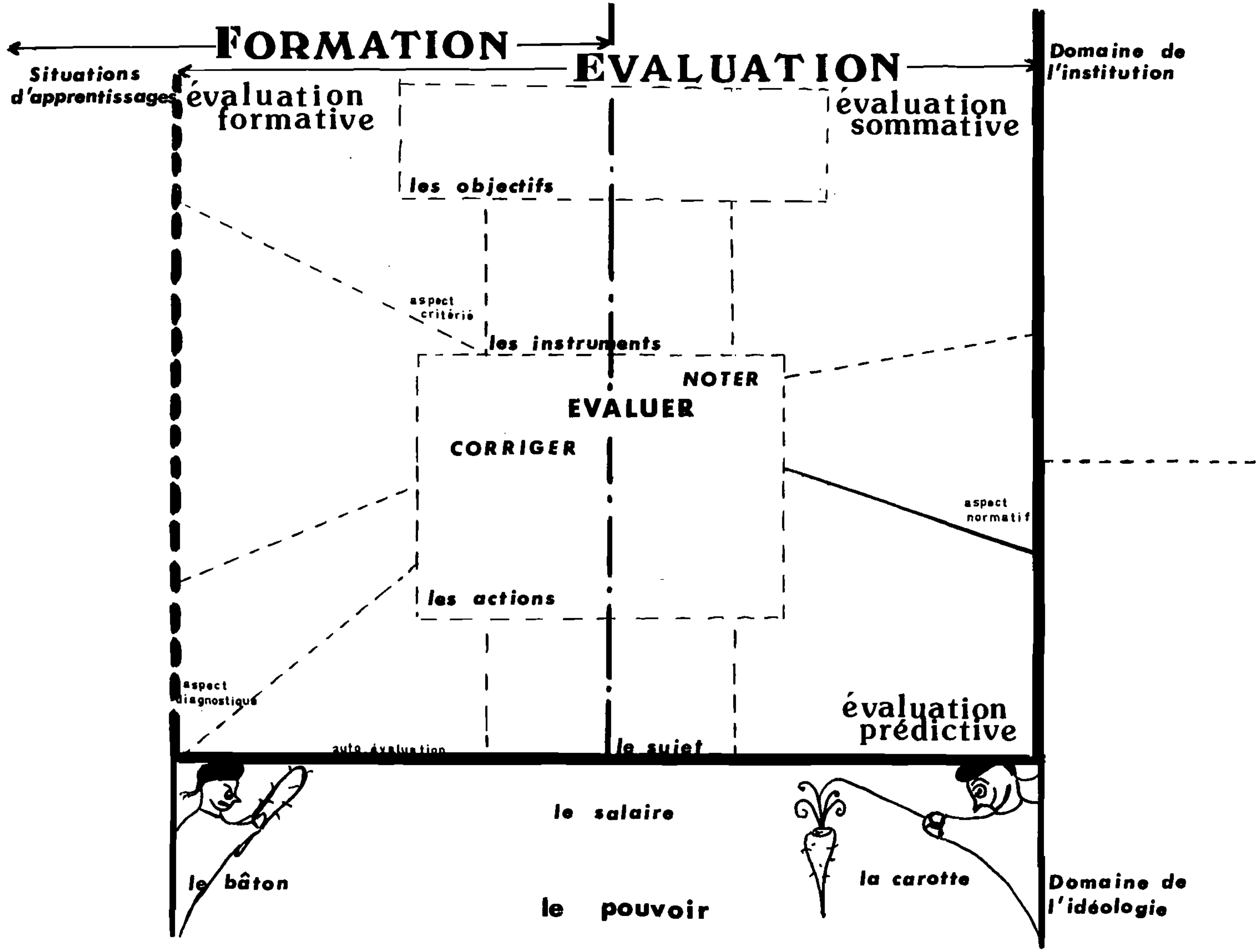
Partis d'une conception somme toute simpliste, nous sommes parvenus grâce à de nombreux apports extérieurs, à reconnaître et à essayer de prendre en compte une part de la compléxité des actions d'évaluation. Le schéma de la page suivante illustre en quelque sorte l'état de nos réflexions. En réalité, ce schéma a été élaboré pour des stages de formation pluridisciplinaire. Chaque stagiaire était invité à écrire cinq mots ou expressions en association libre avec le verbe «EVALUER». En plaçant les mots obtenus dans le schéma, il était alors possible de repérer et d'analyser les tendances et représentations individuelles et collectives. Chacun peut sans doute s'essayer à cet exercice pour se demander quel évaluateur il est ou voudrait être. Pour compléter cet exercice, il conviendrait aussi de répondre à quelques questions :

— Faut-il toujours noter, noter toujours, tout noter ?
— Peut-on évaluer sans noter ?
— Peut-on corriger sans évaluer ? Sans noter ?

Il n'est pas possible de développer ici la signification de ce schéma, on se rapportera à l'abondante littérature sur le sujet.

Attirons seulement l'attention sur deux points :

1) Le discours des sciences de l'éducation sur l'évaluation (partie centrale du schéma) est loin de recouvrir les pratiques habituelles et les représentations des enseignants et des élèves (voir le bas de la page 20).


FORMATION
EVALUATION
Situations d'apprentissages
Domaine de l'institution
évaluation formative
évaluation sommative
les objectifs
aspect critérié
les instruments
NOTER
EVALUER
CORRIGER
aspect normatif
les actions
aspect diagnostique
évaluation prédictive
auto évaluation
le sujet
le salaire
le bâton
la carotte
le pouvoir
Domaine de l'idéologie

2) Formation n'est pas évaluation (et réciproquement). S'il peut y avoir une part d'évaluation dans toute séquence de formation, il ne suffit pas de noter un travail, une production pour faire de l'évaluation. On confond trop souvent situation de formation et situation d'évaluation simplement parce que «tout travail mérite salaire» et que la note est le seul salaire envisagé. Fournir un travail personnel à l'élève peut être nécessité par la situation de formation et être pour l'élève l'occasion d'accroître et d'entraîner ses capacités. L'évaluation, si elle a lieu, devrait alors porter sur les processus, peut-être sur les comportements (investissement personnel, esprit de recherche, organisation du travail, méthodes mises en œuvre...) et non sur les résultats obtenus.

La réflexion qui précède porte sur les fonctions de l'évaluation, elle nous paraît maintenant insuffisante. On sait mieux **POURQUOI** on évalue mais on ne sait pas ce que l'on évalue **(QUOI ?)**. A notre sens, nous devrions évaluer le savoir de l'élève, ce qui exige d'avoir des idées claires sur le savoir lui-même et sur ses manifestations possibles. L'évaluation consisterait alors à repérer la part de savoir que l'élève s'est approprié, et cela de la façon la plus fidèle possible, d'où la nécessité de chercher à optimiser en ce sens les procédures d'évaluation. C'est ce point que nous essayerons de développer dans la section suivante.

III – SAVOIR ET EVALUATION.

Nous posons comme principe que l'évaluation pédagogique se donne pour but la connaissance, avec le moins d'incertitude possible, du savoir de l'élève. Savoir est pris ici au sens large et peut inclure, lorsque cela est nécessaire, aussi bien les niveaux les plus élevés des systèmes de classification d'objectifs pédagogiques que les attitudes et savoir-être. Ceci étant posé, il est clair qu'il ne peut y avoir d'évaluation sans connaissance du Savoir, Savoir considéré ici comme un absolu non médiatisé par les élèves et indépendant des démarches pédagogiques.

Le savoir.

Nous admettrons que : (voir tableaux pages suivantes)

– le SAVOIR peut être considéré comme un système[4] (système $\mathcal{S}$),

(4) Théorie des systèmes : «tentative d'unification des études portant sur les ensembles, appelés systèmes, dont les éléments sont en interaction et constituent une totalité ne se réduisant pas à la somme des parties». (Berbaum J. (1982) : «étude systématique des actions de formation», P.U.F. p. 34).

— le SAVOIR peut être organisé en champs conceptuels[5] ou en domaines de connaissances qu'il est possible de décrire. Ces domaines peuvent être eux-mêmes consédirés comme des systèmes (systèmes A),

— le SAVOIR d'un sujet est un système (système $\mathcal{S}$),

— le SAVOIR d'un sujet relativement à un domaine A est un système (système a).

Le système A, idéal ou idéel, peut se réaliser chez les sujets sous des formes variées qui sont autant d'états possibles du système a. Avant toute évaluation il importe donc de reconnaître et d'explorer ce champ du possible.

L'évaluation.[6]

L'évaluation du système a (savoir du sujet relativement à un domaine A) consiste en un ensemble d'opérations faites dans le but de prendre des DECISIONS :

— repérage ou organisation des situations propices au recueil d'informations sur l'état du système ;

— recueil de l'information pertinente ;

— examen de la compatibilité entre les informations recueillies et un certain nombre d'états possibles du système a. C'est le traitement de l'information ;

— reconnaissance ou hypothèse sur l'état de a, avec prise en compte de la notion de risque.

La décision.

La décision ne se réduit pas, en général, à la reconnaissance de l'état de a. Les décisions prises peuvent revêtir divers aspects :

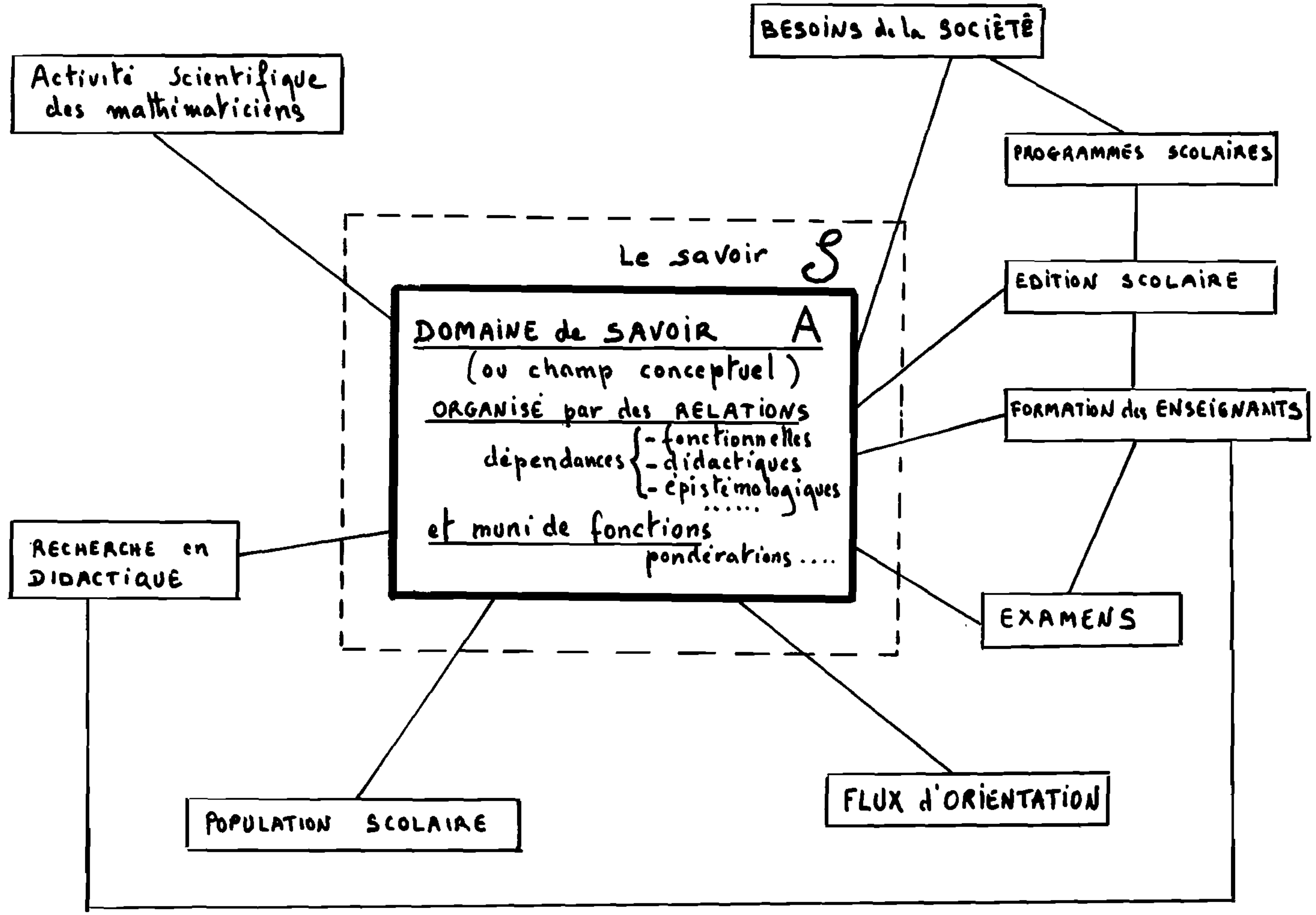
— décision de poursuivre, reprendre ou modifier une séquence pédagogique ;
— décision d'attribution d'une **note** ;
— décision de formulation d'une appréciation, d'un jugement, d'un conseil ;
— décisions de type administratif ;
— etc.

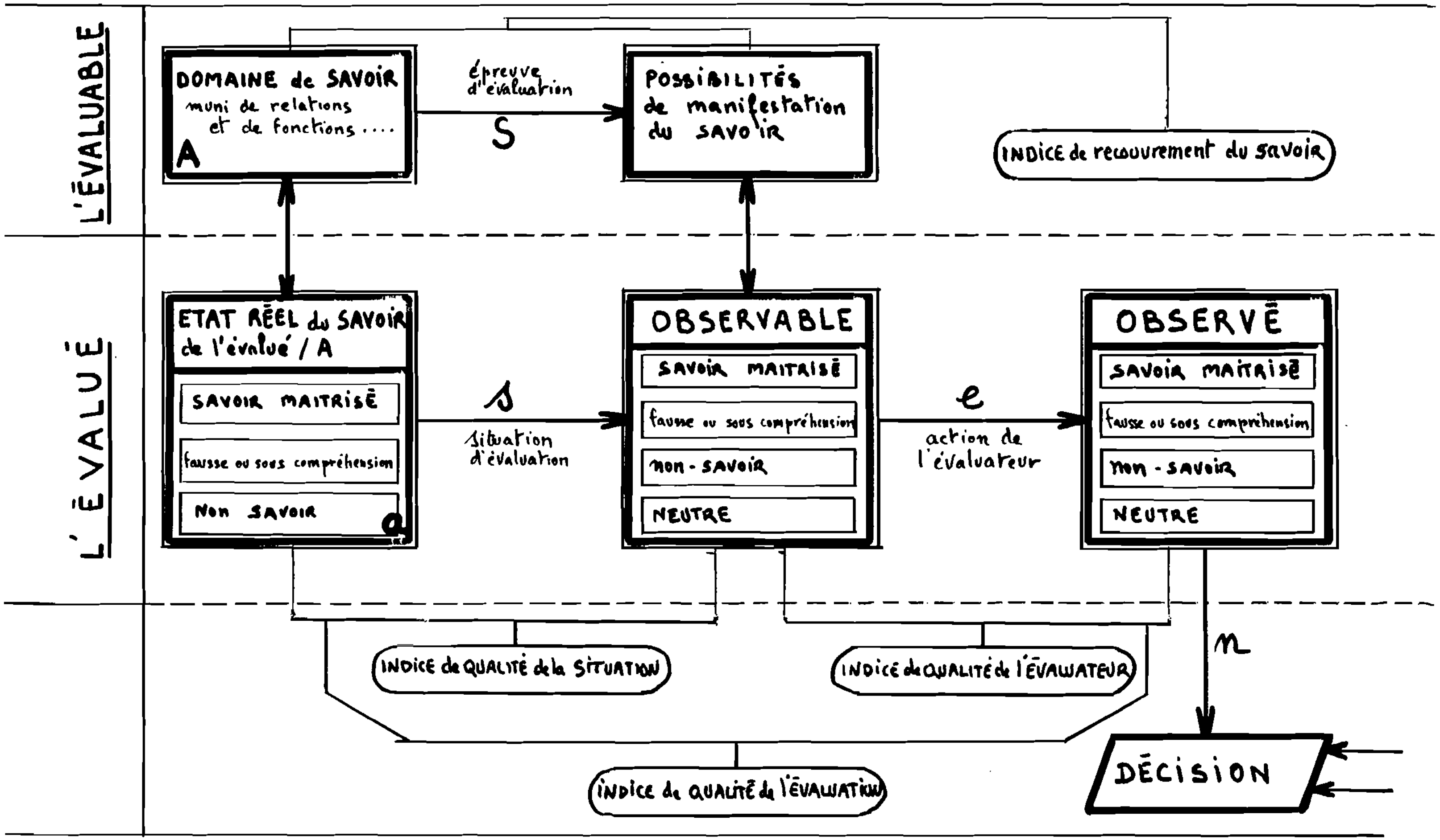
De façon plus ou moins consciente, ces décisions prennent en compte bien d'autres éléments que l'évaluation pédagogique au sens où nous venons de la définir : conceptions personnelles diverses (philosophiques, pédagogiques...), nécessité de réguler les flux, soucis d'harmonisation etc.

(5) Champ conceptuel : Espace de problèmes ou de situations problèmes dont le traitement implique des concepts et des procédures de plusieurs types en étroite connexion (Vergnaud 1981).

(6) Ces définitions s'adaptent facilement à l'évaluation d'un autre type de système (évaluation d'un groupe, d'un projet d'établissement, d'une méthode pédagogique etc.).

LE SAVOIR SCOLAIRE




L'ÉVALUABLE
L'ÉVALUÉ
DOMAINE de SAVOIR
muni de relations
et de fonctions ....
A
épreuve
d'évaluation
S
POSSIBILITÉS
de manifestation
du SAVOIR
INDICE de recouvrement du savoir
ETAT RÉEL du SAVOIR
de l'évalué / A
SAVOIR MAITRISÉ
fausse ou sous compréhension
NON SAVOIR
a
s
situation
d'évaluation
OBSERVABLE
SAVOIR MAITRISÉ
fausse ou sous compréhension
non - SAVOIR
NEUTRE
e
action de
l'évaluateur
OBSERVÉ
SAVOIR MAITRISÉ
fausse ou sous compréhension
non - SAVOIR
NEUTRE
n
INDICE de QUALITÉ de la SITUATION
INDICE de QUALITÉ de l'ÉVALUATEUR
INDICE de QUALITÉ de l'ÉVALUATION
DÉCISION

Evaluation et didactique.

Il est difficile de s'intéresser aux situations d'évaluation sans chercher à les situer par rapport aux situations d'apprentissage. De même on peut s'interroger sur la place d'une recherche en évaluation par rapport aux recherches en didactique des mathématiques. Celle-ci peut-elle être considérée comme appartenant à celle là ? Cela dépend sans doute du sens que l'on donne à certains mots.

Selon Guy BROUSSEAU (1983) :

«la didactique c'est la théorie des situations didactiques».

D'autre part :

«Une situation didactique est un ensemble de rapports établis explicitement et/ou implicitement entre un élève ou un groupe d'élèves, un certain milieu (comprenant éventuellement des instruments ou des objets) et un système éducatif (le professeur) **aux fins de faire approprier à ces élèves un savoir constitué ou en voie de constitution**».

Force est de reconnaître que les situations d'évaluation ne relèvent pas toujours de cette définition. En fait, seule l'évaluation formative est réellement concernée. Contrairement à ce qui est écrit ici ou là : «enseigner c'est évaluer, évaluer c'est enseigner», il importe de distinguer situations didactiques et situations d'évaluation, même si dans le meilleur des cas, les secondes viennent s'insérer à l'intérieur des premières pour en assurer la régulation.

La didactique étudie les conditions de modification du système $\mathcal{S}$ (savoir de l'élève) ou, plus restrictivement de systèmes a (savoir de l'élève relativement à un domaine particulier). L'évaluation s'intéresse à l'état de $\mathcal{S}$ (ou a) a un instant donné. Une recherche sur l'évaluation devrait avoir pour objectif d'optimiser les procédures d'évaluation (mieux rendre compte de ce qui est).

La recherche en didactique utilise largement l'évaluation, mais a parfois tendance à prendre pour vrai ce qu'elle croit observer. Certes dans bien des cas, de multiples précautions sont prises, mais il n'en demeure pas moins vrai que la recherche en didactique utilise un outil mal connu. La reconnaissance des systèmes A et de la variété des systèmes a est du domaine de la didactique, le paradoxe est cependant que pour ces derniers elle ne peut que s'appuyer sur des procédures d'évaluation dont la validité suppose la connaissance de ces systèmes. Ceci pour justifier la nécessité pour la recherche d'utiliser des méthodes d'investigation variées : épreuves papier-crayon de tous types, entretiens individuels etc.

Le tableau de la page 20 est à la fois un schéma d'analyse de l'évaluation concernant le QUOI et le COMMENT et un plan de recherche concernant la validité des évaluations.

Le tableau de la page 20 peut être complété par l'introduction d'indices de qualité : qualité de recouvrement du domaine de savoir par l'épreuve d'évaluation, qualité de la situation d'évaluation, qualité de l'évaluateur. La construction de ces indices devra prendre en compte la fonction assignée à l'évaluation ainsi que les relations et pondérations diverses définies sur le domaine de savoir correspondant. Cette façon de procéder devrait permettre de construire des épreuves valides ainsi que de pouvoir choisir entre plusieurs situations d'évaluation celle qui maximise l'information qu'il est possible de recueillir sur le savoir des élèves.

Il est prématuré d'affirmer que cette méthode d'analyse soit opérationnelle, qu'elle permettra une évaluation plus rigoureuse et plus pertinente. C'est en tout cas la démarche qui inspirera nos recherches ultérieures.

REFERENCES BIBLIOGRAPHIQUES.

BODIN A. et ALL. 1982 : Objectifs et évaluation fascicules 1-2-3. Besançon : IREM de Besançon.

BROUSSEAU G. 1983 : Les objectifs et la didactique des mathématiques. in Contribution à la seconde école d'été de didactique des mathématiques. Orléans : I.R.E.M. d'Orléans.

DE KETELE J.M. 1979 : Observer pour évaluer et éduquer. Bulletin de l'A.P.M.E.P., n° 317.

DE KETELE J.M. 1980 : Observer pour éduquer Berne : Peter Lang.

NOIZET G. et CAVERNI J.P. 1978 : Psychologie de l'évaluation scolaire. Paris, P.U.F.

PLUVINAGE F. 1978 : Questionnaires à questions croisées. Strasbourg, I.R.E.M. de Strasbourg.

VERGNAUD G. 1981 : Quelques orientations théoriques et méthodologiques des recherches françaises en didactique des mathématiques. Recherches en Didactique des mathématiques. Vol. 2.2.

IREM DE BESANCON

Nom :

Etablissement :

Classe :

	Calculez	Ecrivez vos résultats dans cette colonne
1	3,12 + 2,5	
2	27 + 135 + 43 + 17,8 + 24,7	
3	815 - 426,8	
4	245,8 x 37	
5	4440 : 24	
6	318 945,17 + 95 206,03 + 25 019,95 + 43 502	
7	215 647,23 - 67 948,582	
8	247,52 x 93,08	
9	217 : 23 (2 chiffres après la virgule)	
10	745,18 : 28,3 (2 chiffres après la virgule)	

Annexe 1 : test 6A

I R E M de BESANCON

TEST 6A*

nom : _______________ **classe :** __________

Pour les questions 1 à 7, fais au brouillon les calculs indiqués, essaie de vérifier les résultats obtenus, puis reportes ces résultats dans la colonne de droite. Dans le cas des divisions, lorsqu'elles ne tombent pas juste donne le résultat avec deux chiffres après la virgule.		
8 945,1 + 206,03 + 619,95 + 43,502		1
5 647,23 - 948,582		2
247,52 x 93,08		3
3,548 x 17 150		4
413,4 : 26		5
258 : 2,5		6
1 954,77 : 28,3		7
Pour les questions 8 à 14, écris le détail de tes calculs sur cette feuille, à la suite de la question, puis reportes le résultat dans la colonne de droite.		
3 + 5 x 2		8
15 - 3 x 4		
15 - 4 + 6 - 5 + 2		9
25 + 10 - 3 + 7 - 5		
1 + 4 x 11 - 3 x 7 + 2		10
19 - 4 x(3 + 1) - 3		11
17 + [5 x 6 -(7 - 2)]		12
[35 -(27 - 19)] - [(17 - 4) - (3 + 5)]		13
3 x [7 -(12 - 10)]		14

Annexe 2 : test 6A*

Pour les questions 15 à 18, tu n'as pas le temps d'effectuer les calculs indiqués. On te demande simplement, en t'aidant éventuellement d'un calcul rapide de choisir la réponse qui te semble convenir. TU ENTOURERAS CETTE REPONSE.

La somme : 49,731 28 + 253,987 648 + 807,123 456 7

est :	plus petite que 500	comprise entre 500 et 1 000	comprise entre 1 000 et 2 000	comprise entre 2 000 et 10 000	plus grande que 10 000	15

La différence : 237,548 - 174,234 9

est :	plus petite que 50	comprise entre 50 et 60	comprise entre 60 et 70	comprise entre 70 et 80	plus grande que 80	16

Le produit : 98,235 4 x 31,987 065

est :	plus petit que 1 000	compris entre 1 000 et 5 000	compris entre 5 000 et 10 000	compris entre 10 000 et100 000	plus grand que 100 000	17

Le quotient exact de 36,357 par 2,998 789

est :	plus petit que 5	compris entre 5 et 15	compris entre 15 et 30	compris entre 30 et 50	plus grand que 50	18

Complète en remplaçant dans chaque cas les pointillés par l'un des signes <ou>

103,5 110,51	17,23 13,8
16,18 16,108	0,029 0,0209

19

Remplace les pointillés par un nombre tel que les inégalités écrites soient justes. Plusieurs réponses sont possibles, il suffit d'en choisir une.

36 <..... <37	3,4> .....> 3,35

20

Nom :	Classe : Etablissement :		Colonne réservée à la correction
	Calculer :	Réponses	
1	35 -(11 + 7) =		
2	3 + 5 x 2 =		
3	(14 - 3) x 9 -(5 + 3) =		
4	1 + 3 x(8 - 4) =		
5	19 - 4 x(3 + 1) - 3 =		
6	15 - 3 x 4 =		
7	21 -(20 - 3) =		
8	4 x 11 - 3 x 7 + 2 =		
9	25 - [12 -(3 + 6)] =		
10	15 - 4 + 6 - 5 + 2 =		
11	17 + [5 x 6 -(7 - 2)] =		
12	1 + 3 x 8 - 4 =		
13	25 + 10 - 3 + 7 - 5 =		
14	[(4 x 11) - (3 x 7)] + 2 =		
15	[(15 - 4) + (6 - 5)] + 2 =		
16	3 x [7 -(12 - 10)] =		
17	(1 + 3) x (8 - 4) =		
18	[35 -(27 - 19)] - [(17 - 4) - (3 + 5)] =		
19	Mettre les parenthèses de façon à — 7 x 9 - 5 + 2 = 30		
20	ce que les égalités soient justes : — 7 x 9 - 5 + 2 = 56		

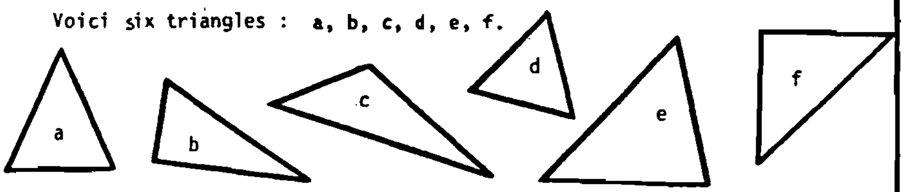
Annexe 3

I R E M de BESANCON	- 74 -	TEST 6E*

nom :_______________ classe :__________

Colonne réservée à la correction

Voici six triangles : a, b, c, d, e, f.



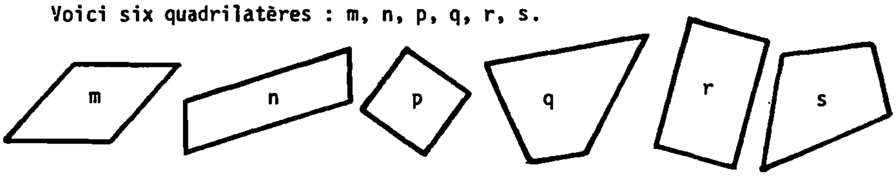
Complète les phrases ci-dessous par l'une des expressions suivantes :

TRIANGLE RECTANGLE, TRIANGLE ISOCELE, TRIANGLE QUELCONQUE, TRIANGLE EQUILATERAL, TRIANGLE AYANT UN ANGLE OBTUS, TRIANGLE RECTANGLE ISOCELE

Choisis à chaque fois l'expression qui convient le mieux. Chaque expression ne doit être utilisée qu'une seule fois.

a est un.........................	d est un.........................	1	
b est un.........................	f est un.........................	2	
c est un.........................	e est un.........................	3	

Voici six quadrilatères : m, n, p, q, r, s.

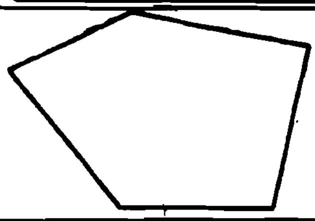


Complète les phrases ci-dessous par l'un des mots ou expressions suivants:

RECTANGLE, LOSANGE, CARRE, TRAPEZE, PARALLELOGRAMME, QUADRILATERE QUELCONQUE.

Choisis le mot ou l'expression qui convient le mieux, chacun d'eux ne doit être utilisé qu'une seule fois.

m est un.........................	n est un.........................	4	
p est un.........................	r est un.........................	5	
q est un.........................	s est un.........................	6	

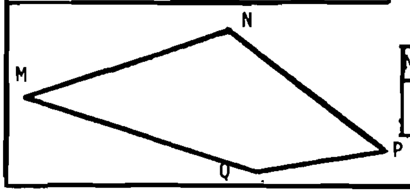


Voici un polygône.

Combien ce polygône a-t-il de sommets ?...
Combien a-t-il de côtés ?... — 7

Trace toutes les diagonales de ce polygône

Combien a-t-il de diagonales ?.... — 8

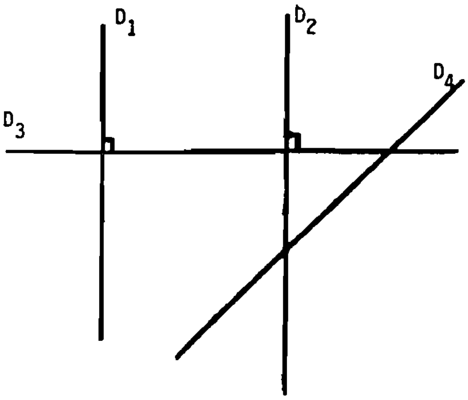


Fais les mesures nécessaires et complète le tableau ci-dessous :

Nom du segment	MN	NP	PQ	QM	MP	NQ
Longueur en centimètres			2,4			

9

Annexe 4



La figure ci-contre est formée de quatre droites : D_1, D_2, D_3, D_4.

Complète les phrases ci-dessous en utilisant les mots suivants :

PARALLELE, PERPENDICULAIRE, SECANTE.

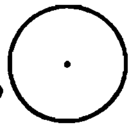
Choisis le mot qui convient le mieux, un même mot peut être utilisé plusieurs fois.

D_1 est.................. à D_3
D_2 et D_3 sont.................
D_2 et D_4 sont................. 10

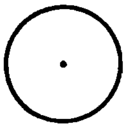
D_1 est.................. à D_2
D_1 et D_2 sont.................
D_1 et D_4 sont................. 11

Dans chacun des cas ci-dessous, fais ce qui est demandé :

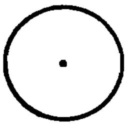
Trace un rayon de ce cercle →



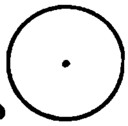
Trace un diamètre de ce cercle →

 12

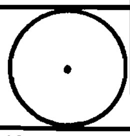
Trace une droite tangente à ce cercle →



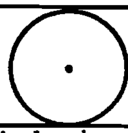
Trace une droite sécante à ce cercle qui ne soit pas un diamètre. →

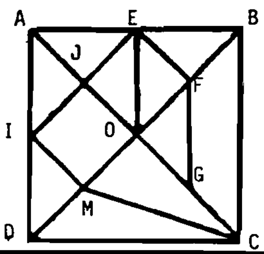
 13

Trace une corde de ce cercle qui ne soit pas un diamètre →



Trace un arc de ce cercle →

 14

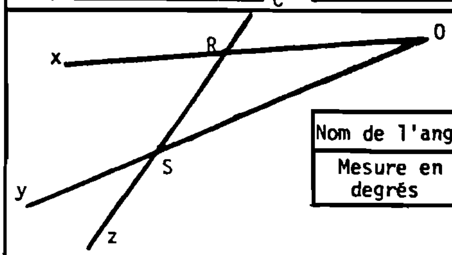


Utilise les mots qui te semblent convenir le mieux pour compléter les phrases suivantes :

JEFO est un..................................
IEFH est un.................................. 15

EFGO est un..................................
IJCH est un.................................. 16

COH est un..................................
BEO est un.................................. 17



Utilise ton rapporteur pour mesurer les angles et compléter le tableau ci-dessous : 18

Nom de l'angle	$\widehat{xOy}$	$\widehat{RSO}$	$\widehat{RSy}$	$\widehat{xRz}$	$\widehat{zRO}$	$\widehat{ySz}$
Mesure en degrés						32

19

20