

Didactique des mathématiques et de la statistique

Cours de Jean-Claude Régnier

6 La question de l'évaluation en mathématiques

Dans ce chapitre consacré à l'évaluation, nous nous contenterons uniquement de tracer les grandes lignes des problèmes posés par l'évaluation dans le domaine de l'éducation et plus particulièrement dans l'enseignement des mathématiques. Nous renvoyons le lecteur au cours intitulé ÉVALUATION.

6A Différentes approches de l'évaluation.

Il est impossible de donner une définition précise et unique de l'évaluation puisque celle-ci dépend de nombreux paramètres :

- la représentation qu'a l'enseignant de sa mission éducative et sociale ;
- la représentation qu'a l'enseignant des mathématiques ;
- l'institution scolaire dans laquelle se place l'évaluation.

Nous pouvons néanmoins dégager cinq approches différentes mais non contradictoires.

6A1 Approche « évaluation : jugement de valeur »

Le Dictionnaire de l'Académie Française de 1822 définit *évaluer* par « apprécier, fixer le prix de quelque chose, réduire l'estimation d'une chose à un certain prix. ». Cette définition se maintient dans le dictionnaire Grand Robert (1986) où *évaluer* signifie « porter un jugement sur la valeur de, le prix de ». Cette approche de l'évaluation renvoie à l'idée de valeur qui dépend évidemment de l'utilisation que l'on fait de l'objet évalué. Il semble donc difficile de conserver cette approche de l'évaluation en didactique des mathématiques, l'attribution de la valeur dépendant de critères sociologiques.

6A2 Approche « évaluation : mesure »

Dans cette approche, *évaluer* c'est simultanément :

- porter un jugement de valeur sur le résultat de cette mesure,
- donner une signification à un résultat par rapport à un cadre de référence, un critère, une échelle de valeurs.

Cette approche semble a priori plus satisfaisante, l'évaluation portant sur un objet matériel, comme très souvent une copie. Pourtant, tous les travaux de **docimologie** depuis bientôt près de cent ans ! ont montré qu'il est illusoire de croire en l'existence d'une vraie mesure. Dès les années 1920, Henri Piéron [II.6-11] réalise des travaux de docimologie à partir de la session de juin 1922 au Certificat d'études primaires. Il proposa le néologisme *docimologie* à partir des racines grecques, pour désigner *l'étude scientifique des examens*. Marie-Claire Dauvisis a abordé cette question en 1982 dans sa thèse [II.6-10] *Objectifs de l'enseignement des mathématiques et docimologie*. M. Chastrette cite dans [II.6-1] que Laugier et Weinberg ont montré que l'écart maximal de notes entre correcteurs d'une même copie de mathématiques au baccalauréat est de 9 points alors qu'il est 13 points pour une même copie de français. De plus ils ont calculé le nombre minimal de correcteurs nécessaire

pour que la note mérite une confiance élevée. Ils ont trouvé 78 correcteurs en français, 13 en mathématiques, 127 en philosophie, etc. De même, J. Nimier rapporte l'expérience menée sur des copies de mathématiques du Brevet des Collèges. 20 copies ont été corrigées par 30 professeurs de mathématiques après l'établissement d'un barème. J. Nimier a aussi corrigé ces copies sans les lire, en regardant l'écriture et la présentation. Le maximum de l'écart pour une même copie était de 11 points sur 20. La note de J. Nimier n'était jamais à une extrémité ! Nous pouvons conseiller la lecture de [II.6-1] pour connaître différents facteurs de cette fluctuation de la notation.

6A3 Approche « évaluation et objectifs pédagogiques »

Dans cette approche, on définit une liste d'objectifs d'acquisition de connaissances, de capacités de compréhension, de maîtrise d'applications, de qualités d'analyse et de synthèse. Chaque objectif peut être décomposé en sous objectifs et certaines classifications ont été élaborées comme celles de B. S BLOOM dans [II.6-2] ou de R. GRAS dans [II.6-3].

Chaque sous-objectif est évalué en terme de « atteint – non atteint » ou « atteint – en voie d'acquisition – atteint ».

Examinons une grille employée à l'école primaire avec de tels objectifs et arrêtons-nous sur le premier objectif : « *compléter une figure par symétrie par rapport à une droite* ». Denise Grenier a repéré dans [II.6-4] un certain nombre de **variables didactiques** pour la symétrisation d'une figure :

- la complexité de la figure à symétriser (segments droits/courbes, angles droits, ..),
- la présence ou non quadrillage,
- l'orientation de l'axe dans la feuille (parallèle ou non aux bords de la feuille),
- l'axe coupe ou non la figure à symétriser.

Ceci montre toute la difficulté à évaluer un objectif en terme « atteint – non atteint » surtout si les valeurs des variables didactiques ne sont pas précisées.

N : niveau 1 ou 2		GÉOMÉTRIE							
Connaître les figures usuelles									
N	Être capable de :	A	B	C	Être capable de :	A	B	C	N
	Compléter une figure par symétrie par rapport à une droite				Calculer l'aire d'un carré, d'un rectangle				
	Calculer le périmètre de polygones usuels				Résoudre des problèmes utilisant périmètres et aires				
A : Acquis B : Acquisition à vérifier C : Non acquis									

Pourtant cette pratique d'évaluation rencontre encore quelques succès. Néanmoins, en plus de l'inconvénient majeur cité plus haut, elle reste critiquable sur d'autres points :

- la parcellisation des tâches ;
- l'absence de prise en compte du tâtonnement ;
- le peu de travail des concepts fondamentaux.

6A4 Approche « évaluation – négociation »

Yves Chevallard a montré dans [II.6-5] que l'évaluation est un élément important de la négociation de la **relation didactique** entre l'enseignant et la classe à propos du savoir. Citons Y. Chevallard : « Trop peu d'exigence, et c'est l'annulation d'une partie du travail qu'il [enseignant] aura fait jusque là. Mais trop d'exigence, en sens inverse, le rendrait non crédible et susciterait de vifs conflits, au risque de rendre la classe ingouvernable : que penserait-on d'un commerçant qui vendrait 59 francs ce que d'autres vendent autour de 15 francs ? ».

Ainsi les élèves et enseignants se livrent à un **jeu didactique**, dissymétrique dans lequel les élèves, face à la stratégie de l'enseignant, usent de tactiques destinées à modifier le déroulement de la vie de la classe en leur faveur. Ainsi dans cette conception didactique de l'évaluation, la note mise sur une copie n'est pas une mesure. Elle est alors considérée plutôt comme une **transaction** entre l'enseignant et l'apprenant. Cette entente implicite rentre dans le cadre d'attentes réciproques de la part de l'enseignant et de l'élève. Ces attentes font partie du **contrat didactique**.

Toujours dans ce texte, Y. Chevallard prend l'exemple d'une classe en analysant les moyennes et les écart-types des notes attribuées à chaque devoir sur une année scolaire. Il montre ainsi qu'après un devoir dit « mauvais » du point de vue de l'enseignant, c'est à dire quand la moyenne des notes qu'il a attribuées à ce devoir était jugée « faible » avec une valeur de l'ordre de « 8 sur 20 » pour un écart-type jugé « élevé » de l'ordre de « 6 sur 20 », le groupe-classe se scindait de lui-même en deux sous-groupes : celui des « forts » et celui des « faibles ». Cette reconfiguration rendait très difficile la gestion pédagogique et didactique par l'enseignant. Dans cette situation observée par Y. Chevallard, chaque fois, l'enseignant, pour sauver la **relation didactique** avec le groupe-classe, proposait un devoir *nettement plus facile* dont les caractéristiques statistiques se modifient dans le sens d'une augmentation de la note moyenne et de la diminution de l'écart-type. Dans l'exemple rapporté la moyenne était passée à une valeur de l'ordre de 11,5 et l'écart-type, de l'ordre de 4.

6A5 Approche « évaluation – décision »

Une autre approche consiste à s'intéresser à la **décision** qui motive l'évaluation. Dans ce cadre J.M De Ketele [II.6-6] définit ainsi l'évaluation : « *L'évaluation est le processus qui consiste à recueillir un ensemble d'informations pertinentes, valides et fiables puis à examiner le degré d'adéquation entre cet ensemble d'informations et un ensemble de critères choisis adéquatement en vue de fonder la prise de décision* ».

Antoine Bodin [II.6-7] résume ainsi cette définition : « *L'évaluation est l'ensemble des procédures et des processus de recueil, de traitement et de communication d'informations effectuées dans le but de prendre des décisions.* »

Citons quelques décisions possibles après une évaluation :

- certifier,
- faire un bilan intermédiaire,
- classer,
- prédire la réussite,
- diagnostiquer,
- sélectionner
- hiérarchiser,
- etc..

C'est à partir de la décision à prendre que l'on va distinguer les différentes formes de l'évaluation.

6B Les fonctions de l'évaluation

Nous distinguons ici trois grandes fonctions de l'évaluation ; les fonctions sommative, formative, et prédictive.

6B1 L'évaluation sommative

C'est ce type d'évaluation que l'on met en œuvre lors de bilans au terme d'un processus d'apprentissage ou de formation. Cette évaluation vise à prendre une décision d'orientation ou de sélection en fonction des acquis. Elle permet de plus de situer les élèves les uns par rapport aux autres. On distingue habituellement « évaluation normative et évaluation critériée ». Une évaluation sera dite **normative** quand elle réfère la performance d'un apprenant aux performances des autres apprenants. En revanche, elle sera dite **critériée** quand on ne compare pas l'apprenant aux autres mais qu'on détermine par la référence à des critères, si, ayant atteint les objectifs, il est en mesure de passer aux apprentissages ultérieurs.

Il est clair qu'un enseignant, quand il fait un bilan trimestriel avec sa classe, son évaluation est de type sommatif et qu'elle relève tout autant de l'évaluation normative que de l'évaluation critériée. Comme l'a souligné Y. Chevallard, l'enseignant doit gérer la classe après l'évaluation et donc son évaluation ne peut être qu'uniquement critériée. Il est contraint de tenir compte de faits contingents comme le niveau de la classe, la dispersion des notes, etc..

Nous verrons dans le paragraphe suivant la difficulté, sur des exemples précis, d'élaborer des critères avec des indicateurs. Nous rappelons qu'un indicateur d'un objectif est un indice visible et objectif pris dans la production de l'élève qui permet de dire si l'objet en question répond au critère. Par exemple, si on s'intéresse à l'objectif « calculer la somme de deux nombres entiers », définir un indicateur, c'est définir clairement et objectivement les conditions avec lesquelles on peut décider que l'objectif est atteint. Très souvent, un indicateur est le nombre de réussites lors d'addition. On acceptera, par exemple, qu'un élève réussisse 8 à 10 additions sur dix pour affirmer qu'il atteint l'objectif. Mais il est possible que ces deux échecs montrent clairement que l'élève n'a pas acquis la technique d'addition dans une situation particulière (par exemple dès qu'il y a une retenue sur les chiffres des unités). Alors que dire sur l'acquisition de la compétence « ajouter deux entiers », peut-on dire qu'elle est atteinte ?

6B2 L'évaluation formative

L'évaluation **formative** intervient dans le cours d'un apprentissage et permet de situer la progression de l'élève par rapport à un objectif donné.

Dans le cadre d'une évaluation formative, l'objectif est d'obtenir une double rétroaction :

- sur l'élève pour lui indiquer les étapes qu'il a franchies et les difficultés qu'il rencontre,
- sur le maître pour lui indiquer comment se déroule son programme pédagogique et quels sont les obstacles auxquels il se heurte.

Ainsi, nous pouvons remarquer :

- que l'évaluation formative s'adresse à l'élève et à l'enseignant,
- qu'elle implique l'élève dans son apprentissage en lui donnant des points de repère,

- qu'elle relève les difficultés de l'élève pour les pallier et non pour les sanctionner,
- qu'elle sert à l'enseignant pour réguler ses stratégies pédagogiques.

Pourtant, dans la pratique enseignante de l'évaluation, la distinction entre évaluation formative et évaluation sommative n'est pas aussi claire qu'on est en droit de penser. Notons, comme Michel Henry dans [II.6-8] que : « Pourtant les résultats obtenus à des tests ou des épreuves en cours d'apprentissage sont le plus souvent chiffrés et interviennent en fin de compte dans un bilan sommatif. On ne peut pas empêcher les élèves d'avoir le sentiment d'une rupture de contrat, même si le professeur tempère ses notes par une appréciation susceptible de rendre compte de l'évolution du travail ou des acquis. »

6B3 L'évaluation prédictive

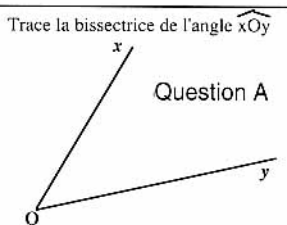
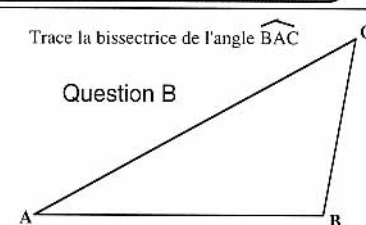
Avec ce type d'évaluation, la décision à prendre est de prédire, à partir d'informations convenables, la réussite future d'une personne, dans une situation scolaire ou professionnelle. Elle suppose qu'il existe un modèle de fiabilité suffisante qui permet de lier les prédicteurs dans l'évaluation aux critères de réussite ultérieure et que les conditions dans lesquelles le modèle a été établi sont réalisées dans le cas particulier concerné. Ces deux conditions sont rarement satisfaites et on a même de grandes difficultés à prouver qu'elles le sont, l'entreprise est donc délicate. Même si le traitement statistique permet de dégager des constantes (par exemple, le redoublement au CP serait un mauvais pronostic pour la réussite scolaire), il est difficile de les réinvestir sur des individus particuliers. On mesure aussi la difficulté de l'orientation fondée sur un rôle prédictif confiée à une évaluation sommative qui s'est en fait déroulée dans un contexte formatif !.

6C Quelques exemples d'analyse d'évaluation

Nous empruntons ces exemples à Antoine Bodin dans [II.6-9]

6C1 La bissectrice

Dans une évaluation, à échelle nationale, organisée par l'A.P.M.E.P au niveau de la classe de sixième, nous pouvons voir ceci :

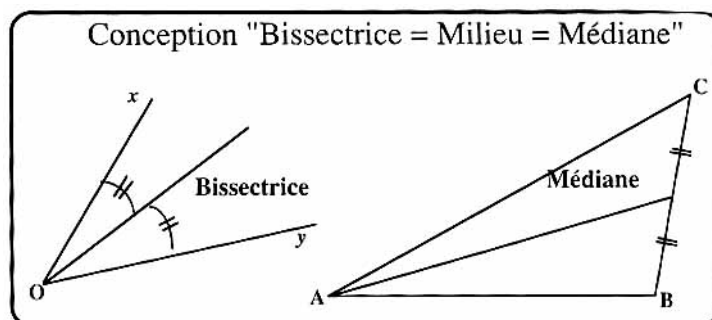
Questions EVAPM 6/87 (fin de Sixième)	
<p>Trace la bissectrice de l'angle \widehat{xOy}</p>  <p>Question A</p> <p>Réussite : 69 %</p>	<p>Trace la bissectrice de l'angle \widehat{BAC}</p>  <p>Question B</p> <p>Réussite : 28 %</p>

Bodin A, (1997), L'évaluation du savoir mathématique, Questions et méthodes, *Recherches en didactique des mathématiques*, (17.1) Grenoble : La Pensée Sauvage

Les deux questions correspondent au même type de tâche : « tracer la bissectrice d'un angle ». Dans les deux cas, la technique est identique et pourtant les réussites sont très différentes. Comment expliquer cette différence ?

Une première hypothèse avancée est qu'il y a trois bissectrices dans un triangle et donc des possibilités de confusion avec ces trois bissectrices. Cette hypothèse ne résiste pas à l'analyse des productions d'élèves : ils ne se sont pas trompés d'angle dans le triangle.

L'hypothèse que formule A. Bodin est la suivante : « Dans les deux cas, les élèves utilisent un intermédiaire qui est un milieu, mais dans le second cas, ils préfèrent prendre le milieu du côté opposé, ce qui, en particulier leur permet d'utiliser une règle graduée, plutôt que le compas, qui leur paraît sans doute d'un usage moins facile. Insistons sur le fait que pour les élèves les deux procédures sont équivalentes. Nous avons sans doute là un bel exemple de conception erronée qui a son domaine de validité propre (triangle isocèle), et qui, de plus, passe souvent inaperçue. En effet, pour que la différence entre médiane et bissectrice apparaisse nettement sur la figure, il faut que l'angle B soit assez grand par rapport à l'angle C, ce qui est rarement le cas dans les situations rencontrées par les élèves »



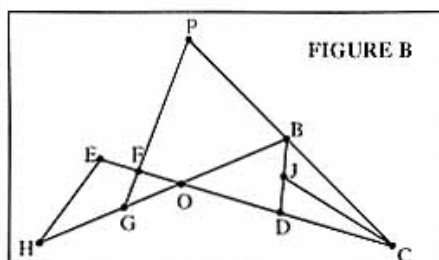
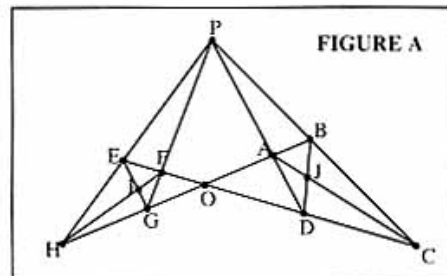
Bodin A, (1997), L'évaluation du savoir mathématique, Questions et méthodes, *Recherches en didactique des mathématiques*, (17.1) Grenoble : La Pensée Sauvage

Cet exemple illustre parfaitement ce que nous disions à propos de la validité de l'atteinte d'un objectif et de la variabilité des résultats selon le contexte de la situation dans lequel le savoir ou le savoir-faire est intégré.

6C2 L'alignement de points

Voici la même question posée à des élèves de CE2 et à des élèves de Sixième.

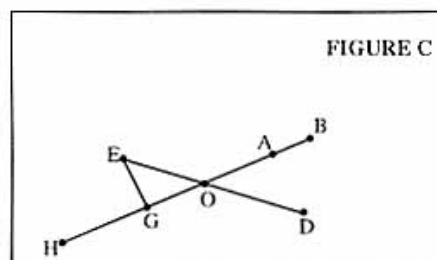
Question SPRESE CE2 1981 - Evaluation du Ministère de l'Education Nationale
Reprise IREM de BESANÇON Sixième 1985



Complète cette figure B pour obtenir la même que celle qui est encadrée en haut de la page (figure A).

TRACE tous les segments qui manquent
MARQUE tous les points qui manquent

Pourcentages de réussites
SPRESE CE2 : 86 %
IREM BESANÇON Sixième : 90 %



Complète cette figure C pour obtenir la même que celle qui est encadrée en haut de la page (figure A).

TRACE tous les segments qui manquent
MARQUE tous les points qui manquent

Pourcentages de réussites
SPRESE CE2 : 50 %
IREM BESANÇON Sixième : 38 %

Bodin A, (1997), L'évaluation du savoir mathématique, Questions et méthodes, *Recherches en didactique des mathématiques*, (17.1) Grenoble : La Pensée Sauvage

Il est frappant de constater que la réussite est à peu près identique pour la figure B mais qu'elle est nettement plus faible en classe de Sixième qu'en classe de CE2 pour la figure C.

Essayons de comprendre cette différence de performances.

Si on examine la tâche demandée à la figure B, une technique assez *naturelle* est de tracer [EP], [EG], [FH], [PD], et [CA] en prolongeant [CJ]. Une autre technique serait de tracer [EG] ou [HF] et de mesurer soit [EJ] soit [HJ] pour placer le point J. On pourrait aussi mesurer [OA] pour placer le point A.

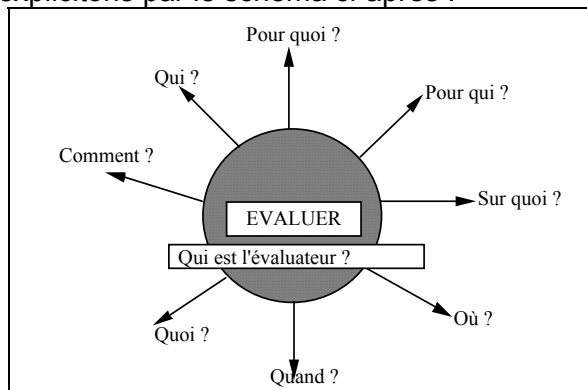
Si on examine la tâche demandée à la figure C, les deux techniques différentes précédentes sont encore applicables. Seulement, pour la première technique, les segments à tracer en premier ne sont pas aussi évidents à trouver. En revanche, mesurer [DC] pour placer le point C se repère assez bien. Laissons A. Bodin expliquer cette différence de réussite :

« Dans la figure C, le point P doit être placé en premier, à moins de mesurer et de prendre en compte les propriétés métriques de la figure. C'est ce que font les élèves de Sixième. Au CE2, les élèves utilisent une procédure affine et ne prennent en compte que les alignements. En Sixième où l'on utilise davantage les instruments de mesure qu'au CE2 (règle graduée), les élèves ont tendance à essayer de s'en servir. L'aspect métrique de la figure est beaucoup plus souvent pris en compte qu'au CE2. Mais chacun admettra qu'il faut être plus adroit pour réussir avec la procédure métrique qu'avec la procédure affine. »

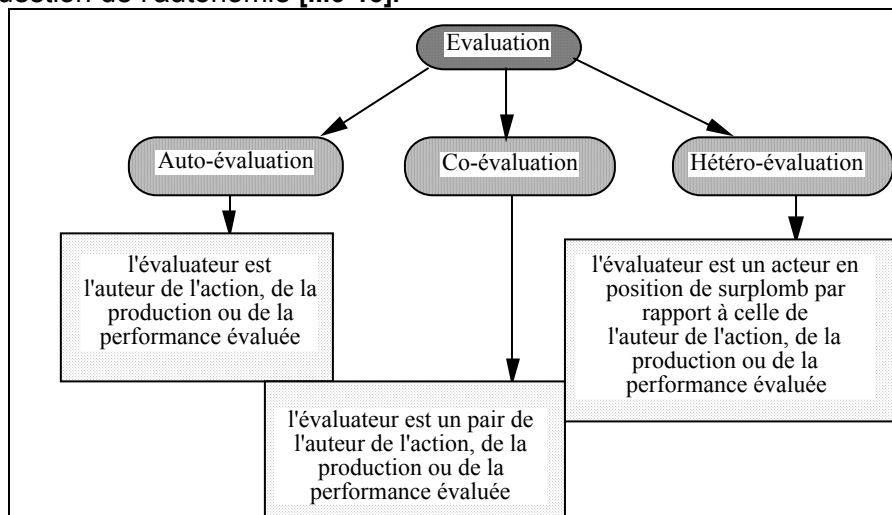
Nous avons là un exemple d'explicitation du **contrat didactique** concernant les constructions de figures géométriques en classe de CE2 et en classe de Sixième.

6D Apport pour l'enseignant : un questionnement fondamental de l'évaluation

Si nous faisons le point sur à quoi répond le processus d'évaluation indépendamment de la conception de l'évaluateur, il appert que nous retombons sur le fait que d'une certaine manière, une situation d'évaluation est toujours la réponse singulière au questionnement fondamental que nous explicitons par le schéma ci-après :



Sinon nous examinons la question centrale celle de l'auteur de l'évaluation : *qui est l'évaluateur ?*, nous pouvons introduire les trois perspectives que nous nommons : auto-évaluation, co-évaluation et hétéro-évaluation. Jean-Claude Régnier a développé ce point de vue dans ses travaux de recherche, en particulier dans [II.6-14] *Auto-évaluation et autocorrection dans l'enseignement des mathématiques et de la statistique*, en tentant de le relier à la question de l'autonomie [II.6-15].



Quand un enseignant organise des situations d'enseignement-apprentissage, il y intègre aussi l'organisation des situations d'évaluation. Le questionnement fondamental lui sert alors de guide d'aide à l'élaboration de la situation d'évaluation.

Ajoutons qu'à l'heure où de grandes enquêtes internationales telle que PISA [II.6-13] sont conduites, il lui est plus que jamais indispensable de recourir à ce questionnement fondamental pour étayer la prise de distance par rapport aux résultats qui en ressortent. Il s'agit d'éviter la dogmatisation des conclusions sans toutefois se priver de leurs apports.

Nous pouvons prendre les résultats en compte comme des informations sérieuses et pertinentes.

Pour aller plus loin...

- [II.6-1] **Chastrette M.** (1989), *Démarches et outils de l'évaluation*, Lyon, IREM de Lyon
- [II.6-2] **Bloom B.** (1979), *Taxonomie des objectifs pédagogiques, domaines cognitifs et affectifs*, Presses Universitaires du Québec
- [II.6-3] **Gras R.** (1979), *Contribution à l'étude expriment ale et à l'analyse de certains acquis cognitifs et de certains objectifs didactiques en mathématiques*, Thèse de Doctorat, Université de Rennes I
- [II.6-4] **Grenier D.** (1988), *Construction et étude du fonctionnement d'un processus d'enseignement sur la symétrie orthogonale en sixième*, Thèse de doctorat, Université Joseph Fourier Grenoble
- [II.6-5] **Chevallard Y.** (1986), Vers une analyse didactique des faits d'évaluation, in J.-M. De Ketele (éd.), *L'évaluation : approche descriptive ou prescriptive ?* De Boeck-Wesmael, Bruxelles
- [II.6-6] **De Ketele J.M.** (1980), *Observer pour éduquer*, Bern, Peter Lang
- [II.6-7] **Bodin A.** (1989), *L'évaluation du savoir mathématique*, APMEP, Bulletin n°368,
- [II.6-8] **Henry M.** (1991), *Une présentation de la didactique en vue de la formation des enseignants*, Besançon, IREM de Franche-Comté
- [II.6-9] **Bodin A.** (1997), L'évaluation du savoir mathématique, Questions et méthodes, *Recherches en didactique des mathématiques*, (17.1) Grenoble : La Pensée Sauvage
- [II.6-10] **Dauvisis, M.C.**, (1982) *Objectifs de l'enseignement des mathématiques et docimologie* Thèse de Doctorat d'État, Univ. Toulouse -le-Mirail.
- [II.6-11] **Piéron, H.** (1969, 1963) *Examens et Docimologie* Paris PUF
- [II.6-12] **Glaeser, G.** (1995) *Fondements de l'évaluation en mathématiques* Brochure N°96 coll. Formation des enseignants Paris A.P.M.E.P.
- [II.6-13] **PISA** L'évaluation internationale PISA 2003 : compétences des élèves français en mathématiques, compréhension de l'écrit et sciences
<http://www.education.gouv.fr/cid4845/l-evaluation-internationale-pisa-2003.html>
- (Organisé sous l'égide de l' O.C.D.E., le programme international pour le suivi des acquis des élèves (PISA) est une vaste enquête qui a lieu tous les trois ans, et vise à évaluer les élèves de 15 ans sur leur capacité à mobiliser et appliquer leurs connaissances dans des situations de la vie de tous les jours. Leurs compétences sont ainsi régulièrement mesurées dans trois domaines : compréhension de l'écrit, culture mathématique, culture scientifique, sous la forme d'exercices ayant fait l'objet d'un consensus entre tous les pays participants.)
- [II.6-14] **Régnier, JC** (2000) *Auto-évaluation et autocorrection dans l'enseignement des mathématiques et de la statistique. Entre praxéologie et épistémologie scolaire*. Note de synthèse HDR Université Marc Bloch Strasbourg
http://demeter.univ-lyon2.fr:8080/sdx/theses/lyon2/2000/regnier_jc
- [II.6-15] **Régnier, JC** (1983) *Évaluation et Autonomie*, Strasbourg : IREM-Université L. Pasteur
- [II.6-16] **Barbier, J.-M.**, (1985) *L'évaluation en formation*, Paris : PUF.