

Université d'été

**L'évaluation des élèves :
acte pédagogique, aide à la
décision, pour l'élève, la
classe, l'institution**

Compte-rendu de Christophe Escartin

Université de Nantes - 8 au 12 juillet 2002

Sommaire

Conférence par Claude THELOT : Les évaluations des élèves dans l'ensemble du dispositif d'évaluation du système éducatif ; quelle utilisation de ces évaluations ?	5
I. <i>L'évaluation : un instrument interne</i>	5
II. <i>L'évaluation : un instrument du débat public</i>	5
III. <i>Les outils et les usages</i>	5
IV. <i>Conclusion</i>	6
Table ronde avec Jean-Claude ÉMIN , Agnès FLORIN et Élisabeth BAUTIER : Les aspects sociaux et les aspects cognitifs des apprentissages.....	7
I. <i>Intervention d'Élisabeth BAUTIER</i>	7
II. <i>Intervention de Jean-Claude ÉMIN</i>	8
III. <i>Intervention de Agnès FLORIN</i>	8
Intervention de Luce DOSSAT : Pourquoi des évaluations diagnostiques ?	9
I. <i>L'évaluation diagnostique, pour quoi faire ?</i>	9
II. <i>Les hypothèses d'apprentissage</i>	9
ATELIER 1 : Qu'est-ce que construire un outil d'évaluation diagnostique ?	10
I. <i>La place de l'évaluation diagnostique</i>	10
II. <i>Les problèmes de construction de l'outil diagnostique</i>	11
III. <i>Conclusion</i>	11
ATELIER 2 : Interpréter les résultats des évaluations, l'exemple des protocoles CE2/6^e	12
I. <i>Points abordés au cours de l'atelier</i>	12
II. <i>Trois points de vue</i>	12
ATELIER 3 : La notion de compétences	13
I. <i>Les participants</i>	13
II. <i>Les moments d'échange</i>	13
III. <i>Les idées centrales exprimées</i>	13
ATELIER 4 : L'accompagnement à l'évaluation et ses usages	14
I. <i>La nouvelle évaluation diagnostique GS/CP</i>	14
II. <i>Dispositifs d'accompagnement à l'usage des évaluations diagnostiques</i>	15
Conférence par Jacques JUHEL : Faut-il –est-il utile de- mesurer pour évaluer les élèves ?	17

<i>I. L'évaluation des élèves</i>	17
<i>II. Qu'est-ce qu'évaluer ?</i>	18
<i>III. Quelle place pour la mesure de l'activité d'évaluation ?</i>	18
Conférence par Alejandro TIANA : Approche de l'évaluation des élèves au niveau international : ce qu'apportent les enquêtes internationales sur les acquis des élèves ; le regard d'un pays étranger sur un système éducatif ; la France : quelle particularité ?	19
<i>I. Quelques idées sur l'approche de la comparaison internationale des acquis des élèves</i>	19
<i>II. Le regard d'un étranger sur le système d'évaluation français</i>	20
Table ronde avec Jean-Claude ÉMIN (France), Manuel PALACIOS (Brésil), Tatiana FORENKO (Russie), Szuzsanna HORVATH (Hongrie), Anna BUTASOVA (Slovaquie), Jean- Richard CYTERMANN (France) : Les systèmes nationaux d'évaluation en France et chez nos partenaires du Brésil, de la Hongrie, de la Slovaquie et de la Russie	21
<i>I. Le système d'évaluation brésilien</i>	21
<i>II. Le système d'évaluation hongrois</i>	21
<i>III. Le système d'évaluation russe</i>	22
<i>IV. Le système d'évaluation slovaque</i>	22
<i>V. Le système d'évaluation français</i>	23
Présentation Les évaluations internationales et leur utilisation : PIRLS par Marc COLMANT ; PISA par Isabelle ROBIN et Thierry ROCHER	24
<i>I. PIRLS : Progress in International Reading Literacy Study</i>	24
<i>II. PISA : Premiers résultats de l'évaluation internationale des acquis des élèves de quinze ans</i>	24
Conférence par André FLIELLER : Évaluation, comparaison, comparabilité	26
<i>I. Les évaluations collectives</i>	26
<i>II. Les évaluations individuelles</i>	27
<i>III. Évaluations sans comparaison (entre individus)</i>	28
Intervention par Josette LE COQ : En quoi les informations chiffrées quantitatives peuvent être utiles à l'enseignant, de quoi l'inspection territoriale peut-elle disposer pour faire de la formation ?	29
ATELIER 1 : La communication des résultats aux évaluations (aux enseignants, aux parents, à l'institution)	29
ATELIER 2 : Comment exprimer les niveaux de compétences des élèves (échelle et calibrage)	29
ATELIER 3 : L'évaluation du non cognitif	29

ATELIER 4 : L'utilisation des résultats et des données par les responsables locaux de l'éducation dans différents contextes et à différents moments	30
Conférence par Michel FAYOL : Évaluation et apprentissage	33
<i>I. Évaluation et apprentissage</i>	<i>33</i>
<i>II. Approche cognitive</i>	<i>33</i>
<i>III. Questions posées par l'assistance</i>	<i>36</i>
Bibliographie provisoire	37
Programme de l'université d'été	38

Conférence par Claude THELOT ¹

Les évaluations des élèves dans l'ensemble du dispositif d'évaluation du système éducatif ; quelle utilisation de ces évaluations ?

I. L'évaluation : un instrument interne

Il y a trente ans, l'industrie avait mis en place des évaluations du produit (en fonction d'un cahier des charges). Or, aujourd'hui, les industriels n'évaluent plus que le processus de production... Si l'on transposait ce système dans l'Éducation nationale, on évaluerait le processus d'enseignement. Cependant, comme dans le domaine de la Santé, le sujet (malade ou élève) contribue à sa réussite, est facteur de sa réussite ; cela ne dépend pas complètement de l'enseignant, de la formation de cet enseignant, etc. Il faut donc mesurer le résultat. C'est pourquoi il faut que les outils d'évaluation soient les meilleurs possibles, soient parfaits, aussi bien pour les évaluations en classe que pour les examens.

II. L'évaluation : un instrument du débat public

Pourquoi doit-on évaluer ? Dans une démocratie, il faut dire à la société en quoi le système éducatif réussit ou échoue. Il faut donner une image réelle de l'état de l'école. Donc l'évaluation des élèves est indispensable pour évaluer les élèves eux-mêmes, les établissements mais également le système éducatif.

Si l'école est transparente, on peut demander sur quoi il faut faire porter les priorités d'apprentissage. L'évaluation n'est pas un but en soi mais un bon instrument « de progrès » ; c'est même le seul. L'évaluation doit permettre d'agir mieux. C'est un instrument de progrès que l'on fait connaître aux acteurs, aussi bien éducatifs (professeurs, chefs d'établissements, recteurs, etc.) que sociaux (parents, etc.).

III. Les outils et les usages

A. Les outils

1) Distinction entre les deux outils « évaluation bilan » et « évaluation diagnostique ».

a) Les caractéristiques de l'évaluation bilan

- L'évaluation bilan est en fin : en fin de cycle, en fin d'année, etc. Il faudrait créer d'autres évaluations à la fin du collège, à la fin du lycée et à l'université (en fin de licence).
- Les évaluations bilan permettent une comparabilité dans le temps (d'une année sur l'autre) et dans l'espace (d'un pays à un autre par exemple).
- L'évaluation bilan doit se détacher des programmes et se rapprocher des compétences : compétences qui sont nécessaires à la vie, qui sont exigées dans la société.
- Nous ne rendons pas les évaluations bilan suffisamment lisibles ce qui amène à des contresens à de mauvaises interprétations.

¹ Président du Haut Conseil de l'Évaluation de l'École à la Cour des Comptes.

b) Les caractéristiques de l'évaluation diagnostique

- Les évaluations diagnostiques sont à faire en début de période : début de séquence, début de cycle, etc.
- Elles doivent être ancrées aux programmes.
- Elles doivent être analytiques.

2) Quelle articulation entre évaluation et notation ?

Il y a une opposition entre évaluation et notation mais l'articulation entre les deux devrait être plus réfléchie. On n'apprend pas assez aux professeurs à noter. De plus, il n'y a pas assez de réflexion à l'usage de la notation comme évaluation. Il faut développer l'idée -et il faudrait l'imposer- que les notes que les professeurs ou les inspecteurs mettent doivent être des notes critériées.

3) La relation évaluation – auto-évaluation

L'auto-évaluation est nécessaire mais ne doit pas être la seule forme d'évaluation. L'évaluation doit être extérieure dans tous les cas. Toutefois, avec l'auto-évaluation, on va entrer plus facilement dans la remédiation.

B. L'usage

1) Quel usage dans le débat public ?

Peu de progrès ont été faits dans ce domaine. Notre communication sur le système d'évaluation est trop faible. On s'est laissé embarquer dans le débat « Le niveau augmente-t-il ou baisse-t-il ? » Il faut trouver des formes de transmission intermédiaires (ni trop schématiques ni trop compliquées) afin d'amorcer le dialogue, pour que le système éducatif soit transparent.

2) Un usage interne pour améliorer les résultats

L'articulation entre l'évaluation et les pratiques pédagogiques n'a pas été assez pensée. Les questions de remédiation ne sont pas assez diffusées par les corps d'inspection. Le problème est de savoir quelle part l'évaluation doit tenir dans l'acte éducatif (quelle place, quelle fréquence...?).

Il faudrait développer l'usage de l'évaluation dans les établissements afin d'améliorer l'état des établissements.

IV. Conclusion

Il manque des compétences pour l'évaluation dans le service public en général et dans le système éducatif en particulier. Il faut donc développer de nouvelles compétences et aller vers une professionnalisation.

Il faut davantage former les enseignants et les chefs d'établissement à évaluer et à intégrer l'évaluation dans les pratiques.

Il faut, pour accroître la qualité éducative, utiliser l'évaluation dans la classe, dans l'établissement, dans l'Académie, dans le Ministère. En l'utilisant, en retour, l'outil s'améliorera et la politique éducative s'améliorera. Seule l'évaluation fournira les moyens de s'adapter à la diversification.

**Table ronde avec Jean-Claude ÉMIN¹, Agnès FLORIN² et
Élizabeth BAUTIER³**

Les aspects sociaux et les aspects cognitifs des apprentissages.

I. Intervention d'Élizabeth BAUTIER

Qu'est-ce que l'évaluation pour les élèves du point de vue de sa potentialité différenciatrice ?

Pour les élèves en difficulté, les évaluations nationales (notamment pour la classe de 6^e) ne correspondent pas à ce que disent les enseignants du travail des élèves. De plus, le discours de l'enseignant est parfois en décalage avec les résultats de l'élève ce qui nous fait nous interroger.

L'évaluation est une tâche scolaire comme les autres. Or les élèves en difficulté se trompent sur l'enjeu cognitif de l'évaluation. Ce qui est important, c'est le processus de réponse (dans une évaluation diagnostique). L'objet cognitif sous-tendu par l'évaluation est ce qui est le moins perçu par l'élève.

A. La logique de l'évaluation

Pour l'élève, de la subjectivité se joue dans l'évaluation. En analysant ce qui se joue, on améliore la portée de l'évaluation. Pour que l'évaluation fonctionne, l'élève doit être dans la même logique que l'évaluateur. Les trois logiques sont le savoir-objet, le savoir-outil et le savoir de la discipline. Le temps de passage de l'objet à l'outil est de plus en plus court. Mais certains élèves croient toujours que le savoir reste objet (par exemple, on donne des définitions pour des définitions). Réciter l'objet ne garantit pas la compréhension de la logique de l'outil. Des formes d'évaluation peuvent faire croire aux élèves que ça va (« je sais cocher », « je sais souligner », « je sais réciter », etc.) mais ne permettent pas un réel apprentissage. Il y a un décrochage de l'intérieur.

Dans l'évaluation diagnostique et dans l'auto-évaluation, il faut amener les élèves à s'interroger sur ce qui est la nature de l'objet de savoir car les élèves finissent par penser que l'enjeu c'est la tâche (« il faut que je coche »).

B. Le méta-langage scolaire

Les consignes reposent sur un méta-langage scolaire ce qui remet en cause la non-compréhension des consignes par les élèves. Cela crée une discrimination scolaire. Il y a donc une nécessité de maîtrise des éléments fondamentaux de la langue : qu'est-ce que les élèves comprennent quand je dis cela ?

C. Le statut du document

Toute évaluation renvoie à du traitement de documents, or ceux-ci sont hétérogènes tant par les supports que par ce qu'ils mobilisent chez les élèves. Les élèves les interprètent au minimum car ils pensent que la réponse est dans le document, or elle est dans le traitement du document. On doit donc apprendre aux élèves à traiter le document. Ainsi, quand je choisis un document d'évaluation, je m'interroge sur comment les élèves l'interpréteront.

¹ Sous-directeur de l'Évaluation à la DPD (Direction de la programmation et du développement).

² Professeur à l'université de Nantes (Psychologie du développement et de l'éducation).

³ Professeur à l'université Paris VIII (Sciences de l'Éducation).

D. L'oral

Pour évaluer l'oral, il faut se poser la question suivante : quels sont les moments d'oral qui permettent à l'élève d'apprendre ?

II. Intervention de Jean-Claude ÉMIN

A. Quels objectifs privilégier dans l'évaluation ?

Les objectifs de l'évaluation dans le système éducatif sont :

- d'une part de rendre compte à la société, aux usagers du fonctionnement du dit service public (idée de transparence) ;
- d'autre part de réguler le service public.

Aujourd'hui, dans le système éducatif français, nous atteignons plus le premier objectif que le second.

Les objectifs à privilégier dans l'évaluation sont variables. Il faut préciser à chaque fois ce que l'on attend : évaluer des progrès, évaluer un niveau, évaluer des compétences, etc.

B. Les différentes évaluations

La distinction entre évaluation bilan et évaluation diagnostique est indispensable. L'évaluation diagnostique est un outil destiné aux enseignants en début de cycle pour leur permettre de connaître le niveau des élèves en terme de savoirs, de compétences, etc. et ainsi de préparer leur progression. Elle n'est pas un ensemble des objectifs qui doivent être atteints à un moment donné ; ceci est réservé à la fin de cycle et donc à l'évaluation bilan. Toutefois, on peut avoir les mêmes épreuves pour les deux types d'évaluation mais il y a une différence fondamentale de traitement des résultats :

- si c'est une évaluation diagnostique, plusieurs items sont présents et vont permettre de progresser ;
- si c'est une évaluation bilan, le traitement est binaire (c'est juste ou c'est faux).

C. Que faut-il faire pour que l'évaluation serve ?

Lorsqu'on parle de « remédiation », on a tendance à vouloir dire aux enseignants ce qu'il aurait fallu faire. Il faudrait dire : « Si je constate telle lacune, telle difficulté à tel moment, qu'est-ce que je fais ? »

III. Intervention de Agnès FLORIN

L'évaluation n'a aucun intérêt si elle ne sert pas à améliorer. Elle est une photographie à un moment donné. Ce qui est observé aujourd'hui ne sera pas vrai demain. De plus, il faut repérer les compétences et les savoir-faire plutôt que les difficultés.

Les troubles des apprentissages apparaissent à l'école élémentaire lors des apprentissages fondamentaux (lire, écrire, compter). L'école maternelle doit jouer un rôle préventif. Mais se posent alors des problèmes : quelles mesures utiliser ? comment passer d'une prédiction statistique à une prédiction personnelle ?

Les évaluations des enseignants sont souvent globale et ont un pouvoir de prédiction important (en maternelle pour l'école élémentaire). On est passé d'une procédure qui permettait de classer les élèves et de les orienter à une procédure d'apprentissage de savoirs et de savoir-faire. Mais finalement, évalue-t-on des compétences scolaires ou sociales ? N'apprenons pas à faire demain ce qui n'existe pas aujourd'hui ?

Quel est le sens de l'évaluation pour l'élève ? A cinq ans, l'enfant sait déjà comment ça fonctionne et sait se situer dans la classe par rapport aux autres.

Intervention de Luce DOSSAT ¹ **Pourquoi des évaluations diagnostiques ?**

Le repérage des compétences des élèves est important pour mettre en place des itinéraires d'apprentissage pour toute l'année. Les élèves n'arrivent pas avec les mêmes connaissances, les mêmes compétences, les mêmes rythmes, or on doit les mener ensemble. On ne peut construire qu'à partir de ce qu'ils savent. Il faut gérer l'hétérogénéité alors que la norme qu'il faudra pour tous est la même.

On peut distinguer trois temps :

- le temps de l'institution (cycles, périodes...);
- le temps de l'enseignement (professeur);
- le temps de l'apprentissage (élève).

I. L'évaluation diagnostique, pour quoi faire ?

Une évaluation diagnostique n'est pas un examen, un contrôle, un outil de classement des élèves, des établissements. Ce n'est pas un outil de comparaison dans le temps. C'est un outil pour le maître qui permet de repérer des acquis, des difficultés, des représentations, etc. À partir de ces repères, avec d'autres outils d'observation, le maître peut analyser les démarches, les procédures, les représentations des élèves, pourquoi ils les utilisent, comment ils se les approprient. Ainsi il va construire des situations d'apprentissage et de ré-apprentissage différenciées. Il va mettre en place cette « remédiation », de nouvelles situations d'apprentissage qui vont amener de nouvelles évaluations.

Mais se pose le problème de la mesure : comment mesurer l'écart entre le départ et le point atteint ?

II. Les hypothèses d'apprentissage

Ce problème de l'évaluation dépend de la représentation que l'on se fait de l'apprentissage, des hypothèses d'apprentissage.

Il est important que l'élève participe de manière active à son apprentissage et qu'il lui donne du sens. On va d'abord construire des outils (plus ou moins implicites) qui, après avoir été utilisés, vont être appropriés par les élèves et donc être institutionnalisés, décontextualisés. Autrement dit, les élèves vont s'approprier les variables didactiques. Il est donc important de repérer les « parasites » dans les représentations des élèves (par exemple, souvent, pour les élèves, un rectangle n'est « bon » que s'il est « couché » et avec certaines proportions).

Les connaissances ne s'élaborent pas de façon linéaire : il y a des ruptures et des reconstructions nécessaires, indispensables (Bachelard parle de construction face à un obstacle).

Apprendre, c'est avoir reconnu ses erreurs, des conceptions erronées. « Apprendre, c'est autant perdre les idées qu'on avait qu'en apprendre de nouvelles » (Hameline).

¹ IEN de Mathématiques – IUFM de Clermont.

ATELIER 1 :

Qu'est-ce que construire un outil d'évaluation diagnostique ?

Animation/intervention : Michel BRAXMEYER¹, Philippe GUIMARD²

Thèmes : Les étapes de la construction des outils, les difficultés rencontrées, les limites ; ce qui différencie un outil de diagnostic d'un exercice de type contrôle ou bilan voire d'un test psychologique dans les aspects méthodologiques et théoriques.

Compte-rendu de l'atelier :

L'atelier s'est déroulé en deux temps :

- le matin a été consacré à une discussion générale sur la place de l'évaluation diagnostique dans les pratiques d'évaluations ;
- l'après-midi visait à proposer une réflexion sur les problèmes de construction de l'outil diagnostique.

I. La place de l'évaluation diagnostique

On a cherché à définir l'évaluation diagnostique (ses finalités, ses champs d'application, ses formes) par rapport aux autres formes d'évaluation qu'il s'agisse des évaluations « bilans » ou des évaluations visant la certification (baccalauréat, par exemple). À cet effet, les différents dispositifs d'évaluation actuellement mis en œuvre (ou en cours d'élaboration) par le ministère de l'Éducation nationale ont été présentés. Parallèlement, la place de l'évaluation diagnostique en psychologie par rapport à l'évaluation réalisée par l'enseignant a été précisée. On a ainsi souligné que l'évaluation diagnostique réalisée par les enseignants visait la régulation des apprentissages de tous les élèves à partir de la prise en compte des difficultés des élèves dans leurs apprentissages. Ce point a permis d'établir une distinction importante entre « difficultés des élèves » (qui apparaissent dans tout processus d'apprentissage) et « élèves en difficulté » (élèves dont les difficultés sont durables dans le temps).

La question du caractère « abstrait » (décontextualisé ?) des évaluations diagnostiques mises à disposition des enseignants a été soulevée. Pour certains enseignants, les évaluations diagnostiques doivent être intimement liées aux situations d'apprentissage réalisées en classe. Il a été alors rappelé qu'un des objectifs poursuivi par les concepteurs des évaluations diagnostiques est précisément que les enseignants s'approprient cette démarche pour qu'ils puissent par eux-mêmes construire des outils d'évaluation.

Les stagiaires ont souligné la difficulté à construire et utiliser ce type d'évaluation, notamment lorsqu'il s'agit d'évaluer les compétences transversales. Dans ce cadre, mais également de façon plus générale, la mise en place d'évaluations diagnostiques nécessite un véritable travail d'équipe : pour harmoniser les pratiques, progresser dans la maîtrise des instruments eux-mêmes, pour interpréter les résultats et mettre en place des actions de remédiation concertées. Ce travail d'équipe peut conduire à développer une véritable « culture » de l'évaluation diagnostique et faire infléchir le recours assez systématique, à certains niveaux d'enseignement (lycée notamment) aux évaluations sommatives.

Nécessité également d'un accompagnement des équipes elles-mêmes pour les aider à se former à l'utilisation des outils diagnostiques et à adapter ces outils dont la présentation est

¹ Professeur d'Histoire-Géographie – Arpajon (91).

² Maître de conférence – Labécé – Faculté des lettres et sciences humaines – Nantes.

parfois trop complexe. Des expériences allant en ce sens ont été rapportées à propos des évaluations GS-CP³.

II. Les problèmes de construction de l'outil diagnostique

La construction de l'outil diagnostique suppose que l'on réponde précisément aux trois questions suivantes : que souhaite-t-on évaluer ? comment peut-on évaluer/mesurer ? comment interpréter les résultats ?

Concernant le premier point, il a été précisé que l'évaluation diagnostique visait l'évaluation des compétences des élèves, en référence, notamment, aux programmes. Une discussion animée s'est alors engagée autour de la définition des notions de compétences/connaissances/capacités.

S'agissant du second et troisième points, divers exemples ont été rapportés par les animateurs et les stagiaires concernant l'importance de la compréhension des consignes (à tous les niveaux scolaires), d'une part, et l'importance de la nature de la tâche proposée aux élèves pour évaluer une compétence, d'autre part. Il a été souligné qu'une compétence donnée pouvait être évaluée de différentes manières. Le recours à une diversité de modalités de l'évaluation (ou tâches) présente l'intérêt de mieux cerner la nature des difficultés de l'élève dès lors que l'évaluateur est en mesure de distinguer la nature des processus et des connaissances que l'élève doit mobiliser dans la réalisation de ces tâches.

On a également souligné la difficulté mais aussi l'intérêt de l'analyse des erreurs. Une présentation d'une typologie des erreurs issue de J.-P. Astolfi (*L'erreur, un outil pour enseigner*, Paris ESF) a été proposée et discutée.

III. Conclusion

Plusieurs stagiaires ont manifesté leur satisfaction à l'égard de la réflexion qui a été menée durant cette journée. Les questionnements convergents ont été jugés finalement « rassurants » et constructifs.

³ Grande section – CP.

ATELIER 2 :
Interpréter les résultats des évaluations, l'exemple des protocoles CE2/6^e
Animation/intervention : Gérard BREZILLON ¹, Brigitte REIBMESTER ²

Thèmes : Les interprétations des résultats des élèves dans l'acte pédagogique sur le plan organisationnel, thématique ou didactique ; au niveau de la classe, de l'interdisciplinarité, entre établissement (inter école ou premier degré, second degré).

Compte-rendu de l'atelier :

I. Points abordés au cours de l'atelier

1. Présentation des principes sur lesquels sont fondés la fabrication des protocoles et les tableaux présentant les champs, les capacités, les compétences, les composantes.
2. Pistes d'exploitation.
3. Présentation rapide de la banque d'outils.
4. Travail pratique : interpréter les résultats de différents points de vue à partir de l'exemple d'un collègue.
5. Présentation du système CASIMIR.

II. Trois points de vue

Du point de vue de l'interdisciplinarité :

- analyse des différences de réussite par items (s'intéresser aux trois ou quatre élèves en difficulté sur des items majoritairement réussis) ;
- repérer à partir des items des « compétences-points forts » et les « compétences-points faibles » dans le groupe classe.

Du point de vue du chef d'établissement :

- dégager deux axes de travail : établissement et secteur (cycle 3 et 6^{ème}) ;
- mettre en place un groupe de réflexion sur les faiblesses constatées (inter-cycles).

Du point de vue de la classe :

- repérer les élèves les plus en difficulté ;
- mettre en place des groupes de besoin brassant les élèves de toutes les classes ;
- travailler sur les écarts de pourcentage entre les résultats de la classe et les résultats nationaux.

¹ Professeur des écoles – Aubervilliers (93).

² Professeur – Paris (75).

ATELIER 3 : **La notion de compétences**

Animation/intervention : Isabelle NOCUS¹, Catherine BERDONNEAU²

Thèmes : Les compétences transversales et disciplinaires, scolaires ou sociales. Est-ce que l'évaluation des compétences s'oppose à l'évaluation des savoirs ?

Compte-rendu de l'atelier :

I. Les participants

Les attentes exprimées par les participants, de fonctions et d'origines diverses, s'organisent autour de plusieurs thèmes complémentaires :

- partager les savoirs et les savoirs-faire sur l'évaluation ;
- repérer précisément les différences d'objectifs (et de liens objectifs/apprentissage), de modalités de conception et d'exploitation entre les différents types d'évaluation ;
- apprendre à mieux utiliser les outils à disposition pour chaque type d'élèves ;
- construire des leviers pour modifier les représentations mutuelles élèves/maîtres/école ;
- réaliser un système d'évaluation local optimisant les échanges entre partenaires, entre disciplines ;
- collecter des pistes pour former à l'évaluation et à son utilisation ;

avec des objectifs d'application directe dans la classe, d'utilisation pour animer des stages de formation initiale et/ou continue et, aussi, nourrir la réflexion de groupes de recherche.

II. Les moments d'échange

Dans le but de mettre à plat et confronter les représentations de chaque participant sur l'évaluation, ont été proposés les thèmes de discussion suivants :

1. Chercher une définition du mot « évaluation ».
2. Faire des associations d'idée avec le mot « compétence ».
3. Réfléchir concrètement à la construction d'un outil, en rapport avec l'évaluation, attaché à un objectif spécifique choisi par le groupe.

III. Les idées centrales exprimées

« Évaluer » sous-tend la notion de valeur ; l'objectif d'une évaluation doit donc toujours être positif.

Le vocabulaire utilisé n'est actuellement pas unifié ; la convergence terminologique ne semble pas prioritaire pour l'action, ce qui est important est de bien définir, au préalable, l'objectif que l'on donne à l'évaluation à construire, puis, de s'assurer, avant l'exploitation des résultats, que l'évaluation réalisée évalue effectivement ce que l'on souhaitait évaluer.

Une évaluation est une prise d'information, à un moment donné, dans un but d'échange et/ou de décision.

Toute compétence ne peut être évaluée qu'à partir de performances.

¹ Maître de conférence – Labéced – Faculté des lettres et sciences humaines – Nantes.

² Professeur de Mathématiques – IUFM de Versailles.

ATELIER 4 :
L'accompagnement à l'évaluation et ses usages
Animation/intervention : Jacqueline MASSONNET ¹, Josette LE COQ ²

Thèmes : Quel accompagnement pour que l'enseignant soit utilisateur des évaluations ? Qui accompagne ? au niveau de la discipline, de la classe, de plusieurs classes, d'un établissement, d'un secteur ou d'un bassin de formation ?

Compte-rendu de l'atelier :

I. La nouvelle évaluation diagnostique GS/CP ³

A. Présentation

Le but de cette évaluation est de réussir les apprentissages du CP. Elle est une aide aux maîtres pour pointer les besoins des élèves pour qu'ils réussissent les apprentissages fondamentaux du CP.

Elle comporte une quarantaine de situations qui permettent d'évaluer des compétences, dont une en particulier, indispensable pour réussir l'apprentissage de la lecture : la représentation du langage dynamique, c'est-à-dire du langage du récit, avec un début, une transformation et une situation finale.

Il est nécessaire d'informer enfants, parents, collègues sur cette évaluation, pour rassurer, pour rendre les parents acteurs de la formation de leur enfant.

B. Débat

Quelles garanties a-t-on pour éviter que ces évaluations diagnostiques ne deviennent des contrôles de connaissances ? Cette évaluation diagnostique n'implique pas de « retour » en termes chiffrés, au contraire des évaluations CE2/6^e. Il n'y a pas d'esprit de contrôle. Dans cette optique, les maîtres ne doivent pas penser que c'est une mesure de leur travail.

Il ne faut pas confondre évaluation diagnostique et évaluation bilan.

Quel rapport l'enseignant a-t-il avec l'évaluation (ses propres représentations à ce sujet) ? Il a un besoin d'accompagnement : les outils des enseignants ne sont pas toujours suffisants. Il faut développer une culture de l'évaluation dans laquelle ces outils ont leur place, dans un ensemble plus vaste.

L'évaluation diagnostique ne peut pas servir à tout : elle est un outil qui participe à la réflexion sur l'efficacité de l'enseignement. Il est nécessaire que le ministère s'interroge sur l'efficacité de son système.

Un chef d'établissement doit pouvoir parler de critères d'évaluation avec ses professeurs.

¹ IEN 1^{er} degré – Conceptrice de l'évaluation diagnostique Grande section/CP.

² IA-IPR de Mathématiques – Versailles.

³ Cf. BO n°31 du 30 août 2001.

II. Dispositifs d'accompagnement à l'usage des évaluations diagnostiques

Travail de groupes dont les objectifs étaient les suivants : qui accompagne qui ? avec quel(s) objectif(s) ? avec quel(s) contenu(s) ? selon quelle(s) modalité(s) ?

A. Groupe 1 : L'évaluation diagnostique CM2-6^e

1) Qui accompagne qui ?

- échanges d'enseignants 1^e et 2^e degré ;
- apports extérieurs par des personnes-ressources (enseignants de l'établissement, IEN, IA-IPR, conseiller pédagogique, chef d'établissement, formateurs IUFM, coordonnateur REP) :
 - en formation continue par circonscription/ par secteur de collège, par stages théoriques, par animation pédagogique (à l'initiative de l'IEN) ;
 - des réunions institutionnelles écoles-collège (suite aux résultats de l'évaluation) à but de remédiation ;
- éventuellement, accompagner des demandes spontanées d'enseignants.

2) Quels objectifs ?

- aider les élèves en difficulté ;
- améliorer la réussite scolaire ;
- assurer la continuité des apprentissages ;
- accompagner les enseignants pour qu'ils aient un recul réflexif sur leur pratique ;
- individualiser les parcours dans une logique d'appartenance au groupe.

3) Quel(s) contenu(s) ?

Il faut des contenus de moins en moins disciplinaires et de plus en plus interdisciplinaires et transversaux.

4) Quelle(s) modalité(s) ?

- tutorat des élèves 6^e/CM2 ;
- équipe mixte 1^e et 2^e degré pour la correction des tests, les projets pédagogiques, l'interaction pour le suivi des élèves en difficulté ;
- stages avec des échanges de pratiques, des histoires de vie, des thématiques transversales.

B. Groupe 2 : Exploitation des évaluations de 6^e

1) Qui ?

Une personne d'un groupe ressource académique, un conseiller pédagogique, un coordinateur du collège identifié (professeur ou chef d'établissement ou son adjoint) s'adresse à une équipe constituée de professeurs d'un collège et des professeurs des écoles du secteur.

2) Objectifs

- Faire évoluer les exploitations de l'évaluation grâce à des heures d'aide et d'accompagnement et à des cours.
- Convaincre.
- Faire des apports théoriques et pratiques.
- Mener une recherche-action.

3) Modalités

Stage sur site à la demande de l'établissement :

- émergence des besoins ;
- apports ;
- travail du groupe sur l'année et bilan.

C. Groupe 3 : Problématique d'un dispositif d'accompagnement

Problématique en trois axes :

- Explicatif :
 - mettre en lumière la polysémie des termes ;
 - expliciter les théories en jeu.
- Incitatif :
 - nécessité d'une articulation intime des politiques qu'elles soient ministérielles, académiques, locales, d'établissement ou de pratique de classe ;
 - utiliser des leviers ou des médiateurs divers tels que l'interdisciplinarité ou le transversal.
- Proche du terrain :
 - prendre en compte les représentations des collègues ;
 - prendre en compte les pratiques d'enseignement.

Conférence par Jacques JUHEL ¹

Faut-il –est-il utile de- mesurer pour évaluer les élèves ?

Objectifs de la présentation :

1. Rappeler la diversité des conceptions de l'évaluation.
2. Aider à comprendre que mesurer c'est aussi et avant tout expliciter.
3. Argumenter en faveur d'une utilisation raisonnée (comprise, raisonnée et raisonnable) de la mesure dans l'évaluation des élèves.

I. L'évaluation des élèves

A. La variété des acteurs

Dans l'évaluation des élèves, les acteurs sont nombreux : institution, professeurs, psychologues scolaires... mais avant tout l'élève lui-même.

B. La diversité des fonctions

AVANT L'ACTION DE FORMATION	PENDANT L'ACTION	APRES L'ACTION DE FORMATION
Évaluation : <ul style="list-style-type: none"> • Diagnostique • Pronostique • Prédicative 	Évaluation : <ul style="list-style-type: none"> • Formative • « Progressive » 	Évaluation : <ul style="list-style-type: none"> • Sommative • Terminale
Fonction : <ul style="list-style-type: none"> • Orienter • Adapter 	Fonction : <ul style="list-style-type: none"> • Réguler • Faciliter (l'apprentissage) 	Fonction : <ul style="list-style-type: none"> • Vérifier • Certifier
Centrée sur : <ul style="list-style-type: none"> • Le producteur et ses caractéristiques (Identification) 	Centrée sur : <ul style="list-style-type: none"> • Les processus • L'activité <div style="display: inline-block; vertical-align: middle; margin-left: 10px;">} de production</div>	Centrée sur : <ul style="list-style-type: none"> • Les produits

Fonctions de l'évaluation selon sa place par rapport à la séquence de formation ²

Les fonctions sont centrées sur :

- les compétences et connaissances des élèves ;
- les modalités d'acquisition des connaissances et des compétences.

C. La multiplicité des approches et des pratiques

On peut distinguer trois logiques.

- La logique de bilan :
 - l'évaluation normative : par rapport à un niveau inter-individuels ;
 - l'évaluation critériée : par rapport à des objectifs.
- La logique d'accompagnement :
 - l'évaluation maîtrise par objectifs ;
 - recherche des structures de l'évaluation, analyse des pratiques d'évaluation ;
 - l'évaluation formative, l'évaluation-régulation, l'évaluation pour la remédiation.
- La logique d'interrogation du sens.

¹ Professeur à l'université de Rennes (Psychologie différentielle).

² Charles Hadji, L'évaluation, règles du jeu. Des intentions à l'action, ESF, Paris, 1989, p.59.

II. Qu'est-ce qu'évaluer ?

A. Une opération de mise en correspondance

L'objet à évaluer n'existe pas en lui-même ; c'est un « construit ». C'est une construction dépendante du modèle de référence de l'évaluateur en fonction des finalités liées à des valeurs, de la logique sous-tendant l'activité de l'évaluation, d'objectifs énonçant les intentions de résultats attendus.

La correspondance entre ce qu'on pense qui devrait être et ce qu'on observe permet d'objectiver un écart.

B. C'est être juste, objectif

La morale de l'équité doit primer lorsque l'évaluation a des fonctions pronostiques, lorsqu'elle certificative, lorsqu'elle est normative.

Dans toute évaluation il y a un souci d'objectivité : « J'essaie de décharger mon discours d'une certaine subjectivité. » (Cf. Bachelard) L'évaluation sommative est donc instrumentée, comparative (aux plans inter et intra-individuel), son référent est « empreint d'extériorité » et elle porte le plus souvent sur des produits.

C. C'est aider, faciliter

L'évaluation est un mode de régulation, de guidage, d'optimisation des apprentissages. C'est une logique pragmatique qui part « de la pratique pour l'apprécier et la clarifier », qui s'appuie sur une représentation réaliste des apprentissages.

Une forme prototypique d'évaluation, l'évaluation à visée formative, met en évidence l'opportunité d'une remédiation, d'une régulation, est inscrite dans une dynamique de changement et se focalise sur les procédures.

III. Quelle place pour la mesure de l'activité d'évaluation ?

A. Qu'est-ce que mesurer ?

Mesurer, c'est assigner des nombres à des objets, selon une règle logiquement acceptable, de façon à représenter leurs propriétés et relations. Mesurer, ce n'est pas seulement quantifier, c'est aussi ordonner, classer.

On ne peut pas répondre à la question « Qu'est-ce que je mesure ? » car tout dépend de l'interprétation.

B. Le modèle métrique de l'évaluation

Le modèle métrique de l'évaluation repose sur l'estimation rigoureuse et objective (modèles de la docimologie, de la psychométrie, de l'édu-métrie).

Les bilans sont intéressants car ils permettent une meilleure connaissance du système éducatif (et peuvent permettre de fixer une norme...)

C. De l'évaluation à l'observation formative

Pour Perrenoud, au lieu « d'évaluation », il vaudrait mieux parler « d'observation formative » car « Évaluer n'est pas mesurer » (Vial ?).

Comprendre, utiliser différemment la mesure pour mieux évaluer ?

Conférence par Alejandro TIANA ¹

Approche de l'évaluation des élèves au niveau international : ce qu'apportent les enquêtes internationales sur les acquis des élèves ; le regard d'un pays étranger sur un système éducatif ; la France : quelle particularité ?

I. Quelques idées sur l'approche de la comparaison internationale des acquis des élèves

Ces derniers mois, on a beaucoup parlé de PISA. ² Les premières données ont été collectées en 2000 et commencent à paraître maintenant. Dans beaucoup de pays, cela a provoqué un débat (notamment en Allemagne). C'est intéressant car ce type d'étude permet d'avoir un autre regard sur son propre système même si ce regard ne va pas au fond du système mais observe seulement certains aspects. Qu'est-ce que peut, ou ne peut pas, nous apporter cette approche ?

Ce projet PISA, même s'il est récent, repose sur le travail de chercheurs impliqués depuis longtemps dans l'évaluation. En 1958, ils se sont réunis à Hambourg (UNESCO) afin de comparer l'évaluation des acquis des élèves et ont créé l'IEA. On a focalisé sur les acquis des élèves, ce qui a été une nouveauté. Ces préoccupations répondaient au désir de trouver des explications pour améliorer les systèmes éducatifs. Le monde devenait ainsi un « laboratoire éducatif » ; on utilise l'écart des caractéristiques des systèmes éducatifs et de leurs résultats afin de comparer, de repérer les points forts et les points faibles, dans le but d'améliorer le système éducatif. Toutefois, l'idée du laboratoire éducatif n'est pas directement transférable à la partie opérante du système.

Le but est d'assurer la comparaison entre tous les systèmes mais cela pose de problèmes. D'abord les objectifs de chaque pays sont partiellement identiques et partiellement différents. En général, les objectifs généraux sont assez proches, au contraire des objectifs particuliers qui sont beaucoup plus divergents. Ainsi sont recherchés des critères communs permettant une comparaison internationale, ce qui amène à se poser les questions : quel *curriculum* est prescrit par l'État ? appliqué en classe par les professeurs ? assimilé par les élèves ? Ensuite, comment peut-on mesurer avec des apports culturels, religieux, etc. différentes ? comment développer des études sans « billet » culturel ? comment assurer la compatibilité d'un instrument traduit dans des langues différentes (par exemple, dans certaines langues, un concept n'existe pas toujours exprimé par un seul mot) ? On doit être conscient des difficultés et des limites d'une comparaison internationale, mais elle permet de porter un regard différent sur son propre système éducatif. Parmi les critères retenus, on relèvera notamment les matières centrales des cursus (langue maternelle, mathématiques, sciences), l'éducation civique, les TICE, les classes maternelles/pré-scolaires...

Il faut noter qu'il existe plusieurs types d'études internationales, mais elles reposent sur des critères différents. Le projet PISA de l'IEA a été construit comme un système séparé de celui de l'OCDE et de plus en plus de pays non-OCDE y participent. On doit distinguer les études internationales qui mesurent le rendement et celles qui mesurent le niveau de formation d'un système éducatif. Les études de l'IEA, comme beaucoup d'études nationales, veulent mesurer

¹ Directeur de l'IEA (International Association for the Evaluation of Educational Achievement), Vice-recteur de l'université de Madrid et Membre du Haut Conseil de l'Évaluation (France).

² PISA : Program for International Student Assessment.

le rendement d'un système scolaire à un degré particulier ; en d'autres mots, elles veulent vérifier si les élèves ont atteint le niveau prescrit. La problématique de PISA est : « Est-ce que la population jeune ³ a développé des capacités à s'adapter aux contraintes du monde ? » Les deux informations sont valables mais différentes. Il faut donc être conscient de ce qu'on mesure afin de tirer des conclusions.

II. Le regard d'un étranger sur le système d'évaluation français

La France a choisi un système d'évaluation lié à l'État et non au marché. Il y a eu des pays, comme l'Angleterre, où l'on a poussé la compétition et où l'on a donc fait intervenir le marché.

En France, il y a eu une volonté d'évaluer. Ce n'est pas une volonté agressive mais une volonté contenue. La France a construit très tôt un système d'évaluation et l'a maintenu dans le temps. Malgré les changements politiques, il y a une continuité de l'effort fait en faveur de l'évaluation et une continuité de fonctionnement. Il y eu une volonté de connexion de l'évaluation du système avec la vie scolaire, avec le terrain. On a essayé de combiner des méthodes différentes (au contraire de la plupart des pays anglo-saxons où l'approche quantitative a été privilégié).

En France, il n'y a pas seulement une volonté de développement de l'évaluation, il y a aussi une réflexion autour de l'évaluation à travers le Haut Conseil de l'Évaluation.

La France offre par ses démarches des possibilités de réflexion pour les pays qui mettent en place d'autres systèmes d'évaluation.

³ Ce système se base sur des jeunes de quinze ans.

**Table ronde avec Jean-Claude ÉMIN (France)¹,
Manuel PALACIOS (Brésil)², Tatiana FORENKO (Russie)³,
Szuzsanna HORVATH (Hongrie)⁴, Anna BUTASOVA (Slovaquie)⁵,
Jean-Richard CYTERMANN (France)⁶**

**Les systèmes nationaux d'évaluation en France et chez nos partenaires du
Brésil, de la Hongrie, de la Slovaquie et de la Russie**

I. Le système d'évaluation brésilien

Au Brésil, le système éducatif est un système fédéral composé de vingt-sept unités. L'évaluation de l'éducation est faite au niveau fédéral, par le gouvernement fédéral.

*Saeb*⁷ a été mis en place en 1991. C'est une évaluation réalisée à la fin du lycée, du collège et de l'école. Les huit premières années ont été évalués les résultats des écoles des municipalités et des écoles de l'État.

En 1998 a été créé l'*Enem*⁸ qui est déterminant pour l'accès à l'enseignement supérieur. Ce n'est pas vraiment l'équivalent du bac car ce n'est pas un diplôme mais une indication.

En 1996 a été créé l'*Exale Nacional de Cursos*. C'est un examen obligatoire pour avoir accès aux diplômes de l'université. C'est un outil de surveillance du système d'enseignement supérieur pour le gouvernement.

L'État de Minas est représentatif de l'hétérogénéité du Brésil. C'est pourquoi il dispose d'un système spécifique d'évaluation visant l'accompagnement des performances des établissements scolaires basiques. Les résultats sont publiés et certaines informations sont envoyées directement aux professeurs et aux écoles. Toutefois, cette évaluation n'est pas assimilable à une compétition.

II. Le système d'évaluation hongrois

Le système éducatif est hérité du système socialiste d'État et cet héritage est lourd. En effet, c'est un système très sélectif marqué par de nombreuses évaluations bilans et évaluations diagnostiques. De très grosses différences apparaissent entre les villes, entre les écoles, etc. Dans une même école, il y a une forte hétérogénéité.

Une évaluation bilan nationale, MONITOR, est menée depuis 1986 autour de la lecture, des mathématiques, des sciences naturelles et de l'informatique. Depuis 1970, la Hongrie participe à de nombreuses évaluations internationales : IEA, OCDE SIALS, PISA...

Trois acteurs évaluent : le ministère, les communautés locales (qui peuvent acheter des services d'instituts privés...) et le corps des enseignants lui-même (l'école définit son programme local et établit sa propre évaluation).

¹ Sous-directeur de l'Évaluation à la DPD (Direction de la programmation et du développement).

² Directeur du CAEd (Centro de Políticas Públicas e Avaliação da Educação) (• Laboratoire de la mesure de l'évaluation).

³ Chercheur à l'université de Moscou.

⁴ Directrice-adjointe de *Kiss Árpád* (• Institut pédagogique national) à Budapest.

⁵ Directrice du département de français à la faculté de pédagogie de l'université de Comenius (Bratislava).

⁶ Directeur de la DPD (Direction de la programmation et du développement).

⁷ Saeb : Sistema Nacional de Avaliação da Educação Básica (• Système national d'évaluation d'éducation basique).

⁸ Enem : Exame Nacional do Ensino Médio (• Examen national d'enseignement moyen).

baccalauréat est conçu, réalisé et corrigé par l'établissement dans lequel sont scolarisés les élèves.

Un projet de coopération avec le ministère de l'Éducation nationale français est en cours. Ainsi, en 2000, un sondage a été réalisé : il s'agissait d'une évaluation diagnostique autour de la langue slovaque et des mathématiques au début de la 5^e année de scolarisation (classe charnière car c'est la première classe du secondaire). Cette évaluation n'a pas été corrigée par les professeurs mais par les chercheurs eux-mêmes car ils n'avaient pas eu les moyens de former les professeurs. En 2002, le nombre d'écoles (et donc d'élèves) sondées a augmenté et les exercices ont été modifiés ; l'inspection générale et l'inspection académique ont été intégrées au projet.

En ce qui concerne l'évaluation bilan, la réforme du baccalauréat est en cours sur le modèle de la Hongrie depuis 1999.

V. Le système d'évaluation français

En France, l'évaluation est prévue par la loi (1989). Il y a une direction statistique, la DPD⁹, qui étudie les résultats.

Il y a une multiplicité des domaines d'activités et des outils :

1. Les évaluations globales qui permettent de comparer les académies et leur évolution.¹⁰
2. L'évaluation des performances des élèves est une réponse aux sollicitations des décideurs et de l'opinion publique et est une aide à l'enseignant, à l'inspecteur, au recteur d'académie.
3. L'évaluation des pratiques pédagogiques et des dispositifs innovants.

Il y a toutefois de lacunes car on ne sait pas trop ce que savent les étudiants dans telle ou telle matière, parce qu'il y a une absence (ou une insuffisance) d'utilisation des évaluations pour concevoir les programmes ou modifier les examens, parce que les évaluations sont difficiles à utiliser sur le terrain (d'où l'abandon des évaluations diagnostiques en seconde).

Pour la France, les enjeux des comparaisons internationales sont de pouvoir se comparer aux autres pour améliorer le système éducatif français et de ne pas laisser le monopole des instruments d'évaluation aux Anglo-Saxons qui sont eux dans la logique de marché.

⁹ DPD : Direction de la programmation et du développement.

¹⁰ Les résultats sont publiés dans *L'état de l'école*.

Présentation

Les évaluations internationales et leur utilisation :

PIRLS par Marc COLMANT ;

PISA par Isabelle ROBIN et Thierry ROCHER

I. PIRLS : Progress in International Reading Literacy Study

PIRLS est une étude pilotée par l'IEA¹ qui porte sur la quatrième année de scolarité obligatoire. Trente-huit pays y participent. En France, le projet est coordonné par la DPD. Il a commencé en 1999. Les résultats seront rendus publics le 8 avril 2003. Cette étude porte uniquement sur la lecture. Il s'agit de comparer les compétences et les habiletés en lecture dans l'espace (entre pays) et dans le temps (sur plusieurs années).

Les aspects évalués sont :

- les processus de compréhension :
 - extraire des informations explicites d'un texte ;
 - effectuer des inférences simples ;
 - interpréter et intégrer des idées et des informations ;
 - examiner le contenu, la langue et les éléments textuels ;
- les buts de lecture :
 - textes narratifs ;
 - textes informatifs ;
- les comportements et les attitudes face à la lecture.

Les outils sont dix cahiers tournants d'épreuves de lecture et quatre questionnaires (un pour l'élève, un pour ses parents, un pour son enseignant, un pour la direction de son école).

II. PISA² : Premiers résultats de l'évaluation internationale des acquis des élèves de quinze ans

A. Présentation de Thierry ROCHER

La première enquête internationale des acquis a été pilotée par l'OCDE. Elle a bénéficié des acquis des enquêtes précédentes (notamment celles de l'IEA). Elle rencontre d'ailleurs les mêmes difficultés.

Le dispositif se présente en trois phases de trois ans chacune. Chaque phase met l'accent sur un domaine : la compréhension de l'écrit en 2000, la culture mathématique en 2003 et la culture scientifique en 2006.

C'est une enquête qui s'adresse à un âge (quinze ans révolus) et non à un niveau. Elle s'est attachée à des compétences plus qu'à des connaissances. Par ailleurs, PISA recueille des informations « contextuelles » auprès des chefs d'établissements notamment.

On peut noter une forte variabilité des résultats.

B. Présentation d'Isabelle ROBIN

Les compétences évaluant la compréhension de l'écrit étaient « s'informer », « interpréter » et « réagir ». C'est dans cette dernière compétence que les Français se sont révélés les plus faibles.

¹ IEA : International Association for the Evaluation of Educational Achievement.

² PISA : Program for International Student Assessment.

C. Présentation de Szabo TAMAS³

Les résultats hongrois ont été inférieurs à la moyenne de l'OCDE. Pourquoi alors que pendant les trente années précédentes les résultats aux autres enquêtes (notamment celles de l'IEA) étaient satisfaisantes ?

C'est peut-être à cause de « l'écroulement des mythes », des épreuves inhabituelles pour les Hongrois, parce que les contenus sont privilégiés dans l'enseignement en Hongrie, en raison de la politique des établissements visant à faciliter le passage des élèves au niveau suivant ou encore de l'apathie du corps enseignant à la suite des changements de grande envergure après la « tournure » (1989-1990).

À la suite de ces résultats, des réactions de la politique hongroise apparurent. D'abord, les résultats précédents des évaluations bilans (nationales et internationales) ont été prises en considération et complétés par l'introduction du système de l'évaluation diagnostique. Ensuite, il y a une consultation ministérielle auprès d'experts. Les programmes scolaires en vigueur ont été analysés du point de vue du développement des compétences. La compréhension de texte a été incluse dans toutes les épreuves du baccalauréat. Enfin, le développement des compétences a été renforcé à tous les niveaux du système.

³ Intervenant hongrois.

Conférence par André FLIELLER ¹ **Évaluation, comparaison, comparabilité**

Le terme *évaluer* a deux sens : c'est d'une part mesurer donc comparer, c'est d'autre part porter un jugement de valeur donc comparer une production à la norme. Donc toute évaluation implique une comparaison.

S'interroger sur la comparabilité amène à se poser quelques questions : quels problèmes ? quelles parades méthodologiques ? quelles limites ?

I. Les évaluations collectives

A. Types de comparaison

1) Comparaisons internationales

Cet aspect a été développé pendant les interventions précédentes.

2) Comparaisons de cohortes

Les comparaisons de cohortes sont des comparaisons dans le temps. Elles ne servent pas seulement à comparer des niveaux pour vérifier si le niveau monte ou si le niveau baisse.

3) Comparaisons d'unités du système éducatif (établissements, académies...)

4) Comparaisons de groupes sociaux ou socio-démographiques

5) Comparaisons d'organisations pédagogiques (programmes, méthodes, filières...)

B. Stratégies méthodologiques pour assurer la comparabilité

1) Méthodes expérimentales

Ces méthodes expérimentales concernent essentiellement les comparaisons d'organisations pédagogiques.

a) Expérimentation vraie

Il faut des circonstances très spécifiques. On dirige les élèves au hasard vers un modèle d'apprentissage déterminé et on observe les différentes évolutions en fonction du modèle.

b) Expérimentation approchée

2) Méthodes statistiques

Avec les méthodes statistiques, on va analyser la covariance, la régression logistique, la régression multiple...

¹ Professeur à l'université de Nancy (Psychométrie).

a) Principes et exemples

Ces méthodes permettent de définir une variable à partir d'autres variables. Le poids des différentes variables est à préciser, à observer.

b) Problèmes

(1) Technique (quelles variables contrôler ?)

(2) Psychologique (compréhension du contrôle)

C. Sélection post hoc

D. Détection des items à fonctionnement différentiel (IFD)

II. Les évaluations individuelles

A. Comparaisons individu-groupe

1) Fiabilité des comparaisons : les enseignements de la docimologie

Un correcteur est confronté à trois types d'erreurs :

- les erreurs aléatoires (erreur de lecture, d'écriture...);
- les erreurs systématiques non-permanentes (par exemple, les effets de position pendant la correction);
- les erreurs systématiques et permanentes liées à l'attente du professeur envers un ou ses élèves (par exemple, les préjugés sociaux).

2) À quel(s) groupe(s) comparer l'élève ?

Par rapport à la classe, le professeur classe assez facilement les élèves mais a du mal à situer la classe par rapport à autre chose de plus large.

3) Normes statistiques (étalonnages)

Les normes statistiques s'érodent rapidement. Il y a donc une nécessité de les actualiser régulièrement.

B. Comparaisons individu-individu

1) Dans le temps

Pour comparer deux individus dans le temps se posent quelques problèmes dont notamment la fidélité des mesures, les effets de retest et l'effet de régression.

2) Entre domaines

Les domaines peuvent tout aussi bien être les disciplines comme les aspects d'une même discipline (par exemple, l'algèbre et la géométrie en mathématiques).

Ces comparaisons individu-individu entre domaines va permettre de contrôler les différences d'échelle. Les différences d'échelle sont contrôlées par exemple par la présence sur les bulletins de la moyenne de la classe, de la plus haute et de la plus basse moyennes face à la moyenne de l'élève en question.

III. Évaluations sans comparaison (entre individus)

A. Les évaluations critériées

Les évaluations critériées visent à mesurer les acquis des élèves.

1) Origines, concepts, principes

Les évaluations critériées ne cachent plus des comparaisons entre les individus mais des comparaisons entre les tâches.

2) Des comparaisons entre individus aux comparaisons entre tâches

Les relations entre les tâches peuvent être l'équivalence, la hiérarchie ou un ordre simple. Ce type de comparaisons peut permettre l'évaluation de la maîtrise d'un objectif ou l'évaluation de la distance à l'objectif.

L'approche peut être a priori, empirique, statistique ou théorique. Cette dernière approche est indispensable si on veut une remédiation efficace car on fera référence à des théories qui vont permettre d'analyser les relations entre les items.

B. Les modèles de réponse à l'item (MRI)

Intervention par Josette LE COQ ¹
En quoi les informations chiffrées quantitatives peuvent être utiles à l'enseignant, de quoi l'inspection territoriale peut-elle disposer pour faire de la formation ?

Différents types d'informations sont utiles pour construire un cours : les contenus ; les réflexions de chercheurs sur les difficultés des élèves ; les travaux de la DPD ; la prise en compte des résultats de l'évaluation diagnostique.

ATELIER 1 :
La communication des résultats aux évaluations
(aux enseignants, aux parents, à l'institution)
Animation/intervention : Luce DOSSAT ², Pierre ROUSSEAU ³

ATELIER 2 :
Comment exprimer les niveaux de compétences des élèves
(échelle et calibrage)
Animation/intervention : Jacques JUHEL ⁴, Manuel PALACIOS ⁵

ATELIER 3 :
L'évaluation du non cognitif
Animation/intervention : Pascal MALLET ⁶, Ruben BARBOSA ⁷

¹ IA-IPR de Mathématiques – Versailles.

² IEN de Mathématiques – IUFM de Clermont.

³ Principal – Clamart (92).

⁴ Professeur à l'université de Rennes (Psychologie différentielle).

⁵ Directeur du CAEd (Centro de Políticas Públicas e Avaliação da Educação) (• Laboratoire de la mesure de l'évaluation).

⁶ Auteur de Motivations et émotions à l'école : évaluer les dimensions conatives à l'école.

⁷ Professeur et Pro-recteur de l'université de Juiz de Fora (Minas Gerais) – Brésil.

ATELIER 4 : **L'utilisation des résultats et des données par les responsables locaux de l'éducation dans différents contextes et à différents moments**

Animation/intervention : Bernard ACHDDOU ¹

Introduction par l'animateur :

Quels sont les enjeux et les finalités des évaluations ? Quelles communications faire des résultats ? À qui sont-elles destinées ? À qui servent-elles ? Quels sont les besoins des utilisateurs ? des destinataires ? Comment permet-on l'appropriation de ces résultats ?

Un effort de communication est à faire aux élèves et aux parents pas seulement sur les points faibles mais aussi sur les points forts. De plus, au moment des évaluations, on pourrait réinterroger le projet d'établissement.

Bilan des échanges de la matinée :

Il y a une absolue nécessité d'accompagnement à l'évaluation. Il faut mettre en évidence l'importance des enjeux et ainsi s'entendre sur la terminologie et trouver des documents nouveaux qui seraient des documents source. Le rôle de l'impulsion revient au chef d'établissement d'une part et aux équipes pédagogiques en inter et transdisciplinarité d'autre part. Il est important de faire des choix dans un souci de faisabilité. Il faut inscrire les résultats aux évaluations dans le projet de l'élève, dans le projet d'école, dans le projet d'établissement. Il a été relevé les difficultés et les freins à ces démarches, les ruptures et les continuités entre le premier et le second degré.

Thèmes des travaux de groupes :

1. Dans vos différents contextes professionnels, recensez et classez, sous forme d'une typologie, les différents usages des résultats aux évaluations.
2. Selon les destinataires de ces résultats, des précautions et des modalités d'information sont-elles prévues ou à prévoir ?
3. Avez-vous rencontré des difficultés à ce sujet ou dans l'utilisation même des résultats et des données au niveau local ?
4. En situant vos besoins au niveau local, de quels outils aimeriez-vous disposer ? Les outils actuellement disponibles sont-ils suffisants ? À adapter ? Si oui, comment ?

¹ IA adjoint – Nantes.

Groupe 1 : Évaluations : les objets de travail en commun

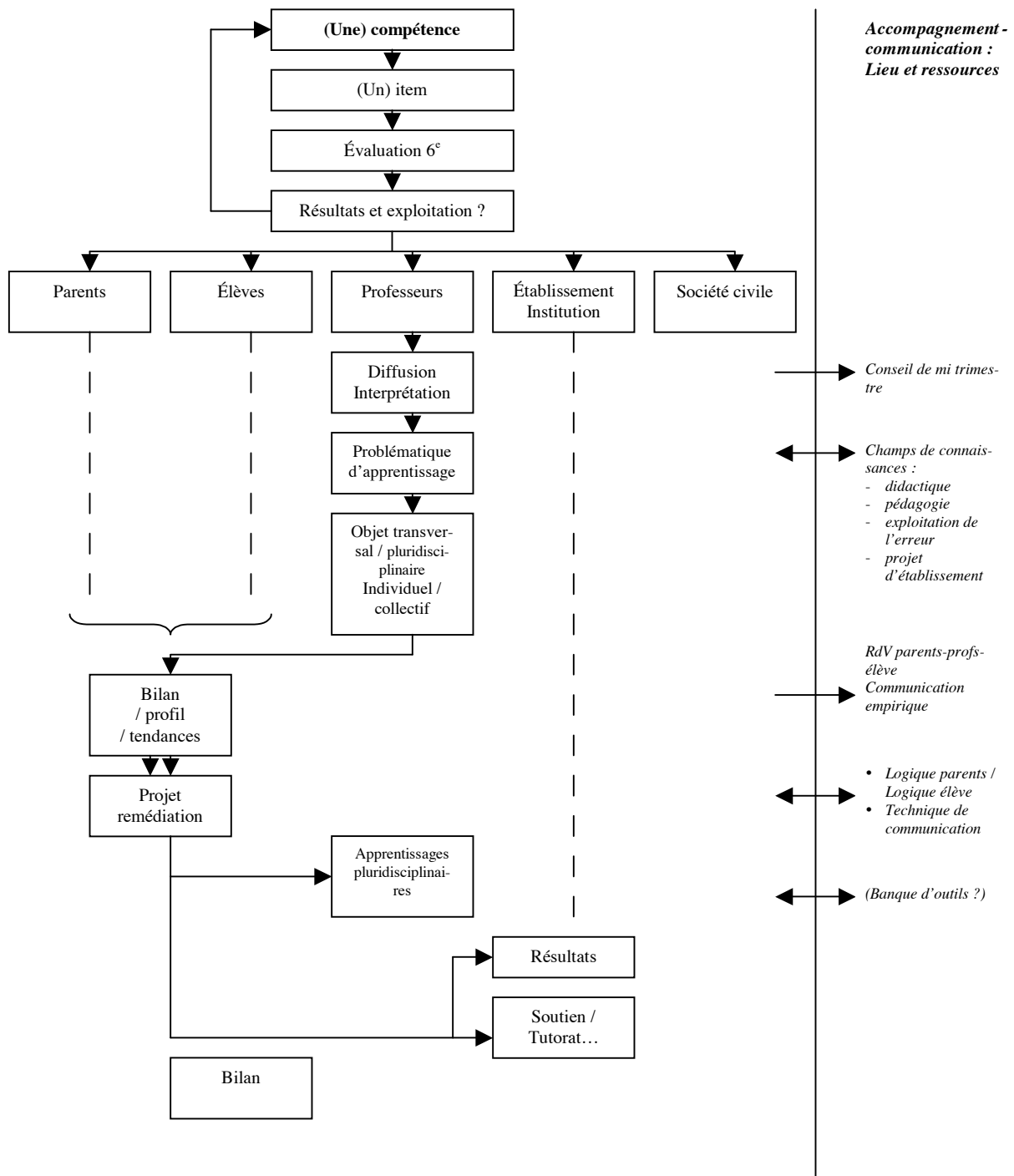
Il y a trois finalités :

	Pour l'élève	Pour l'enseignant	Pour l'institution
Finalités d'apprentissage	<ul style="list-style-type: none"> - donner du sens aux apprentissages scolaires et « donner du sens à l'école » - mesurer ses propres compétences (rendre l'élève acteur) 	<ul style="list-style-type: none"> - aide à la conceptualisation - expliciter l'outil et ses enjeux à l'ensemble des enseignants (polyvalence) - interroger/analyser sa/ses pratiques par rapport à des points forts et faibles - construire des outils de programmation/progression (gestion des apprentissages) pour passer de l'évaluation diagnostique à l'évaluation formative - construire des parcours d'apprentissage (sur le cycle, sur l'année) 	<ul style="list-style-type: none"> - (ré)interroger/construire à travers le projet d'établissement la politique éducative à travers des axes prioritaires (formation continue) - réorienter les axes d'une politique éducative (enjeux politiques forts)
Outils facilitateurs	<ul style="list-style-type: none"> - portfolio - livret de compétences 	<ul style="list-style-type: none"> - aide à la prise de conscience, explicitation des erreurs didactiques/d'apprentissage - mettre en évidence la cohérence dans les apprentissages du cursus scolaire - instaurer les axes prioritaires de la politique éducative de l'école - fonder la culture de l'évaluation, de la gestion des apprentissages 	

Groupe 2 : Utiliser les résultats des évaluations

- Se référer aux livrets du maître.
- Ne pas comparer les résultats entre eux, mais les comparer aux moyennes nationales.
- Accompagnement personnalisé de l'élève :
 - cibler une difficulté, la traiter en petit groupe, en temps limité ;
 - ne pas mélanger des élèves ayant des besoins différents ;
 - se donner un contrat en terme de temps et de contenu ;
 - évaluer à la fin ;
 - faire suivre l'évaluation finale en petit groupe d'une évaluation sur le même point en grand groupe ;
 - à propos d'une difficulté donnée, prendre en petit groupe les élèves ayant obtenu de mauvais résultats avant d'aborder la notion en classe.
- Se donner des priorités.
- Utiliser un certain temps les dispositifs mis en place (par exemple, les heures en « barrette ») même s'ils ne donnent pas entière satisfaction, et se donner les moyens d'en mesurer les effets.

Groupe 3 : L'utilisation des résultats de l'évaluation de 6^e



Conférence par Michel FAYOL¹

Évaluation et apprentissage

I. Évaluation et apprentissage

A. Tests

Si l'on faisait un portrait caricatural, on pourrait dire que les rapports entre ces pratiques (les évaluations et l'apprentissage) se sont cristallisés autour des tests, l'objectif étant d'écarter les élèves qui ne pouvaient pas suivre le cursus. On a construit un instrument d'évaluation « couperet » (Binet).

Le test est toutefois le premier instrument qui vise la facilité d'apprentissage à travers une évaluation de performances. Cette pratique s'est poursuivie en se raffinant mais ce sont restés des instruments statiques.

B. Feed-back

Pour les deux guerres mondiales, il fallait former des militaires rapidement dans des pays où il n'y avait pas de conscription, notamment aux Etats-Unis. On a ainsi développé la méthode qui consiste à donner un feed-back en retour à un acte.

La théorie est très faible mais la technique est extrêmement efficace pour faire apprendre des procédures (orthographe, traitement de texte...). Ce système doit être rapide et sanctionner l'erreur.

Mais rapidement, dans les années 1950-1960, trois obstacles sont nés. D'abord, même sur les animaux, il y a des conditionnements qu'on n'arrive pas à établir (perspective éthologique). Ensuite, les travaux de Piaget (publiés en 1923 mais répandus dans les années 1960) offrent une nouvelle vision de la psychologie des enfants (par exemple, on découvre qu'un enfant, même s'il l'a appris, ne comptera pas le même nombre de jetons selon s'ils sont répartis en ligne, en cercle ou superposés). Enfin, le feed-back n'est pas toujours efficace (par exemple, « l'expérience de la règle » : on fait rechercher sur une règle le point d'équilibre avec un doigt ; mais si la règle est lestée, malgré le feed-back visuel, les enfants n'arrivent pas à trouver l'équilibre... sauf si, curieusement, ils ont les yeux bandés...).

II. Approche cognitive

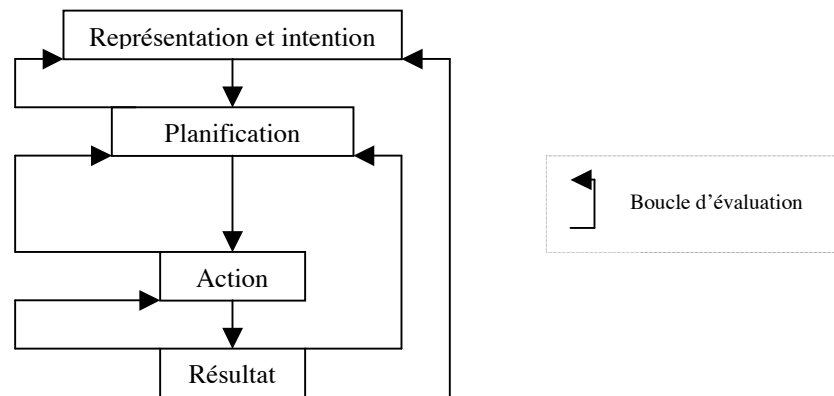
A. Le système cognitif

On va intercaler l'idée d'un univers mental qui intervient.

Lorsque l'action se déploie, il y a une auto-évaluation interne, c'est-à-dire que si je fais quelque chose et que le résultat n'est pas celui escompté, je corrige mon action en fonction du résultat attendu. C'est un fait qui était connu avant, notamment chez les enfants, chez les animaux, etc.

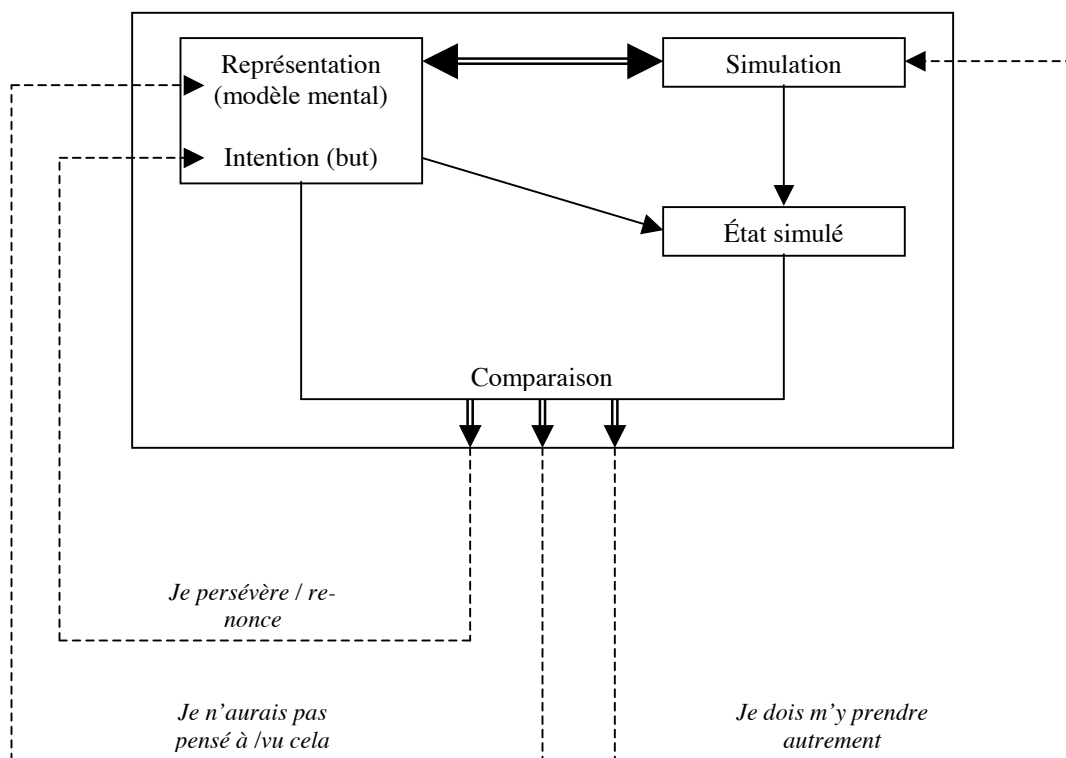
Dans les années 1950, on va développer l'idée d'un système cognitif.

¹ Professeur à l'université de Clermont – Directeur du LAPSCO (Laboratoire de psychologie cognitive).

Le système cognitif

L'évocation de quelque chose de connu induit la représentation symbolique de l'objet (par exemple, si je regarde un visage, une zone du cerveau est activée ; si on me demande de décrire un visage connu, c'est la même zone qui est activée). La combinaison des représentations et des procédures permet de tester avant de mettre en œuvre à l'extérieur (avant de pouvoir observer le feed-back).

Cela a conduit à s'intéresser à l'auto-évaluation.

B. L'auto-évaluation**L'auto-évaluation**

Nous sommes plutôt mauvais en auto-évaluation sauf sur les domaines que l'on connaît très bien ; elle fonctionne très mal et mène souvent à l'échec chez un enfant ou un adulte apprenant. Il faut obligatoirement faire appel à un maître ou un pair.

C. L'attention partagée

On pourrait construire un programme d'évaluation sur les représentations (par exemple, la connaissance des nombres) et on pourrait évaluer les procédures (par exemple, accorder un participe passé). J'évalue mes représentations (savoirs) et mes procédures (savoirs-faire) ; si j'ai bien évalué les deux, cela permettrait de prédire les performances. Mais on en est loin.

1) Par exemple : l'orthographe.

La langue française possède des lettres qui ne se prononcent pas (par exemple, le pluriel « s » ou « ent »). Le fait de savoir qu'il faut mettre un « s » au pluriel (compréhension) n'est pas corrélé au fait de le mettre (production). Les élèves savent mais ne savent pas utiliser. Le problème n'est pas de savoir mais la mise en œuvre du savoir.

Il est question de capacité attentionnelle : à un moment donné et pour un individu donné, l'attention (ou la mémoire) est limitée et ne permet de faire que peu d'actions (c'est plus facile si une des tâches est automatisée). Par exemple, il est facile de parler et de conduire en même temps, mais il est difficile d'écouter et d'écrire en même temps. Lorsque le cumul d'attention requise dépasse la capacité d'action, on cesse l'une des activités, lorsque c'est possible. Par exemple, en automobile, on cesse de parler si un obstacle se présente car il faut faire un effort de concentration. Lorsque c'est impossible, il y a une baisse d'attention sur l'une des activités.

2) Étude de cas

Qu'est-ce qui peut permettre de bonnes performances en dictée ?

- les connaissances en orthographe lexicale ;
- les connaissances en grammaire orthographique (ou orthographe grammaticale) ;
- le niveau de langue, le vocabulaire ;
- les capacités générales.

Cela devrait permettre d'expliquer un pourcentage de la variance (environ 85%). Mais ce n'est pas l'objectif de cette expérience.

L'objectif est de montrer qu'avec des compétences à peu près identiques les enfants qui ont des problèmes de graphisme (écriture lente, maladroite...) ont de moins bonnes performances que ceux qui ont un bon graphisme car plusieurs opérations mentales sont faites en même temps. Ainsi la concentration pour le graphisme se ferait au détriment de la dictée.

L'expérience a montré que les enfants qui ont des difficultés de graphisme au CE2 ne peuvent pas mobiliser leur attention pour les connaissances en orthographe lexical et en grammaire orthographique. La capacité des individus à mobiliser une tâche complexe est donc variable ; ce qui est d'autant plus vrai pour l'écrit qui mobilise de nombreuses connaissances.

III. Questions posées par l'assistance

A. Y aurait-il donc un âge pour la prise de notes ?

Dans la prise de notes sont en jeu deux tâches complexes non automatisées, l'écoute et l'écriture. C'est pourquoi c'est un problème complexe.

B. Que faire quand la capacité attentionnelle d'un élève est limitée ?

Il faut chercher à automatiser pour réduire le coût de certaines tâches ou à apporter des aides pour lui apprendre à gérer progressivement seul la tâche. Si ce n'est pas fait, il y a un risque de simplification, de seulement réaliser le plus simple.

C. Est-ce que cette capacité évolue ou peut évoluer ?

Deux thèses s'opposent à ce sujet. La première défend l'idée que c'est une capacité structurelle donc immuable (comme peut être la couleur des yeux par exemple). La seconde affirme que c'est une capacité qui s'accroît jusqu'à l'âge de 11-12 ans. L'observation d'enfants de 3 à 12 ans peut faire penser que la capacité innée n'augmente pas mais est traitée beaucoup plus rapidement au fur et à mesure.

D. Comment déterminer si l'échec d'un élève à une évaluation est lié à une limite de capacité ou à une autre raison ?

Il faut démonter, séparer chacune des composantes afin de pouvoir le déterminer.

Bibliographie provisoire

Aubret, J. (1995). Psychologie de l'évaluation. In D. Gaonac'h et C. Golder, *Manuel de psychologie pour l'enseignant*. Paris : Hachette Éducation

Belair, L.M. (1999). *L'évaluation et l'école. Nouvelles pratiques*. Paris : ESF.

Bonniol, J.J. & Vial, M. (1997). *Les modèles de l'évaluation : textes fondateurs avec commentaires*. Bruxelles : De Boeck Université.

Depoyer, C. & Noël, B. (eds) (1999). *L'évaluation des compétences et des processus cognitifs. Modèles, pratiques et contextes*. Bruxelles : De Boeck Université.

Hadji, C. (1997). *L'évaluation démystifiée*. Paris : ESF.

Meyer, G. (1995). *Évaluer : pourquoi ? comment ?* Paris : Hachette.

Perrenoud, P. (1998). *L'évaluation des élèves : de la fabrication de l'excellence à la régulation des apprentissages. Entre deux logiques*. Bruxelles : De Boeck Université.

Rémond, M. (2001). Évaluer l'activité de lecture. Les Journées de l'Observatoire : *La lecture de 8 à 11 ans*. Paris : Observatoire National de la Lecture.

Rondal, J.A. (1997). *L'évaluation du langage*. Bruxelles : Mardaga.

Thélot, C. (1993). *L'évaluation du système éducatif. Coûts, fonctionnement, résultats*. Paris : Nathan.

Thélot, C. (1994). L'évaluation du système éducatif français. *Revue Française de Pédagogie*, 107, 5-28.

Publications de la Direction de la Programmation et du Développement (Ministère de l'Éducation Nationale) : *Éducation et Formations, Les dossiers Éducation et Formation, Note d'information, L'état de l'école, Repères et références statistiques sur les enseignements et la formation*.

www.education.gouv.fr/banquoutils

nom d'utilisateur : *outils*

mot de passe : *dpd*

www.iea.nl

Programme de l'université d'été

Lundi 8 juillet	Mardi 9 juillet	Mercredi 10 juillet	Jeudi 11 juillet	Vendredi 12 juillet
	Évaluation diagnostique	Évaluation bilan et comparabilité	Du quantitatif au qualitatif	
	<p><u>9h-10h</u> : Évaluation Diagnostique : Luce Dossat</p> <p>Pause <u>10h30-12h30</u></p> <p>Ateliers/études de cas :</p> <ol style="list-style-type: none"> 1. Construire des outils d'évaluation 2. Interpréter les réponses 3. Les compétences 4. L'accompagnement l'évaluation et à ses usages 	<p><u>9h-10h</u> Du côté de l'international : Alejandro Tiana</p> <p>Pause <u>10h30-12h30</u></p> <p>Table ronde en présence de M. Tiana : les systèmes d'évaluation en France (Jean-Claude Émin) et chez nos partenaires (Brésil, Hongrie, Slovaquie, Russie) et questions/réponses</p>	<p><u>9h-10h</u> Josette Le Coq</p> <p>Pause <u>10h30-12h30</u></p> <p>Ateliers: études de cas</p> <ol style="list-style-type: none"> 1. La communication des résultats 2. Niveaux de compétence, échelle, calibrage 3. Le non cognitif 4. L'utilisation des résultats 	<p><u>9h-10h</u> Rapport des travaux : Quel éclairage, quelles perspectives ?</p> <p>Pause <u>10h30-12h</u></p> <p>Synthèse de l'université Agnès Florin Clôture des travaux : Jean-Richard Cytermann, Bernard Maccario</p>
	Repas	Repas	Repas	Repas
<p><u>13h30</u> : Accueil <u>14h</u> : Ouverture des travaux <u>14h30-15h30</u> : Conférence par Claude Thélot <u>15h30-16h</u> : Organisation de l'UE</p> <p>Pause <u>16h30-18h</u> : Conférence: Élizabeth Bautier, Agnès Florin, Jean-Claude Émin, Dominique Muller</p>	<p><u>14h-16h30</u> Poursuite des ateliers et restitution</p> <p>Pause <u>17h-18h</u> Conférence Faut-il mesurer pour évaluer ? : Jacques Juhel</p>	<p><u>14h-16h30</u> Les évaluations internationales et leur utilisation : PIRLS par Marc Colmant ; PISA par Isabelle Robin et Thierry Rocher et les participants étrangers. Débat</p> <p>Pause <u>17h-18h</u> Conférence Évaluation, comparaison, comparabilité: André Flieller</p>	<p><u>14h-15h30</u> Poursuite des ateliers et restitution</p> <p><u>15h30-16h30</u> Conférence Évaluation et apprentissage : Michel Fayol</p> <p><u>17h30</u> : Apéritif offert par le Président de l'Université de Nantes</p>	

Lieu de l'Université d'Été : Université de Nantes. Faculté des Lettres et Sciences Humaines. Bâtiment de la Censive. Chemin de la Censive du Tertre. 44312 Nantes Cedex.