

Communication C-113
au 21^e colloque de l'Admee-Europe « Evaluation et développement professionnel »
Louvain-la-Neuve, Université catholique de Louvain, 21-23 janvier 2009
(Le contenu de ce texte est publié sous la seule responsabilité de l'auteur)

ETUDE DE L'ACTIVITE EVALUATIVE DE QUELQUES PROFESSEURS DE SCIENCES ECONOMIQUES ET SOCIALES

Vantourout, Marc, Université Paris Descartes – Laboratoire EDA

Goasdoué, Rémi, Université Paris Descartes – Laboratoire EDA

Lanéelle, Xavière, Université de Nantes – Laboratoire CREN

Martins, Daniel, Université Paris Ouest, La Défense

Résumé

Si l'emploi de l'expression activité évaluative tend à se généraliser, peu de travaux s'intéressent précisément à l'analyse de l'activité des évaluateurs notamment lors de la correction de copies. Le travail exploratoire présenté propose une analyse des processus qui conduisent à l'attribution de la note en considérant la correction comme un cas particulier de lecture-compréhension de texte. Nous avons ainsi demandé à des professeurs de corriger une même dissertation de sciences économiques et sociales en verbalisant de manière concomitante à leur lecture leurs remarques, interrogations ou commentaires. Une première analyse des vidéos de ces entretiens-observations montre la valeur heuristique des concepts issus des modèles de compréhension de texte pour analyser leur activité et pour dépasser le constat classique des divergences entre évaluateurs. L'analyse des inférences permet en effet de montrer notamment le rôle des connaissances et des attentes de l'évaluateur lors de sa lecture d'une copie. Ce travail devrait contribuer à renouveler les travaux sur la correction de copies relativement délaissés depuis les travaux pionniers des docimologues expérimentaux, alors que la dissertation reste une épreuve centrale notamment dans le système d'enseignement français.

1. Introduction

Le travail présenté s'inscrit dans un projet de recherche sur l'évaluation à l'université. A titre exploratoire, nous avons entrepris une étude sur l'évaluation de dissertations en sciences économiques et sociales (SES) en classe de terminale. Ce type d'épreuve est très proche de celles sur lesquelles nous allons prochainement travailler, les dissertations en première année de licences de sociologie.

Deux points de clarifications tiendront lieu d'introduction et permettront de situer cette étude dans le champ des recherches sur l'évaluation en éducation.

Le premier concerne l'expression « activité évaluative » qui figure dans le titre même de la communication. Comme l'ont remarqué G. Figari et M. Achouche (2006), cette

expression s'est massivement imposée depuis quelques années dans les textes consacrés aux recherches sur l'évaluation en éducation. Cependant, le recours à cette expression ne renvoie pas, comme c'est le cas pour nous, à une étude de l'activité de l'évaluateur qui est conduite en référence à une théorie de l'activité développée dans le cadre d'une psychologie ergonomique (Leplat, 1997) et qui implique une analyse a priori de la tâche des évaluateurs (Vantourout & Maury, 2006). L'enjeu principal est alors de mieux connaître et comprendre le déroulement de l'activité de correction de professeurs en nous centrant sur la dimension cognitive de leur activité. Ajoutons enfin que nous distinguons deux facettes dans l'activité évaluative, la conception d'épreuves et la correction d'épreuves et que nous n'aborderons ici que la dernière facette.

Le second point concerne les liens entre l'approche que nous présentons ici et les travaux des docimologues (pour une synthèse, voir Piéron, 1963) et des docimologues expérimentaux (pour une synthèse, voir Noizet & Caverni, 1978), travaux que nous prolongeons en quelque sorte en nous intéressant à la correction de copies. Cependant, ces travaux n'ont jamais véritablement porté sur l'évaluateur, ni sur son activité. Le parti pris par les premiers docimologues d'assimiler les divergences entre correcteurs à des erreurs de mesure conduit inévitablement à ne pas s'intéresser aux correcteurs et à leur activité. Il en va de même pour les docimologues expérimentaux : identifier des déterminants externes auxquels seraient sensibles tous les évaluateurs conduit inévitablement à faire disparaître le correcteur derrière la note qu'il a produite. Notre option est d'appréhender la correction de copie comme une tâche particulière de lecture-compréhension de texte. Cela permet d'étudier le déroulement de l'activité de correction de copie en s'intéressant aux caractéristiques de l'évaluateur et à un nouveau type de déterminants internes à l'évaluateur : les connaissances et représentations qui sont impliquées dans les processus de lecture-compréhension. Précisons que l'intérêt porté aux connaissances de l'évaluateur n'a pas pour fin de juger de sa légitimité, mais de mieux comprendre le rôle de celles-ci dans les processus de lecture-compréhension d'une copie.

2. Un nouveau cadre d'analyse de l'activité de correction de copie

2.1. La correction de copie comme un cas particulier de jugement

Considérant l'évaluation comme un cas particulier d'activités de jugement, les docimologues expérimentaux (Noizet & Caverni, 1978) ont transposé des concepts issus des travaux sur la perception, notamment ceux portant sur l'estimation du poids (Noizet, 1967). Outre les notions d'ancrage, de contraste, etc., la notion principale reprise de ces travaux est celle de comparaison. Le jugement est essentiellement vu comme une comparaison, ce qui renvoie indirectement aussi à la notion de mesure. Le jugement évaluatif est conçu comme une comparaison entre la copie à évaluer et le modèle de référence (Noizet & Caverni, 1978). Les déterminants évoqués précédemment (§ 1) affectent ces processus de prise d'information et de comparaison.

La démarche que nous proposons remet en cause la place centrale accordée au processus de comparaison, mais aussi les termes de cette comparaison. L'activité déployée pendant la correction de copies ne peut être assimilée à un prélèvement-accumulation d'indices. Nous considérons au contraire l'activité évaluative comme un processus de compréhension au sens conféré à ce concept dans les travaux sur la compréhension de texte. L'enjeu n'est plus d'analyser ce qui pèse sur le processus de comparaison, mais de comprendre comment s'élaborent les représentations issues de la lecture de la copie. Le jugement évaluatif s'appuie plus vraisemblablement sur cette représentation que sur une accumulation d'indices positifs ou négatifs. Par ailleurs, on

peut également s'interroger sur l'existence a priori du second terme de la comparaison, c'est-à-dire du modèle de référence. Des travaux sur l'évaluation formative en mathématiques (Vantourout, 2007) montrent en effet que certains professeurs élaborent leur modèle de référence au fil de l'activité de correction. Pour certaines épreuves complexes, dont celle que nous étudions, il semble que le jugement ne repose pas sur une comparaison au sens strict, mais sur l'élaboration d'une représentation du texte lu à partir d'un système d'attentes qui n'est ni complètement prédéfini, ni figé (Ardoino et Berger 1986). Cette double mise en cause de la comparaison et des termes de la comparaison nous conduit à changer de cadre théorique de référence et à appréhender la correction de copie comme un cas particulier de lecture-compréhension de texte.

2.2. *L'activité de correction de copie comme une activité de lecture-compréhension*

Si le modèle de l'évaluation-comparaison peut sembler satisfaisant lorsque l'on s'intéresse principalement à la note, sa valeur heuristique semble plus limitée pour étudier l'activité de correction. L'analogie que nous proposons permet de rendre compte d'aspects jusqu'ici négligés, notamment le rôle déterminants internes de l'activité tels que les connaissances de l'évaluateur (mentionnées au § 2.1) ou encore les limites de ses capacités de traitement dans ce contexte particulier de lecture.

La plupart des modèles élaborés dans le domaine de la compréhension de texte (van Dijk & Kintsch, 1983) s'accordent sur la nécessité de distinguer mémoire du texte et mémoire de la compréhension du texte. Comprendre un texte n'est évidemment pas mémoriser une succession d'énoncés, mais consiste à intégrer la succession des propositions dans une représentation cohérente (Blanc & Brouillet, 2003). Les inférences produites lors de la lecture jouent ainsi un rôle déterminant dans la compréhension. La compréhension est alors conçue comme un processus interactif dans lequel les connaissances et représentations du lecteur occupent une place prépondérante.

Dans le cadre de l'activité de correction les inférences produites peuvent être un moyen privilégié d'accès aux connaissances sollicitées lors de la lecture¹. On peut également supposer que la finalité évaluative de cette activité de lecture suscite la construction d'une représentation non seulement du texte et mais aussi l'élaboration d'une représentation évaluative du texte, voire de l'élève évalué.

La notion d'inférence renvoie directement à celle de cohérences (locale/globale) qui sont indispensables pour rendre compte de la dynamique de la construction de la représentation d'un texte. Initialement abordées dans des récits, ces notions pourraient s'avérer pertinentes pour la compréhension du jugement évaluatif. En effet, des exercices tels que la dissertation sont régis par des contraintes formelles précises, censées guider le lecteur dans sa compréhension de l'argumentation. L'incohérence – qu'elle soit locale ou globale, qu'elle se situe à un niveau conceptuel ou organisationnel (plan de la dissertation) – pèse sur le jugement ; elle impose notamment un effort accru de compréhension qui peut s'avérer vain et coûteux à l'égard de l'enjeu sur la note finale.

Les phénomènes qui viennent d'être évoqués à propos de la notion de cohérence renvoient à leur tour, aux limites des capacités de traitement mentionnées précédemment. Cet aspect paraît crucial pour analyser l'activité de correction, car la lecture d'un lot de copies influe nécessairement sur la disponibilité de connaissances en mémoire (activation accrue due à la répétition) ou encore sur la possibilité de mémorisation (textes longs).

Ces quelques illustrations montrent l'intérêt de la transposition des notions issues des

1 Pour le moment les inférences n'ont été étudiées que dans un cadre expérimental ; elles ne sont que partiellement accessibles par la méthode de verbalisation concomitante que nous utilisons.

travaux en compréhension de texte à la correction copies pour renouveler l'analyse de l'activité des évaluateurs et pour mieux comprendre le processus qui mène au produit final qu'est la note.

3. Travail exploratoire

3.1 Choix de l'épreuve de dissertation et des copies à corriger

Le travail exploratoire porte sur la correction de copies en SES. Les copies ont été produites par des élèves de terminale ES (économique et sociale), confrontés lors d'un bac blanc à une épreuve de dissertation avec étude de documents. Le sujet est un sujet de sociologie « classique » (figurant dans les annales du bac) : « *Dans quelle mesure les mutations du travail peuvent-elles expliquer la crise du syndicalisme ?* ».

Nous avons choisi de travailler à partir de la dissertation en SES car il s'agit d'une épreuve particulièrement adaptée aux enjeux théoriques de ce travail (correction de copies comme une activité de compréhension de texte). En effet, les copies produites sont très riches car on demande aux élèves de dissenter en prenant appui sur des documents – 3 textes et 3 documents statistiques, dans le cas présent – qui accompagnent le sujet. De plus, la dissertation occupe une place particulière dans le système éducatif français où elle peut être considérée comme un exercice élitiste auquel sont confrontés beaucoup de lycéens et d'étudiants, indépendamment de leur niveau académique.

Nous avons recueilli un lot de copies² de terminale ES auprès d'un enseignant, Prof. 11 (cf. tab. 1 § 4), lequel a choisi le sujet de la dissertation. En nous appuyant sur les compétences de l'une d'entre nous – sociologue, agrégée de SES et responsable de la formation des professeurs de SES dans un IUFM –, nous avons extrait du lot 3 copies (copies A, B et C), dont 2 « moyennes », sachant que ce sont celles qui posent le plus de difficultés aux évaluateurs. Parmi ces dernières, nous voulions une copie (copie A) qui se lise aisément – nous parlons ici du plan rédactionnel –, mais plutôt « creuse » conceptuellement. C'est de cette copie dont il sera essentiellement question dans le travail présenté ici. A l'inverse, nous voulions une copie (copie B) plus difficile à lire (maladresses) que la copie A, mais beaucoup plus aboutie sur le plan conceptuel. Enfin, nous avons choisi une « mauvaise » copie (copie C). Dans l'orientation retenue, comme nous l'avons mentionné (cf. § 1), l'analyse a priori de la tâche d'évaluation constitue un aspect fondamental, en relation étroite avec le choix des copies. Les 3 copies ont fait l'objet d'une analyse propositionnelle ainsi que d'une analyse de leur contenu sur les plans conceptuel et formel. Faute de place, nous ne développerons pas ce point.

3.2 Situation d'observation

La tâche proposée aux professeurs est une tâche de correction avec verbalisation concomitante, la consigne étant: « *Vous êtes professeur en terminale ES, vous avez choisi ce sujet de type Bac pour préparer vos élèves à l'épreuve de dissertation avec document. Vous corriger, annotez et attribuez une note aux copies proposées. Vous rédigez une appréciation. Cette note comptera pour la moyenne trimestrielle et pour la formulation de l'avis pour le bac. Vous commentez à haute voix votre travail de correction. Vous n'hésitez pas à formuler toutes les remarques que vous jugez nécessaires et utiles à notre compréhension.* ». Les 3 copies manuscrites (photocopies) sont remises au correcteur qui

2 Nous disposons d'une vingtaine de copies, ce qui correspond à l'ensemble des copies de la classe où nous nous les sommes procurées, les élèves n'ayant pas eu le choix entre plusieurs sujets.

se trouve en position d'évaluateur externe (il ne connaît pas les élèves et n'a pas choisi le sujet). L'ordre de correction des copies est imposé : A, B, C. L'activité de chaque professeur est entièrement filmée, le zoom de la caméra fixant la copie.

Pour l'instant, nous avons observé 10 professeurs, tous agrégés. Nous les avons répartis en trois groupes selon leur ancienneté ; premier groupe, 1 à 2 ans d'ancienneté ; deuxième groupe, autour de 10 ans ; troisième groupe, plus de 20 ans.

C'est en comparant l'activité de correction de ces professeurs, mais surtout en procédant à une comparaison intergroupe que nous rejoignons la thématique du colloque, développement et transformation des compétences professionnelles.

3.3 Présentation de la grille de description des vidéos

La grille de description des vidéos (cf. fig. 1) est constituée d'items et de catégories (ces dernières indiquées en italiques) en relation étroite avec le cadre théorique ; elle dépend également des fonctionnalités du logiciel utilisé (Actogram) qui permet des regroupements d'items par catégories. À partir des vidéos et des copies corrigées, nous réalisons un travail de repérage (n° de la page lue et n° de ligne) et de transcription des annotations et verbalisations. Ce sont les annotations et les verbalisations qui sont analysées à l'aide de la grille. Celle-ci est structurée par l'idée de décrire les divers éléments (dont certains étant des déterminants) du jugement évaluatif intervenants au fil de la correction. La première catégorie concerne logiquement, parmi les annotations et verbalisations, celles que l'on peut qualifier de *jugements* (sur les différents aspects du travail évalué). Ces jugements sont eux-mêmes qualifiés par la catégorie *orientation* (quand ces orientations sont explicitement verbalisées ou écrites). Considérant que les *inférences* sont des processus déterminants dans les jugements évaluatifs, nous avons décrit différentes formes d'inférences : inférences de compréhension, inférences d'évaluation sur les contenus et inférences d'évaluation sur les élèves. Enfin, nous relevons des éléments factuels qualifiés de *comportements* qui pourraient également éclairer les jugements, par exemple, les moments de retours en arrière, de lecture des documents.

Cette grille permet pour chaque professeur d'obtenir un relevé d'activité (description temporelle en fonction des catégories et des items de la grille) et des données synthétiques (nombre d'occurrences, fréquence, durée pour les items et catégories)

| | | | |
|----------------------|---------------------------------------|----------------------------------|-------------------|
| <i>Jugements</i> | Jugement sur la forme | <i>Orientation des jugements</i> | Positif |
| | Jugement sur la méthode | | Positif - réservé |
| | Jugement sur le contenu | | Neutre |
| | Jugement sur l'élève | | Négatif - réservé |
| | | | Négatif |
| <i>Inférences</i> | Inférence de compréhension | | |
| | Inférence d'évaluation sur le contenu | | |
| | Inférence d'évaluation sur l'élève | | |
| <i>Comportements</i> | Regarde sujet - documents | | |
| | Ecrit sur une grille d'évaluation | | |
| | Retours en arrière | | |
| | Feuillette la copie | | |

Figure 1 : Grille de description des vidéos – items regroupés par catégories

4. Premiers résultats

4.1 Constat « classique » en docimologie

L'analyse des notes (cf. tab. 1) confirme un constat classique établi en docimologie : l'existence de divergences entre professeurs lorsqu'il s'agit de noter une ou plusieurs copies.

L'ensemble des notes attribuées et la lecture de quelques indicateurs statistiques (moyenne, écart-type) confortent les choix faits lors de la sélection des copies. Les copies A et B sont effectivement des copies « moyennes » ; la copie « C » est bien une « mauvaise » copie.

Concernant les copies A et B, on peut même dire que celles-ci sont « proches » si l'on se contente d'observer leur moyenne (respectivement, 11,55 et 12,18), leur écart-type (respectivement, 2,5 et 2,36), l'amplitude des notes (7 points). Notons, au sein de cet effectif limité, qu'il y a quasiment autant de professeurs qui « préfèrent » la copie A que de professeurs qui « préfèrent » la copie B (préférence surlignée dans le tableau). On retiendra donc, quelle que soit l'ancienneté, qu'il existe des écarts de notations entre professeurs et que certains préfèrent la copie A et d'autres la copie B.

Mais, laissons-là ces résultats attendus. En effet, nous cherchons à les dépasser en étudiant les processus qui conduisent à l'attribution d'une note.

| Profs. | copie A | copie B | copie C | Ancienneté | Groupe |
|------------|---------|---------|---------|------------|-------------------------------------|
| Prof. 1 | 14 | 10 | 7 | 1 | G1 : 1-2 années d'ancienneté |
| Prof. 2 | 13 | 12 | 6 | 1 | |
| Prof. 3 | 13 | 16 | 7 | 2 | |
| Prof. 4 | 9 | 16 | 9 | 7 | G2 : env. 10 années d'ancienneté |
| Prof. 5 | 14 | 12 | 4 | 13 | |
| Prof. 6 | 10 | 10 | 5 | 13 | |
| Prof. 7 | 11 | 9 | 5 | 20 | G3 : plus de 20 années d'ancienneté |
| Prof. 8 | 15 | 13 | 5 | 22 | |
| Prof. 9 | 8 | 13 | 6 | 24 | |
| Prof. 10 | 8 | 10 | 5 | 25 | |
| Prof. 11* | 12 | 13 | 7 | 30 | |
| Moyenne | 11,55 | 12,18 | 6 | | |
| Ecart-type | 2,5 | 2,36 | 1,41 | | |

Tableau 1 : Notes attribuées par chaque professeur aux trois copies, moyenne et écart-type pour chaque copie ; Ancienneté et groupe de chaque professeur (*Pr11 est l'enseignant qui nous a fourni les copies, il note ses élèves)

4.2 Première données quantitatives issues des observations

Ce que nous présentons maintenant s'appuie sur l'analyse de l'activité de 7 professeurs, corrigeant la copie A.

Les premières analyses (cf. tab. 2), réalisées à l'aide de la grille (cf. § 3.3) montrent que le nombre de jugements (138, tous types confondus) est bien supérieur au nombre d'inférences (59, tous types confondus). Cet écart pourrait résulter de la tâche de

verbalisation ; il est probable en effet que la plupart des inférences sont tuées car ce sont des processus extrêmement rapides. Ce trait a conduit, dans les travaux en compréhension de texte, à mesurer par exemple des temps de lecture ou à poser des questions sur les connaissances sollicitées pour une inférence, ce qui s'avérait impossible dans cette première situation d'observation contrôlée. Par ailleurs, il semble se dessiner une différence de stratégies de lecture, différence qui pourrait être décrite par l'écart entre le nombre de jugements et le nombre d'inférences. Nous avons relevé que certains professeurs s'engagent plus nettement dans la compréhension de la copie que d'autres qui semblent avoir une lecture « plus de surface » (il ne s'agit pas d'un jugement de valeur; le professeur peut centrer son activité sur des aspects autres que la compréhension « fine »). Ces tendances doivent être confirmées par l'analyse de l'ensemble du corpus de vidéos.

| Grp | Profs. | Notes | Inf. de compréhension ^o | Inf. éval. / contenu | Inf. éval. / l'élève | Jugem ^t formel | Jugem ^t methodo | Jugem ^t contenu |
|-----|----------|-------|------------------------------------|----------------------|----------------------|---------------------------|----------------------------|----------------------------|
| G1. | Prof. 2 | 13 | 2 | 2 | 1 | 10 | 12 | 12 |
| | Prof. 3 | 13 | 5 | 3 | 0 | 2 | 7 | 18 |
| G2. | Prof. 5 | 14 | 3 | 1 | 2 | 5 | 5 | 16 |
| | Prof. 6 | 10 | 9 | 2 | 2 | 9 | 4 | 10 |
| G3. | Prof. 8 | 15 | 5 | 0 | 2 | 12 | 13 | 40 |
| | Prof. 9 | 8 | 1 | 0 | 1 | 1 | 4 | 19 |
| | Prof. 10 | 8 | 7 | 3 | 8 | 8 | 14 | 23 |

Tableau 2 : Description quantitative des vidéos à partir d'items de la grille de description

4.3 Descriptions qualitatives : illustrations des types d'inférence

Les verbalisations qualifiées d'inférences étant moins communes que celles qualifiées de jugements, nous présentons à l'aide d'un exemple les trois types d'inférences que nous distinguons et qui montrent le rôle des déterminants internes au correcteur dans l'élaboration de la représentation qu'il se fait de la copie.

4.3.1 Inférences de compréhension

Le premier exemple est la verbalisation d'une inférence de compréhension. Le lecteur-correcteur reconstitue une information non-explicite dans le texte pour le comprendre et ce indépendamment de tout jugement évaluatif. La phrase lue dans la copie étant « *les syndicats sont devenus des institutions* », le professeur (Prof. 6) va chercher si le mot « institutionnalisation » figure dans les documents associés au sujet. Il le trouve effectivement dans l'un des documents, mais ce mot n'est pas présent dans la copie. On comprend ainsi que l'expression « devenir des d'institutions » fait référence, pour le professeur, au concept d'« institutionnalisation des syndicats » (inscrit aux programmes) sans que le mot « institutionnalisation » figure dans la copie. La compréhension de la phrase a donc bien nécessité l'intervention d'une inférence, c'est-à-dire l'évocation d'une connaissance non-explicite dans la copie. Un lecteur ne connaissant pas le domaine ne ferait probablement pas cette inférence. Cet exemple montre précisément le rôle des connaissances dans l'élaboration de la représentation de la compréhension du texte et donc indirectement sur l'évaluation. En complément, on pourrait avancer, en s'appuyant sur ce qui a été montré dans le cadre des processus d'intégration lors de la construction de la représentation mentale du contenu sémantique du texte (van Dijk & Kintsch, 1983), que ce professeur résume (peut-être pour mémoriser) la phrase « les syndicats sont devenus des institutions » en faisant appel au

concept d'« institutionnalisation des syndicats ».

4.3.2 Inférences d'évaluation sur le contenu

Le deuxième exemple porte sur le même passage de la copie, comme dans le premier exemple, le lecteur-correcteur (Prof. 3) fait appel au concept d'« institutionnalisation » non-explicite dans le texte et verbalise sur ce contenu en prêtant des connaissances à l'élève : « *l'élève fait appel au concept d'institutionnalisation, c'est le terme canonique qui est dans le programme donc là il a bien repéré, c'est bien vu* ». La fin de cette verbalisation (partie soulignée) montre que l'on passe d'une inférence de compréhension à une inférence d'évaluation sur le contenu, pour déboucher sur un jugement positif sur le contenu (catégorie orientation). La lecture-compréhension du passage « les syndicats sont devenus des institutions » n'est pas valorisée de la même manière par le professeur (Prof. 6) cité au § 4.3.1 qui cherche à vérifier si le contenu vient de l'élève ou s'il est présent dans les documents. Ces divergences sur le plan des inférences à partir d'un même passage, d'une évocation semblable du concept « d'institutionnalisation », contribuent probablement à celles observées au niveau de des notes. Même si un seul cas ne peut expliquer un écart de notation, il se trouve qu'en l'occurrence le premier correcteur (Prof. 6) met 10/20 à cette copie alors que le second (Prof. 3) qui n'a pas recherché l'origine de l'information met 13/20. Ces enchaînements d'inférences et de jugements révèlent de manière implicite la complexité des attentes de l'évaluateur ; le premier correcteur (§ 4.3.1, Prof. 6) regarde quelle utilisation l'élève fait des documents, sans systématiquement valoriser cet aspect du travail. Un même critère peut changer de statut selon le contexte, ce qui montre bien que la correction n'est pas qu'une succession de prélèvement d'indices, mais au minimum une interprétation d'éléments présents dans la copie.

4.3.3 Inférences sur l'élève

Le troisième exemple est la verbalisation d'une inférence sur le type présumé d'élève auquel est censé appartenir l'auteur de la copie. Le lecteur-correcteur (Prof. 10) dit après avoir une nouvelle fois rencontré des difficultés de compréhension : « *c'est quelqu'un qui a appris son cours c'est sûr, c'est terrible les copies comme ça, on voit que c'est l'élève qui a travaillé, qui a des connaissances mais qui n'arrive pas à avoir une analyse rigoureuse et pertinente* » Cette supposition sur l'appartenance de l'élève à une catégorie plus générale illustre clairement la construction concomitante d'une représentation du texte et de son auteur³. Le lecteur-correcteur s'appuie pour cette inférence sur une norme omniprésente dans l'évaluation, celle d'internalité (Bressoux et Pansu, 2003). Il fait des inférences sur ce que l'élève est à partir de ce qu'il fait.

Ces trois exemples montrent bien comment les inférences permettent d'accéder au système d'attentes de l'évaluateur dans ses diverses dimensions (forme, méthode, contenu, élève, etc.). Bien que ces trois formes d'inférences interviennent de manière plus moins directe lors l'évaluation, on ne peut pas, pour le moment, savoir si et comment elles pèsent sur le jugement final.

³ Il nous paraît important d'ajouter que le jugement sur l'élève énoncé par ce professeur (jugement qui repose sur l'inférence d'évaluation sur l'élève rapportée au § 4.3.3) va complètement dans le sens du jugement bienveillant de l'enseignant de l'élève ayant rédigé la copie (Prof. 11). Prof. 11 dit mettre 13/20 à cette copie pour encourager cet (son) élève qui commet des maladrotes mais qui s'investit dans son travail pour progresser et réussir. L'autre professeur (Prof.10) attribue quant à lui 8/20.

5. Synthèse – discussion

Au-delà des premières descriptions de l'activité qui viennent d'être présentées, il semble se dessiner une différence entre deux stratégies de lecture. Ces stratégies pourraient s'apparenter à celles mises en évidence par S. Chen et S. Chaiken (1999) dans le cadre de leur modèle dual du traitement heuristique-systématique, développé pour analyser la réception de messages à visée persuasive⁴. Certains lecteurs-correcteurs semblent procéder à un recueil d'indices qui sont successivement évalués et mis en relation. Cette manière de procéder correspondrait à un traitement heuristique (qui devra être spécifié). D'autres correcteurs fondent leur évaluation sur la représentation de la copie issue d'un traitement systématique du texte (plus coûteux cognitivement). Ces deux stratégies de prélèvement d'indices ou de compréhension plus approfondie peuvent être décrites comme des processus vicariants (Reuchlin, 1978). Elles remplissent globalement la même fonction (servir l'évaluation), mais leurs différences influencent très certainement la représentation que se fait le correcteur de la copie et en conséquence la manière dont il va la juger. On peut également supposer que l'emploi de ces stratégies varie en fonction de trois éléments : (1) le type d'épreuve, (2) les caractéristiques de la copie, (3) les caractéristiques des évaluateurs. (1) L'emploi d'une stratégie de traitement systématique semble en effet moins adapté à la correction d'une dictée que pour celle d'une dissertation. Nous avons par ailleurs constaté dans ces premières observations que les correcteurs « renoncent » parfois à comprendre, ils abandonnent (ou réduisent) la stratégie de traitement systématique lorsque celle-ci s'avère trop coûteuse par rapport à ce qu'une compréhension fine apportera au jugement. (2) Les caractéristiques de la copie semblent déterminantes dans l'organisation de l'activité. Tout ce qui favorise la lisibilité et/ou la compréhension rend l'emploi de cette stratégie de lecture plus fréquent. (3) Les conceptions de l'évaluation chez les correcteurs, ainsi que les techniques et outils de correction, peuvent aussi avoir un effet sur l'orientation vers l'une ou l'autre stratégie. Par exemple, l'utilisation de grilles de correction composées d'une liste de points positifs ou négatifs semblerait inciter à une stratégie de prélèvement d'indices plutôt qu'à un traitement systématique. Ce premier constat devra être confirmé par l'analyse de l'ensemble du corpus.

Au-delà de cette différence de stratégies de lecture, les premières analyses suggèrent également l'importance de la construction de couples d'attentes-confirmations dans l'activité évaluative. La prise en compte d'un indice ou encore la production d'une inférence de la part du correcteur va susciter la construction d'attentes tant sur le contenu que sur la valeur globale de la copie. On retrouve ici les phénomènes de cohérence locale ou globale évoqués précédemment (*cf.* § 2.2). L'incohérence peut orienter le lecteur vers un arrêt du traitement systématique (*cf.* infra. renoncement à la compréhension) au profit d'une stratégie de simple prélèvement d'indices.

Enfin, notons que dans les analyses présentées (*cf.* tab. 1 et tab. 2) rien ne permet d'opérer des distinctions reposant sur l'ancienneté des correcteurs. La taille de l'effectif étudié (7 professeurs) nous paraît ici peu importante, car nous croyons que l'enjeu de ces travaux réside ailleurs que dans la recherche de relations entre stratégies/procédures de correction et ancienneté. Avant tout, il nous semble que la production de connaissances relatives aux diverses manières dont les correcteurs peuvent se réappropriier la tâche de correction de dissertations puis corriger celles-ci pourrait constituer un apport très utile

4 Il s'agit pour l'instant de réfléchir aux apports potentiels de ce modèle lors de l'analyse des observations réalisées et à venir. Notons que pour S. Chen et S. Chaiken (1999), il ne s'agit pas de distinguer deux types de traitements alternatifs, car ces traitements sont tous deux impliqués dans des dosages subtils.

pour la formation initiale et continue des professeurs et contribuer ainsi à leur développement professionnel.

6. Perspectives

Les premières descriptions présentées ici seront ultérieurement complétées par la confrontation à d'autres données, non prises en compte actuellement. L'analyse des appréciations rédigées au début des copies ainsi que celle des commentaires généraux produits au terme de la lecture peuvent s'avérer très riches pour comprendre les représentations forgées au cours de la lecture. Ce travail sera également complété par une mise en relation de l'activité et de l'analyse de la tâche d'évaluation qui repose sur une analyse approfondie de la copie (mentionnée au § 3.1).

Si ce travail exploratoire permet une nouvelle description de l'activité de correction de copie, il ne permet pas pour le moment de relier de manière directe les caractéristiques étudiées de l'activité et la note. Une confirmation d'attente ne conduit pas nécessairement à un jugement positif, de même que l'emploi d'une stratégie de lecture de type prélèvement d'indices n'est pas toujours le signe d'un refus de compréhension comme nous avons pu le constater. Il semble donc nécessaire d'approfondir ce travail de description et probablement de le compléter par une approche plus expérimentale reposant sur des copies construites en fonction des hypothèses dégagées dans la première phase plus descriptive de nos travaux actuels.

Enfin, ce travail est prometteur quant à la valeur heuristique de la transposition des modèles de compréhension de texte pour l'analyse de l'activité de correction de copies, et ce malgré les limites « écologiques » de la tâche de verbalisation concomitante. La redéfinition de l'activité d'évaluation comme une activité de compréhension devrait contribuer à renouveler les perspectives ouvertes par la docimologie expérimentale sur la correction de copies. Malgré le relatif désintérêt depuis une vingtaine d'années pour ce type de recherches les problèmes soulignés par les premiers docimologues demeurent, d'autant que la dissertation reste une épreuve centrale notamment dans le système d'enseignement français.

Références bibliographiques

Ardoino, J. & Berger, G. (1986). L'évaluation comme interprétation, *Pour*, 107, 120-127.

Blanc, N. & Brouillet, D. (2003). *Mémoire et compréhension - lire pour comprendre*. Paris: In Press.

Bressoux, P. & Pansu, P. (2003). *Quand les enseignants jugent leurs élèves*. Paris : PUF.

Chen, S. & Chaiken, S. (1999). The Heuristic-Systematic Model in Its Broader Context. In S. Chaiken & Y. Trope (Ed.), *Dual Process Theories in social Psychology* (pp. 73-96). New York London: The Guilford Press.

Figari, G. & Achouche, M. (2001). *L'activité évaluative réinterrogée : Regards scolaires et socioprofessionnels*. Bruxelles : De Boeck.

Leplat, J. (1997). *Regards sur l'activité de travail - contribution à la psychologie ergonomique*. Paris : PUF.

Noizet, G. & Caverni, J-P. (1978). *Psychologie de l'évaluation scolaire*. Paris : PUF.

Noizet, G. (1967). *Influence de la nature verbale des réponses sur l'ancrage perceptif, Hommage à A. Rey*, Bruxelles : Dessart, 239-257.

Piéron, H. (1963). *Examens et Docimologie*. Paris : PUF.

Reuchlin, M. (1978). Processus vicariants et différences interindividuelles. *Journal de Psychologie*, 2, 133-145.

Van Dijk, T. A. & W. Kintsch. (1983). *Strategies of Discourse Comprehension*. New York : Academic.

Vantourout, M. & Maury, S. (2006). Quelques résultats relatifs aux connaissances disciplinaires de professeurs stagiaires dans des situations simulées d'évaluation de productions d'élèves en mathématiques. *Revue des sciences de l'éducation*, 32(3), 759-782.

Vantourout, M. (2007). Etude de l'activité évaluative de professeurs stagiaires confrontés à des productions d'élèves en mathématiques : Quel référent pour l'évaluateur ? *Mesure et évaluation en éducation*.30(3), 29-58.