

Étude de l'activité évaluative de professeurs stagiaires confrontés à des productions d'élèves en mathématiques : quel référent pour l'évaluateur ?

Marc Vantourout

Volume 30, numéro 3, 2007

URI : <https://id.erudit.org/iderudit/1085728ar>

DOI : <https://doi.org/10.7202/1085728ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (imprimé)

2368-2000 (numérique)

[Découvrir la revue](#)

Citer cet article

Vantourout, M. (2007). Étude 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.
<https://doi.org/10.7202/1085728ar>

Résumé de l'article

Nous étudions l'activité évaluative de dix-huit professeurs stagiaires – professeurs des écoles (enseignement primaire) et professeurs de mathématiques (enseignement secondaire) – confrontés à des productions d'élèves en mathématiques. Les observations réalisées visent principalement l'identification du référent de l'évaluateur. Le dispositif de recherche organisé autour d'une « situation aménagée » permet de proposer des conditions expérimentales analogues pour tous les professeurs. Les analyses réalisées nous amènent à distinguer quatre profils de comportements évaluatifs et à relever plusieurs phénomènes évaluatifs.

Étude de l'activité évaluative de professeurs stagiaires confrontés à des productions d'élèves en mathématiques : quel référent pour l'évaluateur ?

Marc Vantourout

Université Paris Descartes, Laboratoire EDA
(Éducation et Apprentissages)

MOTS CLÉS: Activité évaluative, situation aménagée, référent, professeurs stagiaires, proportionnalité, représentation graphique

Nous étudions l'activité évaluative de dix-huit professeurs stagiaires – professeurs des écoles (enseignement primaire) et professeurs de mathématiques (enseignement secondaire) – confrontés à des productions d'élèves en mathématiques. Les observations réalisées visent principalement l'identification du référent de l'évaluateur. Le dispositif de recherche organisé autour d'une « situation aménagée » permet de proposer des conditions expérimentales analogues pour tous les professeurs. Les analyses réalisées nous amènent à distinguer quatre profils de comportements évaluatifs et à relever plusieurs phénomènes évaluatifs.

KEY WORDS: Assessment activity, simulation, frame of reference, pre-service teachers, proportionality, graphical representation

We have studied the assessment activity of eighteen pre-service teachers – elementary school teachers and secondary school mathematics teachers – during their assessment of students' mathematics problem solving. The main objective of the observations is to identify the frame of reference of the assessor. The research design entails a computer-based situation of simulation that ensures similar experimental conditions for all teachers. The analysis carried out enable us to identify four profiles of assessment behavior as well as several assessment tendencies.

PALAVRAS-CHAVE: Atividade avaliativa, simulação, referente, professores estagiários, proporcionalidade, representação gráfica

Estudamos a actividade avaliativa de dezoito professores estagiários – professores do ensino básico (1º ciclo) e professores de matemática (ensino secundário) – confrontados com as produções dos alunos em Matemática. O principal objectivo das observações realizadas é a identificação do quadro de referência do avaliador. O dispositivo de investigação, organizado à volta de uma “situação simulada”, permite que todos os professores tenham condições experimentais análogas. As análises realizadas levam-nos a distinguir quatro perfis de comportamentos avaliativos e a relevar vários fenómenos avaliativos.

Introduction

Dans cet article, nous nous intéressons à l'activité évaluative de professeurs stagiaires confrontés à des productions d'élèves en mathématiques ; nous étudions le déroulement de leur activité en nous centrant sur sa dimension cognitive, et ce, au moyen d'observations qui visent principalement l'identification du référent de l'évaluateur, tant sur le plan de ses contenus que de ses modalités d'élaboration. L'approche retenue fait appel à la psychologie de l'évaluation scolaire (Noizet & Caverni, 1978) et à la didactique des mathématiques (Joshua & Dupin, 1993).

Actuellement, du point de vue du fonctionnement cognitif de l'évaluateur, il est courant de décrire l'activité évaluative comme une opération de confrontation entre un référent et un référé. Pour désigner cette conception, nous parlerons de « modèle référent/référé » (Hadji, 1999, pp. 43-45) :

- le référent correspond au système d'attentes jugées légitimes de l'évaluateur, système qui oriente la lecture de l'objet à évaluer ;
- le référé désigne un « modèle réduit » de l'objet évalué, construit par l'évaluateur.

Cette description est issue des travaux des docimologues expérimentaux qui ont développé « le modèle explicatif du comportement d'évaluation de productions scolaires » (Amigues, Bonniol, Caverni, Fabre & Noizet, 1975). Chez ces auteurs (*cf.* aussi Noizet & Caverni, 1978 ; Bonniol 1981, cité par Bonniol & Vial, 1997), la composante « référent » est appelée « modèle de référence » ; celui-ci fait l'objet d'une description plus fine et comporte plusieurs éléments dont « le produit norme » et « les produits attendus ». Il est postulé que toute tâche ou problème servant à évaluer conduit l'évaluateur à se représenter abstraitement un ensemble de produits possibles, dont la production à évaluer sera un élément. L'évaluateur, avant même d'être confronté aux productions à évaluer, va restreindre cet ensemble de produits possibles pour élaborer un modèle de référence opérationnel (Bonniol, 1981, cité par Bonniol & Vial, 1997). La restriction s'effectue d'abord en fonction des critères d'évaluation, plus ou moins explicités, qui permettent ainsi de définir le produit norme. Un texte de dictée sans faute ou la réponse juste à un problème sont des exemples de produits normes. L'évaluateur limite également ses attentes en prenant en

compte les données *a priori* dont il dispose et qui sont relatives au producteur du devoir à évaluer ou aux conditions de production du devoir. Les données *a priori* interviennent lors de la détermination des produits attendus, c'est-à-dire des réponses que l'évaluateur juge susceptibles d'apparaître. Le modèle de référence est susceptible de modifications au cours de l'évaluation, mais celles-ci, attribuables à la correction de copies «non attendues», n'affectent pas le produit norme, par essence très stable (Noizet & Caverni, 1978).

Des travaux des docimologues, nous retenons que «l'essentiel de l'évaluation résulte de l'activité d'un comparateur qui confronte d'un côté un modèle de référence et de l'autre côté une production (copie)»; ainsi «la tâche d'évaluation ne peut être exécutée sans que l'évaluateur dispose d'un modèle de référence, inscrit dans ses structures cognitives» (Amigues et al., 1975, p. 794). Dans ce modèle explicatif, la question des connaissances de l'évaluateur a été éludée. Nous adoptons une approche didactique et sommes très attentif aux connaissances impliquées dans le problème et dans les productions à évaluer. Il en résulte que quelques-unes des observations réalisées nous amènent à nous questionner, conjointement en raison de leur filiation et de leur proximité, sur le modèle explicatif et sur celui du référent/référé.

Nous constatons que modèle référent/référé fait l'objet d'un large consensus, tant dans le domaine de l'enseignement (par exemple, Hadji, 1999) que dans celui de la formation (par exemple, Barbier, 1985). Toutefois, à notre connaissance, le modèle explicatif, dont il est directement issu, n'a été établi que sur la base de situations expérimentales de notations et de classements de copies impliquant des professeurs spécialistes de la discipline à évaluer. De plus, pour ses concepteurs, le modèle explicatif renfermait une dimension projective et devait permettre de poser des perspectives ultérieures d'investigations. Ils avaient d'ailleurs encouragé à en étudier la validité en le testant selon les différents niveaux d'enseignement et selon le statut professionnel de l'évaluateur. L'ensemble de ces éléments nous conduit à poser deux questions concernant ces deux modèles :

- dans quelle mesure permettent-ils de décrire l'activité évaluative dans des situations autres que des situations de correction de copies ?
- dans quelle mesure permettent-ils de décrire les variations de l'activité évaluative de professeurs stagiaires ayant des formations différentes (spécialistes et non spécialistes de la discipline) ?

La première question renferme un objectif «clair»: apporter des éléments de réponses relatifs à la légitimité de considérer que ces modèles conviennent pour décrire l'activité d'évaluateurs engagés dans différentes formes d'évaluations, en particulier dans une situation aménagée d'évaluation (décrite ci-après) que l'on peut rattacher à de l'évaluation formative. La seconde prend tout son sens par rapport à la problématique du rôle des connaissances mobilisées par les professeurs stagiaires lors de l'activité évaluative. Pour y répondre, la perspective adoptée a été celle d'une comparaison qui repose sur une distinction *a priori* entre trois catégories de professeurs, même si la méthodologie retenue, essentiellement qualitative, ne permet pas d'opérer une comparaison statistique au sens strict. Faisant l'hypothèse que les rapports aux savoirs mathématiques impliqués dans la tâche d'évaluation peuvent jouer sur l'activité évaluative, nous distinguons les professeurs des lycées et collèges (PLC, enseignants du secondaire spécialistes de l'enseignement des mathématiques) des professeurs des écoles (PE, enseignants du primaire polyvalents), et parmi ces derniers, les PE «scientifiques» ayant reçu une formation que l'on peut qualifier de «scientifique», en ce sens qu'ils possèdent au minimum un baccalauréat scientifique, et les autres PE, qualifiés de PE «non scientifiques».

Nous avons mis au point une expérimentation permettant d'étudier l'activité de professeurs lors de certaines pratiques d'évaluation en mathématiques (Vantourout, 2004). Cette expérimentation permet notamment d'étudier la diversité des référents, tant sur le plan de leurs contenus que de leurs modalités d'élaboration, et de distinguer quatre «profils de comportements évaluatifs» à l'intérieur de la population expérimentale; elle apporte également des éléments de réponse aux deux questions énoncées ci-dessus et met à jour plusieurs phénomènes évaluatifs.

Présentation de l'expérimentation

Sur le plan théorique, nos références appartiennent à plusieurs domaines: la psychologie ergonomique, l'évaluation (incluant la psychologie de l'évaluation scolaire) et la didactique des mathématiques. Nous ne pouvons pas en faire état de manière exhaustive ici, mais, au fur et à mesure de la présentation de l'expérimentation, nous introduirons certaines notions empruntées à ces différents domaines.

Présentation de la situation aménagée

Pour étudier l'activité évaluative, nous recourons à une situation expérimentale que nous qualifions de «situation aménagée», dans un sens qui sera précisé ensuite. En voici la description et le déroulement¹:

- un professeur, face à un ordinateur, découvre le matériel expérimental, lequel présente un énoncé de problème et des travaux d'élèves. Sa tâche consiste à évaluer les productions et l'activité de deux binômes d'élèves «fictifs» (c'est-à-dire absents et qu'il ne connaît pas) ayant résolu un problème de proportionnalité dans lequel intervient une représentation graphique. Cette tâche est plutôt difficile, car les travaux d'élèves qui doivent être examinés ne peuvent pas faire, *a priori*, l'objet d'une évaluation immédiate et tranchée. Le professeur doit répondre à la consigne suivante: «*En quoi les travaux de ces élèves vous paraissent-ils satisfaisants et/ou non satisfaisants?*»;
- le professeur participe ensuite à un débriefing conduit par l'expérimentateur qui adapte la méthode de l'entretien d'explicitation (développée par Vermersch, 1994). Le débriefing débute par la question suivante: «*Je souhaite, si tu en es d'accord, que tu choisisses ce qui te paraît important dans le travail que tu viens de réaliser, un moment qui t'a particulièrement intéressé.*» Puis au bout d'une certaine durée d'entretien, nous utilisons une seconde consigne afin d'orienter davantage le discours vers le vécu de l'activité réalisée lors de la phase précédente: «*Si tu veux bien, je vais te poser une autre question: par quoi as-tu commencé lorsque tu t'es retrouvé(e) face à l'ordinateur?*» Nous recourons aussi à des relances et faisons une utilisation importante de questions en «comment», tout en évitant celles en «pourquoi», afin de ne pas bloquer le discours.

L'expérimentation s'organise autour du problème de mathématiques «Le cycliste»² (cf. annexe 1). Dans ce problème, il s'agit de déterminer la distance parcourue par un cycliste au bout de certaines durées. L'énoncé qui se divise en deux parties – une première «numérique» et une seconde «graphique» – comporte:

- une phrase qui présente globalement la situation – «Un cycliste roule régulièrement, s'arrête, puis roule à allure régulière.»;
- un tableau qui présente le parcours du cycliste en indiquant, en fonction de la durée écoulée (en minutes), la distance parcourue (en kilomètres);
- un graphique que les élèves doivent réaliser en plaçant les données du tableau;

- deux séries de questions :
 - d'abord, trois questions posées dans un cadre numérique: il s'agit de trouver la distance parcourue au bout de 12, 20 et 90 minutes;
 - ensuite, deux questions posées dans un cadre graphique: il s'agit de réaliser (1) le graphique; (2) de trouver la distance parcourue au bout de 25, 37, 45 et 108 minutes.

L'intérêt de ce problème réside dans le fait que la situation n'est que partiellement une situation de proportionnalité, alors que le contexte reste identique. Le parcours du cycliste se décompose en trois phases, caractérisées chacune par une vitesse moyenne et pouvant être modélisées par une fonction: une première phase (durées de 0 à 40 minutes, vitesse $V_1 = 0,25$ km/min) modélisable par une fonction linéaire; une deuxième phase (durées de 40 à 60 minutes, vitesse nulle) modélisable par une fonction constante; une troisième phase (durées de 60 à 100 minutes et plus, vitesse $V_2 = 2 \times V_1$) modélisable par une fonction affine. La question «la plus intéressante» est celle qui consiste à déterminer la distance au bout de 90 minutes. En effet, cette question «numérique» porte sur la partie «affine» ou «non proportionnelle» du problème³.

Le matériel expérimental comprend – outre l'énoncé du problème – une série de documents contenus sur un CD-ROM. Grâce à l'ordinateur, les professeurs accèdent librement à la tâche qui a été proposée à deux binômes d'élèves, aux «résultats» qu'ils fournissent (feuille commune sur laquelle chaque binôme a noté ses réponses), à l'évolution de leurs réalisations graphiques, ainsi qu'à leurs échanges verbaux (transcription des dialogues au sein d'un binôme).

Parmi les productions à évaluer, figure le travail d'un binôme d'élèves, Alexis et David. Leurs réponses aux questions 3, 4 et 5 du problème figurent en annexe 2. Pour trouver la distance parcourue en 90 minutes (question 3), David ne parvient pas à se dégager de la situation de proportionnalité: il transforme le tableau de l'énoncé en conservant pour toute la durée du parcours la situation de proportionnalité. De son côté, Alexis utilise implicitement et judicieusement la propriété des fonctions affines dite «propriété des écarts proportionnels». Lorsque ces élèves passent à la représentation graphique (question 4), ils utilisent les données figurant dans le tableau de l'énoncé et aboutissent au graphique attendu. En observant le graphique, David se rend compte de la justesse de la réponse d'Alexis et barre alors sa propre réponse. Ils répondent ensuite à la question 5 à partir du graphique.

Dans la suite de l'article, quand nous ferons appel à des citations de professeurs qui mentionnent des élèves, il s'agira toujours de ce binôme. En effet, ce sont les productions qui leur sont attribuées qui nous ont permis de réaliser les observations les plus intéressantes pour cet article. Bien évidemment, les professeurs ont également discuté du travail de l'autre binôme d'élèves, mais dans des proportions moindres.

L'utilisation d'un tel dispositif – ordinateur et CD-ROM – se justifie essentiellement par les données précises qu'il procure sur le déroulement de l'activité évaluative. En effet, à chaque utilisation, il se crée un fichier électronique – le «journal des documents affichés» – qui enregistre nominativement les documents affichés, ainsi que leurs durées et ordre d'affichages.

La situation retenue est «aménagée» au regard de la distance qui la sépare de la situation «de référence», c'est-à-dire celle que l'on cherche à reproduire dans l'expérimentation (le terme «référence» relève ici de la psychologie ergonomique qui en donne un sens différent de celui qu'il reçoit en évaluation). L'élaboration de la situation aménagée nous amène, en tant que concepteur, à réduire la complexité de la situation de référence dans des proportions que nous estimons «raisonnables», au sens où, du point de vue du fonctionnement cognitif des évaluateurs, nous faisons l'hypothèse qu'il existe des similarités importantes entre les deux situations.

Nous distinguons deux types d'aménagements (surtout pour en faciliter la présentation, car ils sont en réalité liés) lors du passage de la situation de référence à la situation aménagée: ceux relatifs au contexte dans lequel se déroule l'activité évaluative et ceux qui renvoient à l'élaboration du matériel expérimental, en particulier des productions d'élèves à évaluer.

Aménagements du contexte

Les aménagements du contexte sont nombreux; nous les analysons à l'aide de notions développées en psychologie ergonomique⁴ et en évaluation.

La situation de référence correspond à «une situation d'évaluation formative continue avec régulation interactive» (Allal, 1998, p. 58): le professeur, ayant préparé sa séance et choisi la tâche proposée aux élèves, circule parmi ses élèves; il les observe en train de résoudre un problème, afin de comprendre ce qu'ils font puis de réguler si cela lui semble nécessaire. Il est confronté à une tâche d'évaluation complexe et discrétionnaire (c'est-à-dire, seulement définie par son but) devant être réalisée dans un environnement dynamique, interactif, humain et ouvert.

Dans la situation aménagée, les professeurs n'ont pas préparé la séance, ni choisi le problème à l'origine des productions à évaluer. Celles-ci sont attribuées à des élèves fictifs, représentés uniquement par des traces de leur activité. La tâche, malgré ces réductions, demeure complexe et discrétionnaire; l'environnement reste ouvert, mais il devient statique et passif. Ce contexte expérimental – en particulier, le caractère fictif des élèves – ne permet pas vraiment aux professeurs de s'engager dans une régulation de nature pédagogique, classiquement associée à l'évaluation formative (Scallon, 2000). En d'autres termes, la tâche à laquelle les professeurs sont confrontés est une tâche d'observation qui ne représente qu'une partie des étapes du processus d'évaluation, mais qui est considérée comme fondamentale lors des évaluations formatives. Toutefois, la situation d'évaluation demeure très riche: en effet, comme nous l'avons déjà indiqué, les travaux d'élèves à examiner ne peuvent pas faire, *a priori*, l'objet d'une évaluation immédiate et tranchée (ce qui est souvent le cas avec un problème comportant un graphique).

La notion d'activité est une autre notion que nous empruntons à la psychologie ergonomique. En nous référant à cette discipline, nous considérons que l'activité évaluative renferme, comme toute activité, une dimension intentionnelle. Certains de nos choix reposent d'ailleurs sur la dimension intentionnelle de l'activité. Par exemple, proposer une tâche discrétionnaire, afin que les professeurs puissent se fixer des buts et des sous-buts lorsqu'ils «se redéfinissent leur tâche» (Rogalski, 2003) et se donnent ou se représentent la tâche qui, de leur point de vue, doit être la leur.

Élaboration des productions à évaluer et «tâche attendue»

La didactique des mathématiques nous a servi à élaborer le matériel expérimental, en particulier les productions à évaluer (utilisation de connaissances relatives à la proportionnalité et aux graphiques et, plus spécialement, réalisation d'études préalables ayant pour objet la résolution du problème «Cycliste»). C'est également à elle que nous nous référons lorsque nous considérons l'erreur comme un indicateur du fonctionnement d'une connaissance et que, en conséquence, nous nous attachons, à partir des observables, à identifier et analyser les procédures et leur signification sur le plan des connaissances.

Si nous considérons la finalité de l'évaluation dans l'expérimentation, nous attendons des professeurs qu'ils s'engagent dans une analyse didactique des travaux d'élèves, laquelle montrerait ce qui est pertinent et ce qui ne l'est pas dans leurs réalisations. Pour satisfaire à la tâche, l'activité des professeurs devrait consister, pour une part non négligeable, à tenter d'analyser les

procédures des élèves. Mais «en amont», si nous considérons le modèle référent/référé – si nous nous plaçons à un niveau d'attente intermédiaire qui correspondrait à un sous-but chez l'évaluateur –, nous attendons des professeurs qu'ils élaborent leur référent. Pour y parvenir, leur activité devrait consister d'abord à résoudre le problème afin de disposer des réponses aux questions, mais aussi à envisager les réponses attendues et les procédures susceptibles d'apparaître chez les élèves. Le produit norme serait constitué des réponses aux questions, tandis que les réponses attendues comprendraient le produit norme et les procédures des élèves.

Approche comparative

Le choix d'une situation aménagée permet de proposer une tâche identique à différents professeurs et autorise diverses comparaisons entre les catégories de professeurs, en particulier entre des PLC enseignants de mathématiques, spécialisés dans cette discipline, et des PE, non spécialisés dans l'enseignement des mathématiques.

Le problème «Cycliste» peut être proposé aussi bien à des élèves de CM2 (10-11 ans, étudiants au primaire) qu'à des élèves de 6^e (11-12 ans, étudiants au collège). Lors de l'expérimentation, les élèves sont déclarés être en classe de CM2, si l'enseignant est un PE, ou en classe de 6^e, s'il s'agit d'un PLC.

Participants

Les participants sont dix-huit professeurs stagiaires en fin de formation initiale dans un Institut Universitaire de Formation des Maîtres⁵. Comme nous l'avons indiqué, nous les répartissons en trois catégories: 6 PLC (repérés par des codes allant de Pr1 à Pr6), 4 PE scientifiques (Pr7 à Pr10) et 8 PE non scientifiques (Pr11 à Pr18). Parfois, nous parlerons de professeurs «scientifiques» lorsque nous regroupons les PLC et les PE scientifiques.

Résultats

Consultation de l'énoncé et résolution du problème par les professeurs

Pour répondre à la tâche attendue en début d'activité – résoudre le problème afin d'élaborer un référent –, il semble nécessaire pour les professeurs de consulter l'énoncé et de travailler à partir de celui-ci. Les journaux des documents affichés nous renseignent sur les moments et les durées d'affichage de l'énoncé; ils nous permettent d'obtenir un premier résultat relatif à la

consultation de l'énoncé. En ce qui concerne le travail effectué à partir de cet énoncé, et plus spécialement l'état de résolution du problème (quelles questions ont été traitées?), nous obtenons une première série de résultats grâce à l'analyse des notes figurant sur les feuilles de brouillons dont les professeurs disposent quand ils travaillent dans la situation aménagée.

Ces résultats, essentiellement factuels, sont regroupés dans le tableau 1. Leur analyse nous apporte plusieurs constats :

1. Le temps alloué par les professeurs à l'énoncé et au problème est extrêmement variable, en particulier en début d'évaluation; nous relevons des durées d'affichages allant d'un peu plus de 2 minutes (Pr1, Pr2, Pr17, Pr18) à environ 10 minutes (Pr4, Pr10), voire à près de 30 minutes (Pr13); par ailleurs, les journaux des documents affichés nous apprennent que tous les professeurs amorcent leur activité en consultant l'énoncé, la différence essentielle étant que certains vont s'y attarder, tandis que d'autres vont le quitter rapidement pour s'intéresser à autre chose.
2. Onze professeurs cherchent à résoudre le problème avant d'évaluer les productions des élèves: deux PLC (Pr3, Pr4), trois PE scientifiques (Pr8, Pr9, Pr10) et six PE non scientifiques (Pr11, Pr13, Pr14, Pr16, Pr17, Pr18). Ils appartiennent aux trois catégories que nous distinguons *a priori*. Quatre des cinq scientifiques (Pr3, Pr4, Pr8, Pr9) disposent de la totalité des réponses. En construisant le graphique (parfois il s'agit d'un graphique à main levée, d'une «silhouette») seuls Pr4, Pr8 et Pr9 résolvent totalement le problème. Enfin, un seul PE non scientifique, Pr11, parvient à répondre à l'ensemble des questions.
3. Parmi les sept qui ne résolvent pas le problème, on retrouve des professeurs appartenant aux trois catégories: quatre PLC (Pr1, Pr2, Pr5, Pr6), un PE scientifique (Pr7) et 2 PE non scientifiques (Pr12, Pr15).

Ces résultats témoignent de différences individuelles très sensibles. En revanche, ils ne permettent pas vraiment de clairement discriminer les catégories de professeurs.

Tableau 1
**Durées d'affichage de l'énoncé et résolution du problème
 par les professeurs**

Catégories de professeurs		PLC						PE scientifiques			
Professeurs		Pr1	Pr2	Pr3	Pr4	Pr5	Pr6	Pr7	Pr8	Pr9	Pr10
Affichage de l'énoncé :											
•Durée d'affichage en début d'évaluation		2 min 10	2 min 30	8 min	8 min 40	4 min 20	3 min	3 min 30	6 min 30	4 min 15	11 min 20
•Durée totale d'affichage lors de l'évaluation		5 min	3 min	12 min 20	9 min 20	5 min	4 min 15	3 min 50	18 min 30	9 min	12 min
•% de la durée totale de l'évaluation		19%	11%	28%	21%	14%	11%	7%	34%	18%	25%
Résolution complète du problème		non	non	non	oui	non	non	non	oui	oui	non
Réponses aux questions du problème	12 et 20 min (questions numériques linéaires)			oui	oui				oui	oui	oui
	90 min (question numérique affine)			oui	oui				oui	oui	oui
	25, 37 et 45 min (questions graphiques)			oui	oui				oui	oui	
	108 min. (question graphique affine)			oui	oui				oui	oui	
	Graphique attendu				oui***				oui***	oui	oui***

Catégories de professeurs		PE non scientifiques								
Professeurs		Pr11	Pr12	Pr13	Pr14	Pr15	Pr16	Pr17	Pr18	
Affichage de l'énoncé :										
•Durée d'affichage en début d'évaluation		7 min 20	3 min	29 min 20	7 min	5 min 30	7 min 30	2 min 10	5 min 45	
•Durée totale d'affichage lors de l'évaluation		7 min 20	3 min 50	30 min 40	8 min 10	5 min 30	8 min 20	3 min 20	8 min 15	
•% de la durée totale de l'évaluation		3%	7,5%	47%	19%	13%	18%	7%	23%	
Résolution complète du problème		non	non	non	non	non	non	non	non	
Réponses aux questions du problème	12 et 20 min (questions numériques linéaires)	oui		oui	oui		oui	oui	oui	
	90 min (question numérique affine)	oui		oui*	oui		oui	oui*		
	25, 37 et 45 min (questions graphiques)	oui		oui**						
	108 min (question graphique affine)	oui								
	Graphique attendu						oui***			

Note : oui* signifie que deux réponses (dont une fausse) ont été inscrites et conservées ;
 oui** signifie qu'une réponse fausse à la question «pour 45 minute» a été inscrite et conservée ;
 oui*** signifie que seule une «silhouette» du graphique a été construite.

Analyse des contenus et des modalités d'élaboration des référents des évaluateurs

L'étude des référents, de leurs contenus et de leurs modalités d'élaboration est imbriquée dans l'analyse du déroulement de l'activité évaluative de chaque professeur. Les résultats obtenus proviennent d'une analyse des débriefings (ceux-ci ayant été transcrits) menés auprès des dix-huit professeurs, et de la conjugaison de cette analyse avec les résultats précédents. Cette étude nous a permis de répartir les participants, de manière satisfaisante, selon quatre profils de comportements évaluatifs que nous allons détailler.

Les éléments contenus dans le tableau 2 donnent une première idée de la répartition et devraient aider à suivre la présentation des résultats. Pour déterminer chaque profil, nous nous appuyons sur deux critères. L'engagement (ou le non-engagement) en début d'évaluation dans la résolution du problème (ce choix étant lié à la dimension intentionnelle de l'activité) représente un premier critère. Il nous amène à distinguer, d'un côté, les 1^{er} et 2^e profils, de l'autre, les 3^e et 4^e profils. Pour différencier le 1^{er} profil du 2^e profil, nous utilisons un deuxième critère : nous relevons si les professeurs disposent ou non de l'ensemble des réponses lorsqu'ils découvrent les productions des élèves. Dans l'interprétation de ces profils, nous tenons compte aussi d'éléments montrant le degré de maîtrise et de compréhension de la situation par les professeurs. Ces éléments proviennent des feuilles de brouillons (trace des procédures de résolution) et de l'entretien d'explicitation. Pour distinguer le 3^e profil du 4^e profil, nous tenons uniquement compte de la maîtrise et de la compréhension de la situation par les professeurs (ils ne disposent d'aucune réponse) et faisons appel à la notion de «structure de la solution» (*cf.* ci-après), telle que mise en évidence lors de l'entretien d'explicitation. D'un point de vue pratique, pour étudier les référents, nous relevons les éléments qui semblent «incontournables» selon le modèle référent/référé; il s'agit principalement des réponses aux questions et de la courbe représentant le parcours du cycliste. C'est en portant une grande attention aux éléments qui touchent à la compréhension du problème et à la possibilité qu'ont les professeurs d'exercer un contrôle sur la justesse des réponses qu'ils détiennent que nous inscrivons véritablement notre étude dans une orientation didactique, centrée sur le fonctionnement cognitif et sur les connaissances mathématiques des professeurs (pour plus de détails, *cf.* Vantourout & Maury, 2006).

Tableau 2
*Répartition des professeurs selon les quatre profils
 de comportements évaluatifs*

		<i>1^{er} profil</i>	<i>2^e profil</i>	<i>3^e profil</i>	<i>4^e profil</i>
		Professeurs qui s'engagent dans la résolution du problème		Professeurs qui ne s'engagent pas dans la résolution du problème	
		et qui disposent de l'ensemble des réponses	et qui ne disposent pas de l'ensemble des réponses	mais qui disposent de «la structure de la solution»	et qui ne disposent pas d'une bonne compréhension du problème
Catégories de professeurs	PLC	Pr3 Pr4		Pr1 Pr2 Pr5 Pr6	
	PE scientifiques	Pr8 Pr9	Pr10*	Pr7	
	PE non scientif.	Pr11*	Pr13 Pr14 Pr16 Pr17 Pr18		Pr12 Pr15

Note: Pr10* et Pr11* sont des cas «intermédiaires» entre les profils 1 et 2.

La répartition portée par le tableau 2 nous permet déjà de souligner que la distinction des quatre profils de comportements évaluatifs constitue un résultat qui, contrairement aux précédents, permet de discriminer les catégories de professeurs. Nous allons voir lors de la présentation des profils que, si nous ne tenons pas compte des deux cas «intermédiaires», nous trouvons, d'un côté, les PLC et les PE scientifiques qui relèvent tous du 1^{er} ou du 3^e profil, et de l'autre, les PE non scientifiques qui relèvent tous du 2^e ou du 4^e profil.

1^{er} profil de comportement évaluatif

Il apparaît chez quatre professeurs scientifiques : deux PLC, Pr3 et Pr4, et deux PE scientifiques, Pr8 et Pr9. Il correspond à l'activité de professeurs qui résolvent le problème, disposent de l'ensemble des réponses et exercent un contrôle sur leur validité. Ces professeurs sont les seuls (avec Pr11) à disposer de la réponse attendue à chaque question lorsqu'ils évaluent les productions des élèves (*cf.* tableau 1), soit une assez faible proportion de la population. Nous pouvons les caractériser comme étant ceux qui allouent un temps suffisant à la résolution du problème en début d'évaluation ou, si nécessaire, en cours d'évaluation (*cf.* tableau 1). Ils déterminent les équations des droites pour les phases 1 et 3 du parcours et effectuent, à un moment donné, une résolution algébrique. Chez tous, c'est le besoin de connaître les réponses exactes, pour apprécier la précision avec laquelle les élèves lisent leur graphique, qui est à l'origine du choix de ce mode de résolution. Ils

disposent aussi d'une bonne image du graphique et trois d'entre eux (Pr4, Pr8, Pr9) en font une construction plus ou moins précise. Il existe des différences, liées notamment au moment où ils recourent à l'algèbre. Certains (Pr8, Pr9) résolvent immédiatement, de cette façon, l'intégralité du problème, alors que d'autres (Pr3, Pr4) vont attendre et le faire seulement plus tard, lorsqu'ils jugent que cela est vraiment utile pour vérifier la précision des réponses graphiques. Dans un premier temps, ils utilisent les valeurs du tableau de l'énoncé pour répondre aux questions numériques et évaluent les réponses graphiques en se servant seulement du graphique des élèves.

Quel référent pour l'évaluateur relevant du 1^{er} profil ? – Ces professeurs disposent d'un référent contenant toutes les réponses aux questions. Certains référents sont éventuellement plus complets que d'autres, en particulier pour ce qui touche à la partie graphique. Les procédures algébriques ne correspondent pas aux procédures disponibles chez des élèves de CM2 ou de 6^e ; elles ne devraient pas tenir lieu de réponses attendues. En revanche, elles attestent de la maîtrise du problème et permettent d'exercer un contrôle sur la validité de l'ensemble des réponses. Le déroulement de l'activité de Pr3 et de Pr4 révèle un phénomène évaluatif très intéressant : tout semble se passer comme s'ils hésitaient à employer l'algèbre et avaient la possibilité de recourir à différentes modalités de résolution, plus ou moins économiques, pour trouver les réponses. Cette possibilité est, nous le pensons, fortement liée à la possession de connaissances mathématiques solides (Vantourout & Maury, 2006). Lorsqu'ils accèdent aux productions des élèves, tous mettent en œuvre une activité qui correspond à une confrontation entre un référent et un référé. D'ailleurs, plusieurs déclarent comparer leurs réponses avec celles des élèves. Bref, chez ces professeurs, l'activité évaluative peut tout à fait être décrite à l'aide du modèle référent/référé.

Cas « intermédiaires »

Pr10 et Pr11 sont les cas intermédiaires entre les 1^{er} et 2^e profils. Pr10 est un PE scientifique qui ne résout que la partie numérique du problème et qui ne dispose donc pas de l'ensemble des réponses. Il mobilise cependant des procédures lui permettant de contrôler à la fois ses propres réponses et la validité des réponses graphiques des élèves (il s'appuie sur le calcul des vitesses auxquelles roule le cycliste). Sans entrer dans les détails, sur le plan des modalités d'élaboration du référent, il est très proche des professeurs du 1^{er} profil.

Pr11 est le seul PE non scientifique qui dispose de la totalité des réponses ; sur le plan de la résolution du problème, il est le plus performant de sa catégorie. En lisant le tableau, il reconnaît immédiatement la présence de parties proportionnelle et non proportionnelle ; il trouve cela plutôt difficile et il s'oblige à résoudre le problème. Pr11 est « intermédiaire » car ses réponses sont parfois peu assurées. Pour la question numérique et affine (90 minutes), il dit trouver « *le même résultat qu'Alexis et David* »⁶ (30 kilomètres), lequel lui semble « *intuitivement* » juste. Il reconnaît son incapacité à se justifier. Il utilise en fait, implicitement, la propriété dite « des accroissements proportionnels »⁷, comme le font tous les PE non scientifiques qui trouvent cette réponse. Pour les questions graphiques, il obtient ses réponses uniquement en lisant le graphique construit par Alexis et David. Brièvement, nous dirons que l'activité évaluative de Pr11 correspond à une confrontation entre un référent et un référé. Cependant, nous pensons que le contenu de son référent diffère largement de celui des professeurs du 1^{er} profil : même si toutes les réponses y figurent, sa maîtrise du problème est moindre et il ne dispose d'aucun moyen pour contrôler la validité de certaines réponses.

2^e profil de comportement évaluatif

Il apparaît chez cinq des huit PE non scientifiques : Pr13, Pr14, Pr16, Pr17 et Pr18. Il correspond à l'activité de professeurs qui s'engagent dans une résolution partielle du problème et qui, ensuite, ne disposent pas de l'ensemble des réponses quand ils découvrent les productions à évaluer. Les durées allouées à cette résolution sont extrêmement variables (*cf.* tableau 1). Parmi les dix-huit professeurs, Pr13 est celui qui y consacre le plus de temps, tandis que Pr17 est l'un de ceux qui y consacre le moins de temps. Par ailleurs, ces professeurs ne sont pas toujours en mesure d'exercer un contrôle sur la validité de leurs réponses. Quant à l'intention, et exclusivement sur ce plan, en cherchant à avoir au préalable les réponses, ces professeurs sont relativement proches du 1^{er} profil.

Ils trouvent tous les réponses aux questions numériques et linéaires (12 et 20 minutes, qui sont les plus faciles), puis ils cherchent à répondre à la question numérique et affine (90 minutes). La façon de répondre est ici un excellent indice de la compréhension du problème. Pr14 et Pr16 trouvent la réponse attendue (30 kilomètres) en recourant à la procédure « accroissements proportionnels ». Quant à la maîtrise du problème, ces deux professeurs sont très proches de Pr11 (*cf.* 1^{er} profil). Pr18 marque un point d'interrogation comme seule réponse à cette question et il ne l'évalue jamais. Sa compréhension

du problème, bien que parfois lacunaire, est correcte : outre la distinction des trois phases, il relève que la vitesse double entre les phases 1 et 3. Pr13 et Pr17 trouvent deux réponses, la première étant juste (30 kilomètres), la seconde (22,5 kilomètres) étant fautive. Pr13 est incapable de choisir celle qu'il doit conserver. Cette situation découle de difficultés de compréhension majeures qu'il reconnaît. À aucun moment, il ne s'aperçoit de la modification de la vitesse. D'après ses calculs, elle reste identique (0,25 km/min) durant les phases 1 et 3. La réponse erronée découle de l'application d'une procédure « fonction » ($90 \text{ minutes} \times 0,25 \text{ km/min}$), non valide ici car liée à la proportionnalité. Pr17 se retrouve dans la même position, cependant, en lisant le dialogue entre Alexis et David, il comprend enfin que la situation n'est que partiellement une situation de proportionnalité et qu'il doit rejeter la réponse erronée, obtenue à l'aide d'une procédure « produit en croix ».

Voyons, à partir d'illustrations, comment la partie graphique est traitée. Pr16 est le seul PE non scientifique qui trace la courbe attendue (à main levée). Pr18, en revanche, n'a aucune idée de la forme qu'elle doit prendre, jusqu'au moment où il voit celle d'Alexis et David. Ensuite, il dit se baser sur leur travail, car il sait que ces élèves « *sont dans le vrai* ». Toutefois, il ne contrôle jamais la justesse des réponses, d'après lui, « *ce qui est primordial, c'est la démarche qui a été établie* ». Pr17 déclare à plusieurs reprises « *faire confiance aux élèves* » et s'appuyer sur leurs réponses. Pr13 ne se pose pas la question de savoir comment est la courbe ; en revanche, il est seul à disposer de réponses à deux questions graphiques, obtenues par le calcul (questions « 25 et 37 minutes » qui sont « linéaires »). Il continue de faire comme si l'intégralité du parcours du cycliste se résumait à une situation de proportionnalité, si bien qu'il trouve une réponse fautive et surprenante à la question « 45 minutes », pourtant située pendant la pause et dont la réponse semble « évidente ». Il reprend pour les autres questions les réponses d'Alexis et David.

Quel référent pour l'évaluateur relevant du 2^e profil ? – Tous ces professeurs disposent d'un référent qui est plus ou moins complet. Pour plusieurs d'entre eux (Pr13, Pr17, Pr18), il ne correspond pas vraiment à l'image que l'on peut couramment se faire d'un tel objet : il ne comporte qu'une partie des réponses, avec l'éventualité pour certaines d'être erronées, et le graphique, voire seulement l'idée de la forme de la courbe, en sont souvent exclus. Pour d'autres (Pr14, Pr16), nous estimons que le référent est plutôt proche de celui de Pr11 (cf. 1^{er} profil). Sur le plan des modalités d'élaboration du référent, nous relevons la manifestation d'un phénomène évaluatif assez général qui apparaît aussi avec les 3^e et 4^e profils et qui consiste à

«s'appuyer sur les productions des élèves» pour comprendre la situation ou obtenir certaines réponses. Dans le cas présent, il s'accompagne souvent d'une incapacité à exercer un contrôle sur les réponses qu'il permet d'obtenir et est à l'origine de modifications qui portent, notamment, sur les réponses erronées des professeurs. Tout ce qui vient d'être évoqué dépend de la maîtrise du problème; l'hétérogénéité des niveaux de compréhension génère des différences importantes entre ces professeurs dès que le problème se complique.

Lorsque ces professeurs accèdent aux productions des élèves et qu'ils disposent des réponses, ils mettent en œuvre une activité qui correspond à une confrontation entre un référent et un référé. Cette description reste acceptable lorsqu'ils n'ont qu'une réponse fautive, bien que nous émettions des réserves quant à la pertinence de l'évaluation qui en résulte. Nous nous interrogeons sur la possibilité d'une telle description quand ils évaluent une question alors qu'ils n'en connaissent pas la réponse, ou pour laquelle ils se fient aux réponses des élèves. Nous reviendrons sur ces interrogations lors de la présentation du 4^e profil.

3^e profil de comportement évaluatif

Il apparaît chez cinq professeurs scientifiques, quatre PLC, Pr1, Pr2, Pr5 et Pr6, et un PE scientifique, Pr7. Il correspond à l'activité de professeurs qui ne s'engagent pas dans la résolution du problème mais qui posséderaient, selon nous, «la structure de la solution». D'après leurs dires, ces professeurs ne cherchent jamais à résoudre le problème, ce qui fait qu'ils ne disposent d'aucune réponse lorsqu'ils découvrent les productions des élèves. Cette manière de faire correspond à un véritable choix, en accord avec leur intention de suivre les raisonnements et de comprendre les façons de procéder.

Si ces professeurs ne résolvent pas le problème, que font-ils avec l'énoncé et que retirent-ils de sa lecture? Globalement, ils développent un mode de relation à l'énoncé qui se caractérise par la présence de plusieurs phénomènes évaluatifs (décrits ci-après), lesquels proviendraient de relations réciproques entre leur façon de lire l'énoncé et le fait de disposer de la structure de la solution. Cette expression désigne pour nous, qui nous référons à Mendelsohn⁸ (1981), un modèle général de résolution pour une classe de problèmes donnée. Ici, il s'agit de la classe dont relève le problème «Cycliste» et le rôle de cette structure serait de faciliter une lecture et une compréhension plutôt globales du problème. Ce dernier serait reconnu comme une situation impliquant les notions de fonctions linéaire et affine, une courbe affine par morceaux, etc.

Ces professeurs lisent l'énoncé et admettent le faire trop rapidement (cf. tableau 1); certains ne seraient pas attentifs aux valeurs numériques (Pr2, Pr5). En fait, bien qu'ils laissent des éléments de côté, ils relèvent la plupart des éléments clés nécessaires à la compréhension du problème. Pr2, par exemple, repère dans le tableau le moment de la pause et distingue grossièrement les phases du parcours; il reconnaît prêter attention aux questions et *«comprendre où on veut en venir: d'abord faire travailler les élèves sur une situation de proportionnalité puis affine, ensuite leur faire tracer la courbe, puis s'en servir simplement pour du repérage»*. En citant l'un d'eux (Pr7), nous dirons que tout se passe un peu comme s'ils arrivaient *«à prendre de l'information sans réfléchir, à voir ce qui est important tout de suite»*. Enfin, à l'exception de Pr1 qui survole l'énoncé et imagine un autre problème (cf. ci-après), ils ont en tête une image graphique adéquate et plus ou moins précise.

Ce mode d'appréhension peut conduire également à une compréhension partielle de la situation, voire inadéquate (cf. Pr5, Pr1 ci-après). Cependant, cela reste toujours provisoire, d'une part, parce qu'ils s'appuient sur les productions des élèves, et, d'autre part, parce qu'ils ont conscience de leur position vis-à-vis de l'énoncé. Pr5 a une approche large du problème. Il dit voir ce qu'est la situation, mais aussi ne distinguer ni le moment de la pause, ni la modification de la vitesse. C'est en lisant les dialogues qu'il repère les éléments qui lui ont auparavant échappé. Il en va de même pour Pr1. Celui-ci ne comprend pas au départ que le cycliste ne s'arrête qu'une seule fois et croit qu'il peut *«le faire 50 000 fois»*. Mais, il se rend compte rapidement qu'il se trompe car cette situation est *«trop compliquée pour des 6^e»*, surtout la réalisation du graphique correspondant à cet autre problème. La structure de la solution jouerait ici un rôle essentiel: sa prégnance constitue une piste plausible pour expliquer ces rétablissements rapides et aisés. La façon dont ces professeurs utilisent les productions des élèves est un élément majeur pour décrire leur activité. Ils s'en servent pour suivre et comprendre la progression des élèves, parfois pour améliorer leur compréhension de la situation, mais jamais pour obtenir les réponses. Ils disent avoir l'impression *«de faire le problème en même temps que les élèves»* (Pr2) et *«d'essayer d'avancer avec les élèves»* (Pr7). Bref, l'utilisation des productions des élèves, en raison de l'attitude critique développée vis-à-vis de celles-ci, est totalement différente de ce qui se produit avec les 2^e et 4^e profils.

Ces professeurs quittent très rapidement l'énoncé pour aller lire les dialogues qui vont les aider à comprendre les raisonnements des élèves. Auparavant, trois d'entre eux (Pr1, Pr5, Pr7) recopient le tableau afin de

mieux suivre les dialogues dans lesquels les élèves font souvent référence aux nombres qu'il contient. Ils disent être «*sans idée préconçue*» (Pr2), voire dans une situation de découverte «*en aveugle*» (Pr1). Cependant, cette autoanalyse doit être nuancée car nous avons relevé chez la plupart d'entre eux des attentes vis-à-vis des productions des élèves, mais celles-ci restent assez vagues et ne s'expriment jamais sous la forme de réponses numériques. Pr7, par exemple, pense que les élèves ont vu que la distance parcourue restait la même entre 40 et 50 minutes, et ce qui l'intéresse c'est de voir comment ils repèrent les trois phases du parcours dans le tableau. Pr2 sait qu'il va y avoir d'abord des calculs de proportionnalité, mais sans pouvoir dire quelles sont les données du tableau et les procédures qui vont être utilisées. Pour ce qui est de l'évaluation des réponses proprement dites, nous avons relevé quelques différences. Pr1, après avoir lu les dialogues, retourne sur l'énoncé pour vérifier l'exactitude des résultats des élèves, sans préciser comment il procède. Pr5 voit en lisant que «*les élèves répondent bien*», mais ensuite, en utilisant les nombres du tableau, il fait des calculs pour contrôler les réponses graphiques. Pr6 et Pr7 «*font confiance aux élèves*» pour leurs réponses numériques et graphiques.

Pour terminer, un dernier phénomène évaluatif exclusif à ce 3^e profil. Certains (Pr2, Pr6 et Pr7) s'interrogent au cours de leur activité pour savoir si les bornes de la pause figurent parmi les valeurs du tableau. Ce questionnement n'est pas banal car l'absence de l'une de ces bornes complique énormément le problème. Nous croyons qu'il ne peut pas renvoyer uniquement à une lecture «partielle» de l'énoncé, mais qu'il dépend aussi de leur connaissance du rôle joué par ces bornes, c'est-à-dire qu'il serait en relation avec la structure de la solution.

Quel référent pour l'évaluateur relevant du 3^e profil? – Nous considérons que la structure de la solution tient lieu de référent pour ces professeurs. Cependant, elle ne serait pas mobilisée ici pour résoudre le problème, mais pour en permettre une lecture et une compréhension globale. Ce référent renfermerait les principales notions mathématiques impliquées dans le problème, une image adéquate de la courbe, ainsi que des attentes plus ou moins précises, comme le recours à certaines procédures. En revanche, aucune réponse précise aux questions n'étant formulée, il n'y aurait donc pas de «véritable» produit norme. La description de l'activité évaluative de ces professeurs ne renvoie pas à la comparaison référent/référé sous sa forme la plus «classique» (c'est-à-dire telle qu'elle apparaît avec le 1^{er} profil). Cela considéré, nous pensons qu'il s'agit d'une opération de confrontation impliquant, d'un côté, une lecture et une analyse attentives des raisonnements

des élèves, parfois de leurs résultats (nous pensons aux deux professeurs qui effectuent cette vérification à un moment), et de l'autre, un cadre de possibles, à la fois ouvert à beaucoup d'éventualités et limité par la structure de la solution. Bref, en nous référant à Hadji (1999, pp. 43-45), nous pensons que cette activité évaluative peut être décrite comme une lecture orientée des productions des élèves, basée sur le système d'attentes de l'évaluateur jugées légitimes.

4^e profil de comportement évaluatif

Il apparaît chez deux PE non scientifiques, Pr12 et Pr15. Il correspond à l'activité de professeurs qui ne cherchent pas à résoudre le problème et qui ne disposent d'aucune réponse lorsqu'ils découvrent ensuite les productions des élèves. Ces professeurs amorcent leur activité en lisant brièvement l'énoncé (cf. tableau 1), puis ils recopient le tableau et les trois premières questions pour pouvoir suivre les dialogues. Nous devons nous garder de conclure hâtivement à une proximité entre ce profil et le 3^e, car il ne s'agit que de ressemblances de «surface» : en effet, ces professeurs sont loin de maîtriser le problème, au contraire ils rencontrent des difficultés importantes pour le comprendre.

Au début de son activité, Pr15 décide de ne pas résoudre lui-même le problème car il pense que cela lui prendrait trop de temps et qu'il le fera «*automatiquement après, en évaluant le travail des enfants*». Il essaye alors seulement de le faire «*vaguement et dans sa tête*», mais il se fabrique une idée totalement fautive de la situation : il ne voit pas la phrase décrivant les trois phases dans le parcours puis, lorsqu'il lit le tableau, il croit qu'aucune des phases ne correspond à une situation de proportionnalité, que «*le cycliste n'avait jamais un rythme régulier*». De là, il imagine une courbe qui «*n'est pas une droite, mais une courbe ascendante qui à la fin monte de plus en plus*». Comme d'autres, il se rend compte de son erreur en lisant le dialogue entre Alexis et David où est mentionnée la proportionnalité. Il reprend le tableau, repère alors les trois phases, et comprend que la dernière est une situation de non-proportionnalité. Il a désormais une image adéquate de la courbe attendue. Lors du débriefing, il admet que d'avoir une idée fautive de la réponse attendue est «*un petit peu gênant*».

Quant à Pr12, on peut s'interroger sur les raisons qui le conduisent à ne pas résoudre le problème. Serait-ce parce qu'il trouve que «*l'énoncé est clair*» et qu'«*il n'y a pas de piège, rien de subtil pour les enfants*»? La phrase qui décrit les phases du parcours est le point qui lui paraît le plus important ; il la

retient, bien avant le tableau. Quant à sa compréhension de la situation, il voit « deux sortes de proportionnalités qui se suivent, coupées par une pause » et insiste à plusieurs reprises sur le caractère proportionnel de la troisième phase du parcours. Par ailleurs, il oublie de regarder la partie graphique du problème.

Ces professeurs se mettent très rapidement à lire les dialogues. Ils disent s'intéresser avant toute chose à l'engagement des élèves dans leur travail et au fonctionnement social du groupe ; Pr12, par exemple, « essaye de voir s'il y a un enfant qui est prépondérant dans la réalisation de l'exercice ». Le tableau leur permet de suivre la manière dont les élèves procèdent pour répondre aux deux premières questions (12 et 20 minutes). Hormis pour ces deux questions, ils ne donnent pas l'impression de s'intéresser aux réponses proprement dites ; aucune autre réponse n'est évaluée. Ils regardent en revanche si les points du tableau ont été correctement placés sur le graphique. Ils s'appuient abondamment sur les productions des élèves pour conduire leur évaluation. Mais, contrairement aux professeurs relevant du 3^e profil, ils n'exercent aucune attitude critique vis-à-vis de celles-ci, et sembleraient même plutôt en dépendre totalement. Ce phénomène est visible chez Pr15 qui choisit d'évaluer les productions « finales » des élèves, après avoir suivi les étapes de la construction de leurs réponses, parce que « là, [il] pense maîtriser complètement l'exercice et ce qu'il faut faire ».

Quel référent pour l'évaluateur relevant du 4^e profil ? – Il nous semble assez difficile d'attribuer un référent – au sens où on l'entend habituellement – à ces professeurs. D'abord, parce qu'ils ne parviennent pas, ou très difficilement, à modéliser correctement la situation et parce qu'ils ne disposent d'aucune réponse. Ensuite, parce que le phénomène évaluatif, très répandu, qui consiste à s'appuyer sur les productions des élèves intervient ici dans des proportions que nous jugeons « extrêmes ». En effet, il ne s'agit plus de « s'appuyer », mais davantage de puiser de l'information, voire des connaissances (par exemple, distinguer la proportionnalité), au sein des productions à évaluer, afin de disposer des réponses qui paraissent justes, tout en étant incapable d'en contrôler la validité. À partir de ce constat, d'une part, rien ne nous empêcherait de considérer qu'un référent puisse être « pauvre », c'est-à-dire ne contenir aucune réponse, et ne renfermer qu'une modélisation erronée de la situation ; bref, qu'il puisse se situer à une distance considérable d'un éventuel produit norme. D'autre part, nous pourrions voir dans le phénomène évaluatif mentionné (« s'appuyer ») qu'une modalité particulière d'élaboration du référent, et continuer de décrire, une fois tout cela admis, l'activité évaluative comme une comparaison référent/référé. Cependant, la notion de

réfèrent s'intègre dans une conception de l'évaluation où il est question d'une lecture du référé orienté par le réfèrent; or, il semble qu'ici l'évaluation devienne uniquement une lecture orienté par le référé, par les productions, les dialogues et les réponses des élèves. Bref, ce renversement total nous conduit à nous interroger sur la pertinence qu'il y aurait à conserver ce modèle pour le 4^e profil de comportement évaluatif. Nous pensons atteindre avec ce type de comportement une limite du domaine d'application du modèle réfèrent/référé, si l'on tient à rester en accord avec la conception de l'activité évaluative qu'il cherche à véhiculer.

Discussion des résultats

Le but de cet article était de présenter une étude de l'activité évaluative de professeurs stagiaires confrontés à une situation d'évaluation de productions d'élèves en mathématiques, situation où interviennent la proportionnalité et une représentation graphique. Il s'agissait d'interroger la description courante de l'activité évaluative qui, du point de vue du fonctionnement cognitif de l'évaluateur, considère cette activité comme une opération de confrontation entre un réfèrent et un référé. Plus précisément, nous voulions savoir si les modèles réfèrent/référé et explicatif permettaient aussi de décrire l'activité évaluative dans des situations autres que des situations de corrections de copies, d'une part, et s'ils permettaient de décrire l'activité évaluative de professeurs non spécialistes de la discipline à évaluer, d'autre part. Il s'agissait également d'étudier le réfèrent développé par chaque professeur, tant sur le plan de ses contenus que de ses modalités d'élaboration. Au terme de cette étude, il apparaît que le modèle réfèrent/référé, ainsi que le modèle explicatif sous-jacent, peuvent convenir pour décrire :

- l'activité évaluative dans des situations autres que des situations de correction de copies, en particulier la situation aménagée que nous avons utilisée et qui s'apparente à une sous-tâche d'évaluation formative (analyse des réponses et procédures des élèves);
- l'activité de professeurs spécialistes et non spécialistes de la discipline à évaluer, respectivement des PLC et des PE, que ces derniers soient scientifiques ou non scientifiques.

Cependant nous devons nuancer ces résultats car nous avons relevé des cas où il nous paraît plus difficile d'utiliser ces modèles. Nous allons compléter ces résultats en rappelant trois phénomènes évaluatifs qui interviennent lors de l'élaboration du référent. Nous n'en mentionnerons que les aspects les plus marquants, ceux qui pourraient servir à caractériser un profil de comportement évaluatif et les professeurs qui y sont associés.

Le premier phénomène évaluatif n'apparaît que chez certains professeurs scientifiques – PLC et PE scientifiques – rassemblés au sein du 1^{er} profil. Ceux-ci s'engagent dans la résolution du problème et élaborent un référent complet: ils disposent de l'ensemble des réponses aux questions et sont en mesure d'en contrôler la validité. Le phénomène qui les caractérise consiste à pouvoir choisir entre plusieurs procédures, plus ou moins coûteuses, pour obtenir efficacement ces réponses. La description de leur activité évaluative correspond pleinement à celle du modèle référent/référé et renvoie à la représentation «canonique» que nous pourrions avoir de l'évaluation des productions scolaires. Signalons tout de même que seulement quatre des professeurs engagés dans l'étude (soit moins d'un quart) correspondent à cette représentation «canonique».

La mobilisation vraisemblable de la structure de la solution est un deuxième phénomène évaluatif qui n'apparaît que chez des professeurs scientifiques, PLC et PE. En particulier, chez ceux qui sont rassemblés au sein du 3^e profil et qui ne s'engagent pas dans la résolution du problème⁹. Selon nous, ils peuvent s'en dispenser car ils disposent de la structure de la solution qui leur permet d'accéder à une compréhension globale du problème et sur laquelle ils se reposent pour mener leur activité. Les concepts impliqués pour comprendre la situation devraient en principe figurer au sein de leur référent, accompagnés de quelques attentes plus ou moins précises vis-à-vis des productions à évaluer. En revanche, aucune réponse aux questions ne s'y trouverait et, par conséquent, nous ne voyons pas vraiment ce que pourrait être ici le produit norme. Toutefois, bien que sur le plan de ses contenus ce type de référent ne corresponde pas à ce que l'on imagine pour un tel objet, nous considérons que, sur un plan fonctionnel, la structure de la solution serve de référent. L'activité évaluative de ces professeurs peut être décrite comme une lecture orientée des productions d'élèves, basée sur le système d'attentes de l'évaluateur jugées légitimes (Hadji, 1999), c'est-à-dire comme une opération de confrontation entre un référent – particulier, il est vrai – et un référé. Bref, il s'agit bien d'une confrontation référent/référé, mais pas sous sa forme la plus classique.

Nous avons relevé un troisième phénomène évaluatif qui consiste à s'appuyer sur les productions des élèves pour obtenir les réponses à certaines questions du problème (il s'agit de la forme particulière d'un phénomène plus général) et qui s'accompagne très fréquemment d'une incapacité à contrôler les réponses ainsi obtenues. Il n'apparaît que chez des PE non scientifiques qui relèvent des 2^e et 4^e profils. Les premiers (2^e profil) s'engagent dans la résolution du problème, mais celle-ci reste partielle. Ils ont un référent, mais pour certaines questions du problème il ne correspond pas vraiment à l'image « courante » associée à un tel objet : il ne comporte qu'une partie des réponses, dont certaines qui peuvent être erronées. Quant aux autres (4^e profil), ils ne s'engagent pas du tout dans la résolution du problème et rencontrent d'importantes difficultés pour le comprendre. Lorsque les premiers détiennent les réponses, leur activité peut-être décrite à l'aide du modèle référent/référé. Mais la question essentielle est de savoir si cette description reste acceptable quand un professeur dispose d'une réponse fausse ou quand il n'en a aucune et qu'il se fie uniquement à celle des élèves. Ces observations nous amènent à nous éloigner de la position tenue par Bonniol (1981, cité par Bonniol & Vial, 1997). Selon cet auteur « il semble bien qu'il y ait toujours un produit norme quand il y a évaluation, quel que soit le caractère vague et flou qu'il possède dans les représentations de l'évaluateur, quel que soit le degré de conscience que l'évaluateur peut en avoir » (p. 64). À plusieurs reprises, nous avons avoué nos doutes quant à la présence systématique de ce produit au sein de tous les référents et il nous paraît plutôt difficile d'entrevoir un tel produit, même « vague et flou », pour les professeurs qui relèvent du 4^e profil. Bref, il nous semble qu'il devient parfois impossible de penser l'existence d'un produit norme et que nous atteignons une limite du domaine d'application des modèles référent/référé et explicatif. En effet, même si nous acceptons l'existence d'un référent extrêmement « pauvre », sans le moindre élément pouvant s'approcher d'un produit norme, nous gardons à l'esprit que la notion de référent s'intègre dans une conception de l'activité évaluative selon laquelle c'est le référent qui oriente la lecture du référé. Or, il se produit ici un renversement complet qui donne lieu à une lecture uniquement orientée par le référé, ce qui rend difficile, voire impossible, une description à l'aide de ces modèles. Ajoutons, si nous revenons à la citation de Bonniol, que la prégnance de la structure de la solution, mise en avant pour comprendre l'activité de certains professeurs (3^e profil), pourrait renvoyer à leur faible degré de conscience du produit norme.

En observant la répartition des professeurs concernés par ces phénomènes, il apparaît que, dès que l'on s'intéresse aux contenus et aux modalités d'élaboration du référent, la séparation entre les professeurs ne s'effectuerait pas entre les spécialistes et les non-spécialistes de l'enseignement des mathématiques (PLC vs PE). Elle s'établirait plutôt entre ceux qui maîtrisent le problème et ceux auxquels il pose des difficultés de compréhension, autrement dit, elle se situerait sur le plan des connaissances disciplinaires des professeurs. Le fait que la non-résolution du problème n'ait pas les mêmes incidences, selon que l'on est, soit PLC ou PE scientifique (3^e profil), soit PE non scientifique (4^e profil), nous semble tout à fait illustrer cette séparation. Les phénomènes évaluatifs décrits témoignent, selon nous, du bien-fondé de prendre effectivement en compte les connaissances disciplinaires des professeurs pour décrire leur activité évaluative. Lors de l'introduction, nous disions que les docimologues avaient éludé la question des connaissances de l'évaluateur. Il faut reconnaître que ces derniers mentionnent les connaissances dans leurs travaux en tant que première « entrée d'information » dans la schématisation du modèle explicatif (Amigues et al., p. 796). Dans la modélisation du comportement évaluatif, elles interviennent cependant uniquement pour l'obtention « du produit norme qui est issu des objectifs pédagogiques et du corps de connaissances impliqués dans la tâche prétexte à l'évaluation » (p. 797). Ainsi, les connaissances semblent juste considérées, à notre avis, comme des éléments dont il suffit de mentionner la présence, soit parce qu'elles se situeraient du côté de la tâche d'évaluation, quel que soit l'évaluateur, soit parce qu'il semblerait acquis pour les docimologues que les spécialistes de la discipline à évaluer auxquels s'applique leur modèle disposent forcément de ces connaissances. C'est certainement pour ces raisons qu'elles ne sont jamais par ailleurs étudiées, ni même simplement présentées.

À ce stade de la discussion, on peut se demander si les interrogations relatives aux deux modèles n'apparaissent pas uniquement lorsque l'évaluateur est un non-spécialiste de la discipline à évaluer, discipline qu'il maîtriserait insuffisamment? Il semble que non: dans une étude récente conduite auprès de professeurs de mathématiques libanais titulaires, Nabbout (2006) relève que certains de ces professeurs, lorsqu'ils évaluent des productions d'élèves en probabilités, ont des comportements évaluatifs très proches de ceux que nous avons distingués chez des PE non scientifiques. Bref, ces modèles, malgré leur pertinence, ne conviendraient pas pour décrire l'ensemble des activités évaluatives de productions scolaires.

Conclusion

Quelques remarques et perspectives pour conclure.

La situation d'évaluation retenue est expérimentale et, comme nous l'avons signalé, elle renferme les limites indissociables de ce type d'approche. Elle en présente également les avantages ; en particulier, en décomplexifiant la tâche, les réductions ont effectivement permis d'étudier l'activité évaluative des professeurs sous son aspect cognitif. En quelque sorte, ce travail pourrait se situer dans la continuité des travaux des docimologues expérimentaux qui ont montré l'intérêt des approches expérimentales pour étudier l'activité cognitive de l'évaluateur. Il nous semble que ce qui vient d'être présenté pourrait modestement illustrer le point de vue de De Ketele (2001) qui déclare « si la docimologie était la discipline qui étudie le processus évaluatif (non seulement certificatif, mais aussi formatif) dans un contexte d'apprentissage, [le paradigme docimologique] devrait reprendre du sens » (p. 103).

Bien évidemment, l'approche par « situations aménagées » que nous n'avons utilisée qu'avec des professeurs stagiaires doit donner lieu à des études complémentaires. Actuellement, nous en projetons deux. D'une part, nous allons reconduire sous peu les expérimentations réalisées dans le cadre de notre thèse (Vantourout, 2004), en faisant appel cette fois à des professeurs « expérimentés » répartis dans les trois catégories : PLC, PE scientifiques, PE non scientifiques. Les résultats obtenus permettront d'effectuer une étude comparative entre professeurs débutants et professeurs expérimentés. D'autre part, nous avons commencé à nous intéresser à l'activité de PE dans leur classe, lorsqu'ils sont engagés dans des situations d'évaluations formatives continues avec régulations interactives (c'est-à-dire, dans des situations analogues à notre situation de référence). Certains résultats obtenus grâce aux situations aménagées (catégorisation des connaissances des professeurs, profils de comportement évaluatifs) serviront de base à la constitution de grilles de lecture pour appréhender l'activité évaluative des professeurs observés dans leur classe¹⁰.

Nous nous sommes déjà justifié en partie sur les raisons qui expliquent la mise au point d'un dispositif expérimental « sophistiqué ». L'une des principales fut, dès l'origine de la thèse, la volonté de disposer d'un dispositif de recherche qui – dans un deuxième temps et une fois aménagé – pourrait servir de dispositif de formation à l'évaluation formative. Pareillement, certains résultats actuels aideront à mieux appréhender l'activité des formés lors des séances d'évaluation. Nous pensons que ce dispositif, dans le cadre

d'une didactique professionnelle (Pastré, Mayen & Vergnaud, 2006) appliquée au métier d'enseignant, pourrait aider les formés à «pragmatiser» (p. 167) leurs connaissances théoriques dans les domaines de l'évaluation et de la didactique de la discipline concernée, et ce, afin d'en faire un bon guide de leur activité évaluative. Ce projet s'accompagne d'une réflexion théorique que nous menons actuellement en confrontant les situations aménagées au cadre de la théorie des situations développé par Brousseau (1998) pour les mathématiques, et ce, afin de savoir si les situations aménagées peuvent être considérées comme des «situations fondamentales» pour l'apprentissage de l'évaluation formative. Tous ces éléments, très succinctement présentés, contribuent au projet global d'aborder la question de l'articulation entre didactiques des disciplines et évaluations.

NOTES

1. Cette situation s'intègre dans un dispositif expérimental plus important qui n'est présenté que partiellement ici du fait de sa complexité. La mise en œuvre du dispositif se déroule sur trois séances, chacune faisant appel à un matériel spécifique sur le plan des contenus et s'organisant autour d'un problème de mathématique où interviennent la notion de proportionnalité et des représentations graphiques. Les données utilisées pour cet article proviennent de la troisième séance. Pour une présentation détaillée du dispositif, se référer à Vantourout (2004).
2. Cet énoncé s'inspire d'une situation présentée par Galai, Gerente, Grenier & Rivoire (1989).
3. Afin d'alléger le texte, nous utiliserons parfois des expressions plus courtes pour désigner les questions. Par exemple, «question 90 minutes» ou «question numérique et affine».
4. On trouvera chez Rogalski (2003) toutes les explications relatives aux concepts ergonomiques mentionnés dans ce paragraphe et dans le suivant.
5. En France, les IUFM sont les établissements qui assurent la formation professionnelle initiale des enseignants des premier et second degrés.
6. Les italiques signalent qu'il s'agit d'expressions utilisées par les professeurs lors du débriefing.
7. Par manque de place, nous ne pouvons pas expliciter davantage les procédures de résolution mentionnées dans l'article. Pour les connaître, se référer, par exemple, à Galai et coll., 1989.
8. En utilisant cette expression, nous nous inspirons de Mendelsohn (1981) qui dissocie, dans les procédures d'un sujet, ce qui renvoie à la logique structurale et ce qui dépend des contenus et des significations (p. 174). Le niveau structural permet d'accéder à la cohérence interne des conduites, de dégager ce qui est commun à toute une série de comportements, en épurant ces conduites de tout ce qui est contingent à une situation particulière (pp. 173-174). Parler de la structure de la solution, toujours en référence à cet auteur, revient à considérer l'isomorphisme des épreuves et des réponses. Cet isomorphisme renvoie à une épreuve largement maîtrisée, laquelle n'est rien d'autre qu'une épreuve dans laquelle les sujets finissent par choisir la procédure la plus économique (p. 194).

9. Il est probable que les professeurs qui relèvent du 1^{er} profil possèdent aussi la structure de la solution. Cette notion pourrait, par exemple, très certainement servir à rendre intelligible cette possibilité de réaliser des choix parmi des procédures plus ou moins coûteuses (*cf.* note précédente).
10. Dans les études à venir, pour que la situation aménagée s'approche davantage de la situation de référence (évaluation formative à régulation interactive), nous ajouterons à la consigne expérimentale actuelle, centrée sur les démarches d'analyse des réponses et des procédures des élèves, une demande de propositions d'intervention et d'interaction visant la régulation des apprentissages des élèves.

RÉFÉRENCES

- Allal, L. (1998). *Vers une pratique de l'évaluation formative*. Bruxelles: De Boeck.
- Amigues, R., Bonniol, J.-J., Caverni, J.-P., Fabre, J.-M., & Noizet, G. (1975). Le comportement d'évaluation de productions scolaires: à la recherche d'un modèle explicatif, *Bulletin de psychologie*, 28(318), 793-799.
- Barbier, J.-M. (1985). *L'évaluation en formation*. Paris: Presses Universitaires de France.
- Bonniol, J.-J. (1981). *Déterminants et mécanismes des comportements d'évaluation d'épreuves scolaires*. Thèse, Université de Bordeaux.
- Bonniol, J.-J., & Vial, M. (1997). *Les modèles de l'évaluation*. Bruxelles: De Boeck.
- Brousseau, G. (1998). *Théorie des situations didactiques*. Grenoble: La Pensée Sauvage.
- De Ketele, J.-M. (2001). Évolution des problématiques issues de l'évaluation formative. In G. Figari & M. Achouche, *L'activité évaluative réinterrogée* (pp. 102-108). Bruxelles: De Boeck.
- Galai, M.-C., Gerente, M., Grenier, D., & Rivoire, R. (1990). Analyse de deux situations problèmes autour de la proportionnalité. *Petit x*, 22, 5-22.
- Hadji, C. (1999). *L'évaluation démystifiée*. Paris: ESF.
- Joshua, S., & Dupin, J.-J. (1993). *Introduction à la didactique des sciences et des mathématiques*. Paris: Presses Universitaires de France.
- Mendelsohn, P. (1981). Analyse procédurale et analyse structurale des activités de permutation d'objets. *Archives de psychologie*, 49, 171-197.
- Nabbout, M. (2006). *Enseignement des probabilités en terminale au Liban: études des représentations et des pratiques dans des situations aménagées*. Thèse de doctorat non publiée, Université René-Descartes - Paris 5 Sorbonne.
- Noizet, G., & Caverni, J.-P. (1978). *Psychologie de l'évaluation scolaire*. Paris: Presses Universitaires de France.
- Rogalski, J. (2003). Y a-t-il un pilote dans la classe? Une analyse de l'activité de l'enseignant comme gestion d'un environnement dynamique ouvert. *Recherches en didactique des mathématiques*, 23(3), 343-388.
- Pastré, P., Mayen, P., & Vergnaud, G. (2006). La didactique professionnelle. *Revue française de pédagogie*, 154, 145-198.
- Scallon, G. (2000). *L'évaluation formative*. Bruxelles: De Boeck.
- Vantourout, M. (2004). *Étude de l'activité et des compétences de professeurs des écoles et de professeurs de mathématiques dans des situations « simulées » d'évaluation à visée formative en mathématiques*. Thèse de doctorat non publiée, Université René-Descartes - Paris 5 Sorbonne.

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.

Vermersch, P. (1994). *L'entretien d'explicitation en formation initiale et en formation continue*. Paris : ESF.

ANNEXE 1

Énoncé du problème

Énoncé Le cycliste

Un cycliste roule régulièrement, s'arrête, puis roule à nouveau à une allure régulière.

Le tableau suivant présente son parcours:

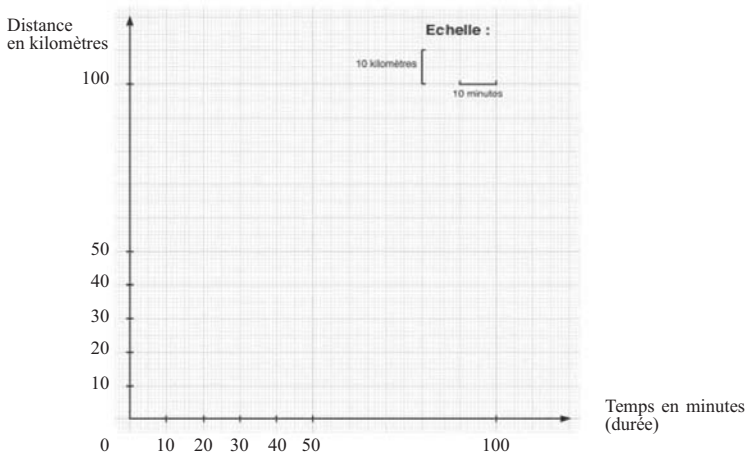
Temps (durée) en minutes	4	10	16	40	50	60	80	100
Distance parcourue en kilomètres	1	2,5	4	10	10	15	25	35

Questions (première partie):

1. Pouvez-vous trouver la distance parcourue par le cycliste au bout de 12 minutes?

Expliquez votre réponse.

2. Même question pour 20 minutes.
3. Même question pour 90 minutes.



En utilisant les axes ci-dessus, faites un graphique correspondant au tableau de données de la première partie du problème.

En utilisant le graphique, trouvez la distance parcourue par le cycliste au bout de 25 minutes, 37 minutes, 45 minutes et 108 minutes.

ANNEXE 2

Productions des élèves (extraits du travail d'Alexis et de David)

Réponse écrite Alexis et David

3) Nous ne sommes pas d'accord

Méthode de David: si en 50 min il fait 10 km alors qu'il devrait faire 12,5 km car en 10 min il en fait 10 km en 40 min

Ensuite il revient à une allure régulière. En 120 min il doit en faire 30 car en 60 min il a parcouru 15 km donc en 30 min il parcourt 7,5 km. En 90 min il parcourt 22,5 km.

40	50	60	80	100	120
10	12,5	15	20	25	30

[Tableau construit
par David]

Distance parcourue en 90 min

$$120 - 30 = 90 / 30 = 7,5 \quad / 90 = 22.5 \text{ km}$$

Méthode d'Alexis:

Distance parcourue en 90 min

En 90 min il fera 30 km

Car si de 80 min à 100 min il fait 10 km en 90 min il fait 30 km

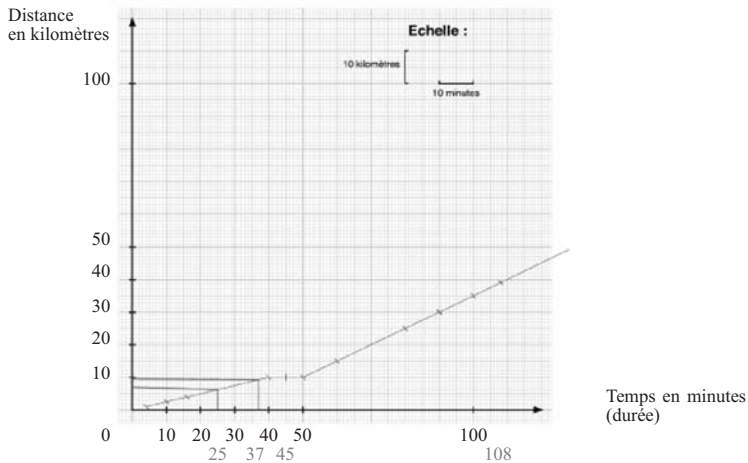
5) En traçant avec la règle on découvre qu'il fait 7 km en 25 min de même pour 37/45/108

$$37 \text{ min} = 9,5 \text{ km}$$

$$45 \text{ min} = 10 \text{ km}$$

$$108 \text{ min} = 39 \text{ km}$$

Graphique Alexis et David



Remarque: les traits diagonaux sur la réponse de David sont présentés comme ayant été inscrits par David lui-même.