Professionnalisation de l'évaluation aux Pays-Bas

Professional evaluation instruments in the Netherlands

Profesionalización de la evaluación en Holanda

Frank van der Schoot et Paul van Dam

Édition électronique

URL : <http://journals.openedition.org/ries/2528>

DOI : 10.4000/ries.2528

ISSN : 2261-4265

Éditeur

Centre international d'études pédagogiques

Édition imprimée

Date de publication : 1 juin 2000

Pagination : 73-81

ISSN : 1254-4590

Référence électronique

Frank van der Schoot et Paul van Dam, « Professionnalisation de l'évaluation aux Pays-Bas », *Revue internationale d'éducation de Sèvres* [En ligne], 26 | juin 2000, mis en ligne le 01 juin 2003, consulté le 30 avril 2019. URL : <http://journals.openedition.org/ries/2528> ; DOI : 10.4000/ries.2528

Professionnalisation de l'évaluation aux Pays-Bas

*Frank van der Schoot
Paul van Dam*

Résumé

Aux Pays-Bas, l'Institut national d'évaluation en éducation (CITO) a développé cette dernière décennie des programmes d'évaluation divers afin de mesurer les performances scolaires des élèves de l'enseignement primaire. Chaque programme a été développé dans un but spécifique mais un certain nombre de programmes ont un caractère plus ou moins multifonctionnel, les données pouvant ainsi être employées à plusieurs fins. Cet article décrit différents programmes d'évaluation de l'enseignement.

Professional evaluation instruments in the Netherlands

Over the past decade, the Dutch National Institute for the Evaluation of Education (CITO) has developed various evaluation programmes to measure the performance of primary schoolchildren. While programmes are generally devised for a specific purpose, other are multifunctional in approach, providing data for a wide range of applications. In this article, the writers describe various programmes used to evaluate teaching quality.

Profesionalización de la evaluación en Holanda

En Holanda, el Instituto Nacional de Evaluación en Educación (CITO) ha desarrollado durante la última década varios programas de evaluación de los performances escolares de los alumnos en la enseñanza primaria. Cada programa tiene un objetivo específico pero algunos revisten un índole más o menos multifuncional, sus datos tienen múltiples finalidades. Este artículo describe diferentes programas de evaluación de la enseñanza.

73

Dans la pratique quotidienne de l'enseignement, c'est l'enseignant qui détermine dans quelle mesure les élèves ont compris la matière enseignée pendant la journée ou pendant la période écoulée. Dans ce but, l'élève exécute les exercices en relation avec le sujet traité en cours. L'enseignant contrôle ainsi le travail de ses élèves, détermine si des actions supplémentaires sont nécessaires et, dans l'affirmative, quelles actions entreprendre pour l'ensemble du groupe ou pour l'un des élèves pris individuellement. L'action de l'enseignant est alors orientée sur son propre groupe, dans le cadre de sa propre méthode, sur la base d'un certain nombre d'exercices. C'est une forme d'évaluation nécessaire pour le contrôle du progrès quotidien de l'élève. Cette évaluation est dite «fermée» car elle ne reflète que la situation de l'enseignement. Il manque en effet à cette forme d'évaluation :

- des références externes, telles que les performances des élèves, de la classe et de l'école par rapport à d'autres élèves, d'autres classes et d'autres écoles ;
- des indications sur les déficiences de niveau de connaissances et de compétences qui pourraient se manifester à plus long terme chez l'élève, dans la classe ou dans l'école ;
- des exigences de qualités professionnelles permettant de tirer des conclusions valides et fiables.

Étant donnée la portée limitée des décisions prises par les enseignants sur la base des données de cette évaluation fermée, on admet volontiers que les aspects plus généraux cités précédemment fassent défaut. En revanche, lorsqu'il s'agit d'évaluer la qualité de l'enseignement proprement dit, cette forme d'évaluation est insuffisante. Une «évaluation ouverte» s'avère alors nécessaire prenant en compte la position de l'élève, de la classe ou de l'école par rapport aux groupes de référence, à l'échelon régional ou national.

On peut distinguer trois niveaux d'évaluation ouverte.

De même qu'il est important pour les écoles d'aller voir ce qui se passe hors de leur terrain de jeu, il est important dans le contexte de la politique à mener, d'aller voir hors des frontières

Au micro-niveau, l'évaluation de l'élève et de la classe

L'évaluation au micro-niveau concerne l'ensemble des performances de chaque élève ou de la classe en tant que groupe et permet de prendre des décisions d'orientation scolaire. On peut distinguer à nouveau trois types d'évaluation.

L'évaluation à court terme concerne les informations recueillies sur le progrès quotidien de l'acquisition des connaissances. Les différentes façons d'ob-

tenir ces informations sont bien connues : observer, poser des questions, faire exécuter des tâches, corriger le travail écrit et faire passer de courts tests en fin de chapitre ou d'unité de matière à étudier.

De telles activités font partie de l'enseignement quotidien et sont à peine considérées comme des activités d'évaluation. Les informations recueillies ont pour but de vérifier si les élèves et le groupe ont atteint de manière satisfaisante l'objectif pédagogique fixé. En général, on ignore presque tout de la valeur des instruments employés et des procédures d'évaluation.

L'évaluation à long terme concerne le besoin d'informations relatives aux progrès des élèves sur de plus longues périodes, par exemple, sur six mois. Il s'agit alors de savoir si les élèves ont réalisé les progrès souhaités. L'enseignant devrait disposer pour cela de données de référence prises hors de son propre établissement et, de préférence, à l'échelon régional ou national. De plus, les décisions prises en matière d'enseignement à suivre peuvent avoir de lourdes conséquences en cas de retard des élèves. Les enseignants doivent alors disposer d'informations fiables et valides sur les acquis de leurs élèves et du groupe dans le processus d'enseignement, qui vont bien au-delà des limites de la classe.

C'est la raison pour laquelle l'Institut national d'évaluation en éducation (CITO) a développé le système de suivi des élèves (SSE) pour l'enseignement aux Pays-Bas. Le SSE comprend des instruments d'évaluation pour toutes les classes

Tableau 1
Domaines du système du suivi des élèves (SSE)

Domaine	groupe/classe*
Rythme/rapidité de lecture	groupe 4 - 8
Lecture centrée sur la compréhension du texte	groupe 4 - 8
Compréhension orale	groupe 3 - 8
Expression écrite	groupe 6 - 8
Orthographe	groupe 4 - 8
Aptitudes de soutien de la langue	groupe 4 - 8
Vocabulaire	groupe 3 - 8
Calcul/mathématiques	groupe 5 - 8
Nombres et opérations	groupe 5 - 8
Mesures, temps et argent	groupe 5 - 8

* Aux Pays-Bas, l'école élémentaire comprend six classes, allant du groupe 3 (CP) au groupe 8 (CM2), l'école élémentaire en France n'en comprend que cinq.

de l'enseignement primaire (y compris la maternelle) et pour les matières suivantes : calcul/mathématiques, néerlandais, disciplines relatives à la découverte du monde. Le tableau 1 donne un aperçu des principaux domaines du SSE.

On fait passer les tests du SSE à un échantillon représentatif d'élèves, à l'échelon national. L'examen des résultats permet à l'enseignant de constater la position de ses élèves par rapport à la population d'élèves pertinente. Ce système a été développé sur la base de modèles de réponses par items. Cela permet de suivre l'évolution individuelle des élèves pratiquement tout au long de l'enseignement primaire. Une série de tests a été élaborée pour chaque domaine permettant de mesurer tous les six mois la compétence de l'élève. Les résultats de ces tests semestriels peuvent être représentés sur la même échelle de compétences. Les acquis en connaissances et compétences des élèves sont ainsi mis en évidence et, en outre, peuvent être comparés à la progression constatée dans l'échantillon.

Les résultats des tests informent ainsi l'enseignant sur :

- les différences du niveau de compétence des élèves de sa classe,
- le niveau de compétence des élèves par rapport au groupe de référence,
- le développement de la compétence individuelle des élèves.

Par ailleurs, l'élaboration de ces tests garantit l'objectivité, la fiabilité et la validité des informations.

Le choix du type d'enseignement

Le choix de l'orientation pour le secondaire est une décision importante qui doit être prise à la fin de l'enseignement primaire (groupe 8 aux Pays-Bas, soit CM2 en France). À cette fin, le CITO a développé «l'épreuve finale de l'enseignement primaire». Cette épreuve fournit des informations objectives aussi bien à l'enseignant qu'aux parents de l'élève ; ce qui permet à l'enseignant de les comparer avec son propre choix et aux parents de vérifier avec leur enfant si le niveau d'enseignement secondaire souhaité est réaliste. En vue de ce choix, l'école peut faire passer l'«épreuve d'entrée» à la fin du groupe 7 (CM1). Cette épreuve est, quant au contenu, analogue à l'épreuve finale de l'enseignement primaire et informe l'enseignant sur un éventuel retard scolaire de l'élève. L'enseignant peut alors examiner avec les parents et l'élève les résultats obtenus et, éventuellement, prendre les mesures nécessaires afin de rattraper le retard constaté et de le réduire au maximum.

Les enseignants éprouvent un grand besoin d'informations pertinentes et fiables. Cela se dégage de l'implantation dans tout le pays, du SSE, de l'épreuve d'entrée et de l'épreuve finale de l'enseignement secondaire.

Au moyen niveau, l'évaluation de l'école

Les écoles sont de plus de plus attentives à la qualité de l'enseignement donné dans leur établissement et les parents demandent davantage de comptes sur cette question. La qualité peut se définir de différentes manières, mais il est évident que le rendement de l'enseignement y occupe une place essentielle. Les données recueillies au micro-niveau peuvent être réunies et conduire, au moyen niveau, à l'évaluation interne et externe de l'école. En ce qui concerne l'évaluation interne de l'école, c'est l'école elle-même, c'est-à-dire le directeur et l'équipe pédagogique, qui examine les résultats scolaires réalisés par l'école. La multifonctionnalité des outils d'évaluation ressort de façon évidente, car les données peuvent servir non seulement à l'évaluation de l'élève mais peuvent également être interprétées au niveau de l'école. Il est possible d'évaluer les résultats scolaires aussi bien transversalement dans le temps que longitudinalement par cohorte d'élèves. Dans le premier cas, l'équipe de l'école analyse et compare les performances des élèves d'un groupe/d'une classe sur des années consécutives. Dans le deuxième cas, l'équipe de l'école analyse et évalue un groupe d'élèves tout au long de leur trajectoire scolaire.

À un autre niveau, les responsables des écoles et les municipalités veulent être tenus au courant des rendements des écoles placées sous leur autorité. On parle alors d'évaluation externe de l'école. Les responsables de la gestion scolaire ont besoin d'informations afin de gérer au mieux les écoles dont ils ont la responsabilité. Les municipalités doivent également pouvoir mener une politique responsable et fondée en matière d'enseignement et, par conséquent, disposer de données sur le rendement de l'enseignement dispensé. Au premier abord, il s'agit de signaler les difficultés rencontrées, ensuite de pouvoir évaluer l'efficacité de la politique menée. On peut constater un besoin grandissant d'informations aussi bien chez les responsables de la gestion scolaire que de la part des municipalités. Cela est certainement dû à la décentralisation de la politique d'enseignement aux Pays-Bas qui a donné plus de responsabilités et de pouvoirs aux instances locales.

77

Au macro-niveau, l'évaluation du système éducatif

Enfin, on constate également un besoin d'informations à l'égard du rendement du système éducatif et des résultats de la politique menée par le gouvernement dans ce domaine. La question se pose alors de savoir si le rendement de l'enseignement primaire répond aux objectifs du programme scolaire. Cette évaluation trouve son fondement dans la responsabilité des pouvoirs publics pour la qualité du système éducatif.

Le CITO se sert, à cette fin, de trois programmes de recherche. Le sondage périodique du niveau de l'enseignement (SPNE) rapporte les résultats scolaires dans l'enseignement primaire et les confronte aux objectifs du programme de l'enseignement primaire. Le projet PRIMA s'oriente plus spécifiquement sur les recherches concernant l'évaluation de la politique menée à l'égard des élèves ayant un retard scolaire. De même qu'il est important pour les écoles d'aller voir ce qui se passe hors de leur terrain de jeu, il est important dans le contexte de la politique à mener, d'aller voir hors des frontières de son propre pays et de participer à des recherches internationales comparatives.

Une construction modulaire

En 1986, le ministère de l'Éducation, de la Culture et des Sciences a mis sur pied le programme SPNE dans le but d'effectuer des sondages périodiques. Le SPNE permet «de recueillir de façon systématique des données sur les différents types d'enseignement proposés et sur les résultats de l'enseignement au profit de la discussion sur le contenu et le niveau de l'enseignement primaire.» Ce projet doit permettre de produire un rapport sur les résultats scolaires dans l'enseignement primaire au sens large du terme comme dans le détail.

Au sens large, dans la mesure où non seulement on recueille des données en calcul/mathématiques et en néerlandais, mais en outre, on tente également de faire un bilan sur tous les domaines d'apprentissage de l'enseignement primaire (découverte du monde, anglais, musique, éducation physique, initiation à la circulation routière, éducation artistique). Ces sondages ont lieu auprès des élèves à la fin de l'enseignement primaire. Des sondages ont également lieu à mi-parcours en calcul/mathématiques et en néerlandais. Dans tous ces cas, les résultats de l'enseignement sont rapportés de façon détaillée. Les champs d'application ou les unités de matière enseignées dans chaque domaine, et qui sont bien délimités, forment le point de départ. Les unités de matière sont caractérisées par le fait qu'on peut les considérer comme formant des unités pédagogiques significatives. Dans le sondage effectué en calcul/mathématiques, par exemple, on peut distinguer et faire un bilan sur vingt-quatre champs d'application différents. Pour le néerlandais, les aptitudes intégrées (compréhension écrite, compréhension orale, expression écrite et expression orale) mises à part, on peut distinguer également différentes aptitudes partielles telles que la grammaire, le style, l'orthographe, la ponctuation, l'analyse logique, l'analyse grammaticale, la réflexion sur la langue, la compréhension des mots. Dans les sondages concernant les disciplines relatives à la découverte du monde, on fait un rapport sur chacune des matières suivantes : histoire, géographie, biologie et physique et, au sein de chaque matière, un bilan sur différents champs d'application. En histoire, on distingue par exemple différentes périodes allant de la préhistoire au XX^e siècle, ou la notion de temps, ou bien encore les relations politiques.

Le caractère exhaustif de cette évaluation importe précisément pour la

discussion sur le contenu et le niveau de l'enseignement, et pour la planification du travail de recherches, car la « construction modulaire » en unités de matière reflète le nombre d'aptitudes enseignées. Elle permet également de confronter les résultats des élèves aux objectifs de l'enseignement primaire. La discussion sur le rendement d'une discipline enseignée peut alors porter de façon concrète sur des aptitudes individuelles et sur les choix à faire parmi ces aptitudes.

Sondages et flexibilité

La construction modulaire rend le sondage également flexible car elle permet d'ajouter de nouveaux sujets à la grille et de supprimer les sujets qui ne sont plus actuels. Il est aussi relativement facile d'apporter des modifications dans la conception du sondage. Certains changements de priorités apportés dans le curriculum ont souvent des répercussions sur un ou plusieurs modules, ce qui ne serait pas le cas dans des évaluations de niveau général.

Ceci nous amène au « caractère périodique » du sondage qui permet de mettre en évidence certaines tendances à travers le temps. En principe, le sondage effectué dans un domaine particulier est renouvelé au bout d'un certain temps. Depuis 1986, trois sondages ont eu lieu, à intervalle de cinq années pour chacun des domaines suivants : calcul/mathématiques, néerlandais et disciplines relatives à la découverte du monde. De cette manière, des comparaisons sur le rendement de l'enseignement sur plusieurs années sont possibles. Les discussions concernant les développements dans l'enseignement primaire ont acquis ainsi une base empirique. La conception du système du sondage périodique du niveau de l'enseignement offre aussi la possibilité d'analyser les aptitudes à l'aide d'autres moyens que les questions écrites, souvent « fermées ». C'est ce que l'on appelle l'« évaluation authentique ». Cette forme d'évaluation est moins souhaitable dans des projets à grande échelle, qui demandent un traitement rapide et efficace des données. Le SPNE essaie cependant d'analyser également des aptitudes telles que, en néerlandais, l'expression écrite et l'expression orale ; en éducation physique, le niveau d'entraînement physique des élèves ; en musique, des aptitudes musicales comme le chant, la notion du rythme et de la mesure ; dans les disciplines relatives à la découverte du monde, un aspect comme la formation de la compréhension et du jugement. Tout en continuant à construire sur la base des données recueillies dans le passé, on pourra peu à peu mettre l'accent sur des formes plus authentiques de la recherche en évaluation.

79

Information et politiques nationales

Le projet de recherche PRIMA a été élaboré, en premier lieu, pour fournir aux pouvoirs publics des informations sur les effets de la politique menée au

niveau national. On peut citer, par exemple, la politique concernant les élèves ayant un retard scolaire, qui a pour objectif de faire diminuer le nombre d'élèves orientés vers l'enseignement spécialisé. Un grand nombre d'écoles primaires et d'écoles d'enseignement spécialisé font passer, tous les deux ans, des tests et des épreuves dans les groupes 2, 4, 6 et 8. Ces tests et ces épreuves concernent l'intelligence, la langue, le calcul, le vécu scolaire de l'élève et les aspects du comportement socio-émotionnel.

La forme et le caractère de recherche de PRIMA permettent de faire des comparaisons non seulement transversales mais aussi longitudinales. Cela signifie que la plupart des élèves qui participent aux épreuves dans les groupes 4, 6 et 8 ont déjà participé aux épreuves deux ans auparavant, s'ils fréquentaient alors la même école et s'ils n'ont pas pris entre-temps de retard scolaire. Le fait que, dans la recherche PRIMA, des groupes d'élèves soient soumis aux mêmes épreuves, permet d'évaluer l'effet des décisions politiques prises à l'égard de l'enseignement. Beaucoup d'écoles faisant déjà passer des épreuves d'évaluation des élèves, dans le cadre du SSE, on demande au projet PRIMA d'en tenir compte dans sa recherche et d'essayer d'intégrer les données de ce système. Le tableau 2 résume les divers grands programmes d'évaluation de l'enseignement primaire néerlandais.

Comparaisons internationales

80

Pour terminer, les Pays-Bas participent régulièrement à des recherches comparatives internationales. Ce type de recherche livre également des informations pertinentes en matière de politique à mener, car elle permet de comparer les prestations des élèves néerlandais à l'échelon européen et même mondial. Ainsi, par exemple, le projet de recherche de l'IEA¹ concernant la compréhension centrée sur la lecture du texte a contribué à la création d'un centre d'expertise pour le néerlandais. Elle montre du point de vue international les atouts et les faiblesses du pays, offrant ainsi un point de départ pour la politique à suivre.

¹ Association internationale pour l'évaluation des résultats des élèves.

Tableau 2
Relations entre les programmes d'évaluation
aux Pays-Bas