

Évaluations nationales en France et contexte ultrapériphérique : un enjeu significatif pour l'école à La Réunion ?

Azzedine Si Moussa

Volume 38, Number 1, 2015

URI: <https://id.erudit.org/iderudit/1036550ar>

DOI: <https://doi.org/10.7202/1036550ar>

[See table of contents](#)

Publisher(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (print)

2368-2000 (digital)

[Explore this journal](#)

Cite this article

Si Moussa, A. (2015). Évaluations nationales en France et contexte ultrapériphérique : un enjeu significatif pour l'école à La Réunion ? *Mesure et évaluation en éducation*, 38(1), 31–56. <https://doi.org/10.7202/1036550ar>

Article abstract

From 1989 to 2012, nation-wide standardized testing in education took place in France in primary and early secondary schools. In overseas regions, such as Reunion Island, national testing needs to be assessed in relation to specific issues including the degree of local implementation of the education system and the historical, cultural and socio-economic contexts. As the national testing has been suspended, the question of its local relevance is raised while the level of school achievement remains low. This paper will discuss underlying theoretical issues of national testing and educational arguments raised by reflexive contributions of some local teachers in Reunion Island.

Évaluations nationales en France et contexte ultrapériphérique: un enjeu significatif pour l'école à La Réunion?

Azzedine Si Moussa

École supérieure du professorat et de l'éducation de La Réunion

MOTS CLÉS : évaluation nationale, contextualisation, La Réunion, département d'outre-mer

La France a organisé, entre 1989 et 2012, des évaluations standardisées sur l'ensemble du territoire national à l'attention des élèves en école primaire ou en début d'enseignement secondaire. Dans les départements d'outre-mer, comme celui de La Réunion, ces évaluations nationales représentent un enjeu particulier au regard du degré de développement local du système éducatif et des caractéristiques contextuelles très spécifiques de l'environnement ultramarin sur les plans historique, social, culturel et économique. Alors que ces évaluations sont actuellement suspendues dans leur application à l'échelle nationale, leur pertinence sur le plan local est interrogée, tandis que les résultats scolaires témoignent de difficultés importantes et récurrentes. La discussion intègre des arguments tenant de l'intérêt de ce type d'évaluations pour ce qu'elles représentent au point de vue des enjeux théoriques sous-jacents ainsi que des arguments issus des réflexions pédagogiques recueillies auprès des enseignants des écoles de La Réunion.

KEY WORDS: national testing, contextualization, Reunion Island, overseas region

From 1989 to 2012, nation-wide standardized testing in education took place in France in primary and early secondary schools. In overseas regions, such as Reunion Island, national testing needs to be assessed in relation to specific issues including the degree of local implementation of the education system and the historical, cultural and socio-economic contexts. As the national testing has been suspended, the question of its local relevance is raised while the level of school achievement remains low. This paper will discuss underlying theoretical issues of national testing and educational arguments raised by reflexive contributions of some local teachers in Reunion Island.

PALAVRAS-CHAVE: avaliação nacional, contextualização, Reunião, departamento ultramarine

Entre 1989 e 2012, a França organizou avaliações padronizadas em todo o território nacional aos alunos do 1.º e 2.º ciclos do ensino básico. Nos departamentos ultramarinos franceses, como o de Reunião, estas avaliações nacionais representam um desafio particular em termos do grau de desenvolvimento local do sistema de ensino e das características contextuais muito específicas do ambiente ultramarino nos planos histórico, social, cultural e económico. Embora essas avaliações estejam atualmente suspensas na sua execução nacional, discute-se a sua pertinência no plano local, uma vez que os resultados escolares demonstram dificuldades importantes e recorrentes. A discussão integra argumentos sobre o interesse deste tipo de avaliações por aquilo que elas representam do ponto de vista dos desafios teóricos subjacentes, bem como argumentos resultantes das reflexões pedagógicas recolhidas junto dos professores das escolas de Reunião.

NOTE DE L'AUTEUR : La correspondance liée à cet article peut être adressée à Azzedine Si Moussa, ESPE de La Réunion, allée des Aigues marines, 97487, Saint-Denis Cedex, France, téléphone : 02 62 90 43 71, ou à l'adresse courriel suivante : [azzedine.simoussa@univ-reunion.fr].

Entre pilotage du système éducatif et développement des départements d'outre-mer

La France se caractérise par un système éducatif largement centralisé : tous les élèves de l'enseignement primaire et secondaire suivent les mêmes programmes d'enseignement et valident des certifications ou des diplômes à dimension et à valeur nationale. Les enseignants suivent des cursus de formation accrédités à l'échelle nationale. Les élèves et les enseignants de l'Académie de La Réunion, région ultrapériphérique située dans la zone sud-ouest de l'océan Indien, intègrent donc ce schéma relativement uniforme, moyennant quelques rares et symboliques aménagements portant par exemple sur des parties du curriculum¹.

Concernant les données relatives au fonctionnement du système éducatif, dont font partie les résultats des évaluations nationales, la diffusion pour l'Académie de La Réunion, comme pour les autres départements d'outre-mer (DOM), apparaît a priori dans les mêmes publications que pour les autres régions françaises, mais le plus souvent au sein d'un alinéa ou d'un chapitre spécifique dédié à l'outre-mer. Cette présentation démarquée trouve sa justification dans les particularités contextuelles attachées à ces territoires, hors-normes sur les plans démographique, économique, social et culturel (Le Gall & Roinsard, 2010; Ascaride & Vitale, 2008).

Ces particularités, associées à la contrainte de l'éloignement géographique et des coûts de production y afférant, expliquent en partie la restriction de nombreuses études, recherches ou enquêtes dites nationales aux seuls contextes des régions françaises métropolitaines, dans le domaine éducatif comme dans d'autres champs d'investigation. Il s'avère également que, dans le cadre de la construction d'un échantillon représentatif de l'ensemble du pays, le nombre d'élèves scolarisés outre-mer pèse insuffisamment pour représenter un sous-échantillon de taille statistiquement exploitable. Dans le même temps, il existe peu de travaux concernant spécifiquement les DOM, sans doute là aussi essentiellement pour des raisons économiques².

Dans ces conditions, les évaluations nationales mises en œuvre dans les DOM font en quelque sorte figure d'exception, au point que leur pertinence peut être interrogée : à quoi servent-elles ? Sont-elles précisément indispensables dans le contexte particulier de l'outre-mer, à l'heure où le principe même des évaluations nationales semble remis en cause ? Existe-t-il des arguments spécifiques sur le plan contextuel devant orienter vers leur maintien ou, au contraire, vers leur suppression ?

L'examen de l'évolution historique des évaluations nationales en France et de leurs principaux enjeux fournit un premier cadre de réflexion global, enrichi par un aperçu nécessairement sélectif des apports internationaux de nombreuses contributions portant sur les évaluations standardisées³. Enfin, d'un point de vue local, la contextualisation opérée à travers l'analyse perçue (par les enseignants des écoles) et effective (de façon statistique et modélisée) des principaux résultats de ces évaluations éclaire le débat de façon à construire une argumentation dont les aboutissants apparaissent essentiels et d'autant plus nécessaires au regard de l'ampleur des difficultés scolaires à surmonter⁴.

Des évaluations nationales sur échantillon aux évaluations exhaustives en France : un outil de pilotage évolutif et en discussion

En 1979, à l'initiative de Claude Seibel⁵, pour la première fois en France, une évaluation à dimension nationale est organisée afin d'établir un constat des connaissances des élèves à l'école primaire. Il s'agit « d'observer l'évolution des résultats pédagogiques des élèves pendant la durée du cycle préparatoire » (Seibel, 1984, p. 14). Le dispositif s'appuie sur un échantillon représentatif au plan national et sur la passation de tests standardisés (conditions de passation et consignes de correction identiques). L'objectif ne se résume pas à évaluer les élèves, les enseignants ou les écoles, mais vise une vue macroscopique du système à travers des résultats mesurés à partir des performances des élèves. Ce protocole d'évaluation sera proposé à différents ordres d'enseignement, primaire et secondaire, et dans de nombreuses disciplines entre 1979 et 1987.

Progressivement se construit un dispositif d'évaluation plus centré vers deux disciplines de référence, le français et les mathématiques, et destiné à mesurer l'acquisition de l'ensemble des savoirs de base de l'école

primaire, en particulier le triptyque lire-écrire-compter avant l'entrée au collège. Cette transition annonce la mise en place des évaluations nationales diagnostiques en septembre 1989, organisées de façon systématique et exhaustive à l'entrée en CE2 et en 6^e sur l'ensemble du territoire national. L'objectif est multiple : sensibiliser et former tous les enseignants à l'évaluation à des fins d'amélioration des pratiques pédagogiques ; mobiliser l'ensemble des acteurs de l'institution scolaire sur les plans local et scolaire (recteurs, inspecteurs et enseignants) ; et, surtout, permettre d'exploiter les résultats des évaluations de façon comparative entre les classes, les écoles, les circonscriptions et les académies.

Apparaît ainsi une fonction plurielle des évaluations nationales, conçues comme un outil de pilotage éducatif à travers des actions de formation initiale et continue ; de repérage des élèves en difficulté ; de mise en relation des éléments de diagnostic et des pistes de remédiation ; de réflexion sur l'organisation pédagogique ; et de mise en place de programmes personnalisés de soutien pédagogique. Le protocole est présenté officiellement comme suit aux enseignants, selon les propos relatés par des inspecteurs généraux (IGEN & IGAENR, 2005, p. 13-14) :

La centralisation des résultats a permis de constituer une base de données dont l'utilité se révèle déjà considérable pour les recteurs, les inspecteurs d'académie, les inspecteurs départementaux de l'éducation nationale, les chercheurs et de calculer les résultats nationaux qui vous parviennent aujourd'hui. [...] Cette évaluation a été conçue comme un outil mis à votre disposition pour déceler, de façon précise et dès le début de l'année scolaire, les difficultés de vos élèves et vous permettre, dans toute la mesure du possible, d'y apporter rapidement une réponse.

Dans le même temps, les évaluations sur échantillon national n'ont pas disparu, s'orientant vers des études ciblées sur les compétences et les acquisitions des élèves en lecture, en écriture, en mathématiques ou en langues vivantes à différentes étapes du parcours scolaire et selon des intervalles temporels à échéance régulière, offrant une possibilité de comparaison diachronique. Ceci se concrétise par la mise en place en 2003 d'un dispositif qui demeure peu médiatisé, le cycle des évaluations disciplinaires réalisées sur échantillon (CEDRE), et se situe stratégiquement à la fin de l'école primaire et à la fin du collège.

Le dispositif CEDRE s'appuie sur la mesure des compétences d'un point de vue disciplinaire (en mathématiques, en histoire et géographie, en éducation civique, en langues vivantes, en sciences expérimentales), mais s'intéresse également aux « attitudes des élèves à l'égard de la vie en société » et à des compétences transversales telles que « la maîtrise du langage et de la langue française ». Dans le cadre des objectifs définis par le socle commun, l'évaluation mesure également, de façon graduée, des compétences générales telles que « prélever l'information », qu'elle soit explicite ou implicite, « organiser l'information prélevée », « synthétiser l'information » à l'école primaire et « exploiter l'information de manière complexe » au collège (Ministère de l'Éducation nationale, 2010, p. 1-2).

Le positionnement des situations d'évaluation dans des champs disciplinaires mais aussi en référence à des dimensions de la vie quotidienne rappelle opportunément le cadre défini à l'échelle internationale par le Programme international sur les acquis des élèves (PISA) de l'Organisation de coopération et de développement économiques (OCDE), dont les évaluations CEDRE s'inspirent largement. Le fonctionnement par cycle de six ans et la constitution d'un échantillon représentatif d'élèves en sont une illustration, de même que la construction d'une échelle de compétences par groupe de 0 à 5, selon les lacunes ou la fragilité graduelle des acquis des élèves, de nature à rendre plus difficile une formation qualifiante ou une poursuite d'étude, par exemple pour les groupes 0, 1 et 2. Les élèves du groupe 3 sont considérés comme moyens, tandis que ceux des groupes 4 et 5 sont considérés comme ayant des acquis solides ou très solides.

Selon le Haut Conseil de l'Éducation (HCE)⁶, l'apport principal du dispositif CEDRE réside dans la possibilité d'analyser l'évolution des compétences, par exemple entre 2003 et 2009 pour la « maîtrise du langage et de la langue française » (Haut Conseil de l'Éducation, 2011, p. 14) à la fin de l'école primaire : près de 40% des élèves de CM2 avaient des acquis insuffisants ou fragiles, proportion relativement stable sur la période, sachant que le même protocole d'évaluation a été proposé aux élèves.

Le HCE a par ailleurs remis en cause le bien-fondé des évaluations nationales exhaustives mises en place en 2009, aux niveaux CE1 et CM2, en lieu et place des évaluations diagnostiques de CE2 et de 6^e, mais avec la particularité de se situer en milieu d'année scolaire pour les CM2 et au mois de mai pour les CE1. Ces évaluations nationales ont été présentées par le ministère de l'Éducation nationale comme des évaluations-bilans.

« Ces évaluations situent les acquis de chaque élève par rapport aux objectifs définis dans les programmes. À partir de ce constat, les enseignants apportent une aide personnalisée aux élèves qui en ont besoin. » Elles ont aussi pour but d'« informer les parents et [de] leur permettre de suivre les progrès de leur enfant ». Enfin, il s'agit de « disposer d'indicateurs fiables des acquis des élèves pour mieux piloter le système éducatif » (HCE, 2011, p. 9-10).

Or, le HCE estime au contraire que le pilotage du système éducatif ne peut rigoureusement pas s'appuyer sur ce type d'indicateurs des acquis des élèves. La maîtrise des compétences du socle n'est pas assurée par l'observation de résultats relatifs à une proportion d'élèves ayant, en moyenne, des acquis solides. De plus, les conditions de préparation et de passation des tests sont sujettes à caution en raison de pratiques très variables des enseignants, conduisant par exemple à l'enseignement pour l'examen (*teaching to the test*), autrement dit à une forme de bachotage.

Depuis 2009, sans doute en lien avec un contexte politique national relativement tendu en matière d'éducation, les critiques se sont multipliées contre ces évaluations, essentiellement sur leur changement de calendrier et sur leur véritable finalité politique. L'argumentaire se situe à la fois sur le plan de la rigueur méthodologique, de l'apport pédagogique et de la perception même des évaluations, parfois considérées comme un instrument de contrôle hiérarchique des enseignants par l'administration.

Depuis 2012, l'exploitation des résultats des évaluations est suspendue par le ministère de l'Éducation nationale : « Les résultats des évaluations qui se sont déroulées en CE1 et CM2 du 21 au 25 mai 2012 ne sont pas centralisés au niveau national. » (Peillon & Pau-Langevin, 2012, p. 2) Ils sont collectés et analysés à l'échelle des écoles, et seulement à ce niveau, avec l'appui des inspecteurs des circonscriptions. Ces évaluations servent également de « support aux échanges avec les familles » et, s'agissant des évaluations réalisées en CM2, elles « favorisent la liaison avec le collège de secteur ». Pour l'avenir, la concertation traitera de la « refondation de l'évaluation tant du système éducatif que des acquis des élèves » (Peillon & Pau-Langevin, 2012, p. 2). Dans le cadre de la Loi d'orientation et de programmation pour la refondation de l'école votée en mai 2013, afin de mieux piloter les politiques éducatives et évaluer leurs résultats, le ministère de l'Éducation nationale propose de mettre en place le Conseil national de l'évaluation du système scolaire (CNESCO).

Ce rapide aperçu historique de l'évolution des évaluations en France montre que les politiques éducatives dans ce domaine ont suivi un cheminement parallèle et sans doute lié à l'évolution des évaluations standardisées à l'échelle internationale. Les enjeux et les réflexions y afférant ont donc aussi concerné la France, avec peut-être un point d'interrogation concernant un éventuel débat sur le contexte particulier traité ici.

Des concepts et enjeux des évaluations standardisées vers le questionnement sur la contextualisation

Les évaluations françaises se rattachent à l'ensemble des évaluations standardisées organisées de par le monde, soit à l'échelle nationale par de nombreux États, soit à l'échelle internationale, avec notamment le PISA de l'OCDE et les protocoles *Trends in International Mathematics and Science Study* (TIMSS) et *Progress in Reading Literacy Study* (PIRLS) développés par l'*International Association for the Evaluation of Education Achievement* (IEA). Une littérature scientifique abondante existe dans ce domaine, en particulier sur l'analyse du rôle de ces évaluations, devenues progressivement un outil incontournable de régulation des politiques éducatives en matière de résultats des systèmes éducatifs, mais aussi de modèles d'organisation et de mise en œuvre des méthodes pédagogiques. Au total, au moins trois entrées conceptuelles principales s'en dégagent, dessinant autant de thèmes de réflexion et de discussion, parfois caractérisés par de vifs débats⁷:

- les notions de performance et d'efficacité associées à la réussite éducative et donc mesurées, au moins partiellement, par les résultats des évaluations ;
- l'attente sociale vis-à-vis du système éducatif et de ses résultats, exprimée par le concept de redevabilité (*accountability*) ;
- en lien avec le cadre de l'évaluation d'une politique publique, l'enjeu de l'organisation administrative et de la décentralisation de l'éducation, et du rôle des acteurs à différents échelons du système.

Dans le cadre de la présente réflexion, centrée sur l'analyse du rôle des évaluations nationales dans un contexte régional ultramarin, trouver un angle d'analyse le plus convergent ou le plus approchant de ces entrées générales reste problématique, en l'absence de corpus théorique spécifique au-delà de premières contributions exploratoires (Si Moussa, 2003).

Comme la première fonction de l'évaluation consiste à donner une appréciation des résultats des élèves de façon à fournir une indication mesurée sur la qualité du système éducatif, l'enjeu de la diffusion de cette information une fois sa collecte réalisée semble important dans le cadre du contexte régional traité dans cet article. Il s'agit d'informer à la fois les acteurs de l'institution scolaire et les partenaires ou acteurs extérieurs, au premier rang desquels figurent les parents d'élèves et, plus généralement, la société civile. S'exprime ici une réponse immédiate à une attente sociale, en matière de redevabilité, sur les résultats comparatifs des écoles et du système en général. Cette attente peut être considérée comme particulièrement forte à La Réunion au regard des conditions historiques et socio-économiques de développement de l'île.

L'attente d'information est également institutionnelle et hiérarchique : à différents niveaux de responsabilité administrative et politique, les résultats obtenus représentent une indication essentielle pour le pilotage local, régional ou national de l'éducation, opérant une nécessaire connexion entre les différents acteurs impliqués (Woessmann, 2007). Cette connexion n'est pas neutre : dans le cadre de la décentralisation et de l'autonomie accordée aux établissements scolaires se jouent à la fois le contrôle social des acteurs locaux du système, la répartition des pouvoirs entre les uns et les autres, et l'attribution de marges de manœuvre plus ou moins étroites (Maroy, 2008 ; Broadfoot, 2000). C'est aussi par ce canal que peuvent plus directement se concrétiser des orientations pédagogiques et cognitives, avec une responsabilisation accrue à l'échelon d'une école ou d'une circonscription scolaire, voire d'une académie. En France, l'utilisation des résultats des évaluations nationales par les corps d'encadrement (administrations déconcentrées, corps d'inspection, chefs d'établissement) semble constituer une forme de valorisation des prérogatives hiérarchiques, mais aussi d'outils de pilotage, en particulier dans le cas des évaluations diagnostiques (Mons & Pons, 2006 ; Demailly, 2001). Ces différents enjeux restent là aussi prépondérants au regard d'un contexte ultrapériphérique marquant un mode de fonctionnement particulier qui mériterait sans nul doute d'être analysé.

Les conséquences de la mise en œuvre d'évaluations standardisées sont également analysées dans le cadre de la lutte contre les inégalités sociales de réussite scolaire (Hong & Youngs, 2008). L'uniformisation relative des attentes induites par la passation de tests identiques aurait aussi vocation

à encourager la progression des résultats des enfants des groupes sociaux défavorisés, en les poussant vers le haut. Cet argument revêt une importance particulière dans le contexte réunionnais, marqué par une forte différenciation sociale (Si Moussa, 2005). Il introduit aussi l'idée que l'écart de résultats entre La Réunion et la métropole (voir section suivante) doit tendre à se réduire au fil des années, comme il a pu être observé par ailleurs selon certains indicateurs, par exemple le taux de réussite au baccalauréat.

Il reste toutefois qu'au-delà de ces convergences sur des enjeux communs, le contexte de La Réunion se caractérise par deux logiques d'interprétation s'orientant vers des directions divergentes, l'une s'apparentant à une transposition totale du modèle éducatif hexagonal, l'autre s'attachant au contraire à la construction d'un *subystème* (Si Moussa & Tupin, 2005) priorisant au moins partiellement la prise en compte des spécificités locales. La réalité éducative oscille en permanence entre ces deux logiques qui, sans être nécessairement extrêmes, s'opposent, plaçant les acteurs du système dans une recherche récurrente de compromis. Au moins trois exemples significatifs illustrent ce phénomène :

- le calendrier scolaire : construit sur un rythme proche du calendrier métropolitain, il en diffère cependant par la longue coupure en milieu d'année scolaire (5 semaines de congé) et des vacances en cours d'année situées à des périodes différentes ;
- la langue officielle d'enseignement et la prise en compte de la langue régionale, le créole, à l'école : le débat demeure, certes limité par les dispositions constitutionnelles et législatives monolingues, mais donne lieu à des pratiques pédagogiques tentant d'intégrer (ou pas) le défi de l'apprentissage du français en milieu créolophone ;
- les contenus des enseignements : au-delà des aménagements mineurs en histoire-géographie évoqués précédemment, des choix pédagogiques peuvent être faits (ou pas) pour contextualiser davantage les situations d'apprentissage, en sensibilisant les élèves à ce qui les entoure, soit des éléments parfois très différents de la métropole.

Ces différents éléments conduisent finalement à la mention d'une quatrième entrée conceptuelle, largement analysée dans le cadre des travaux sur les évaluations internationales : celle des biais culturels (Vrignaud, 2002). Comme il sera présenté plus loin, les enseignants invoquent parfois la présence de biais dans les évaluations nationales. Cependant, en principe, la question de la présence d'un biais de construit (Van de Vijver

& Tanzer, 1997), qui traduit l'inconsistance de la variable mesurée par le test, ne se pose pas si l'on considère que cette consistance est garantie par la référence au programme d'enseignement qui est commun à tous les élèves. Le biais de méthode, imputable par exemple à des différences de degré de familiarisation avec les procédures utilisées lors de la passation des tests, paraît également peu envisageable, sauf si les élèves scolarisés à La Réunion étaient assujettis à des méthodes pédagogiques et à des protocoles d'évaluation sans rapport à ce qui se pratique en France métropolitaine, hypothèse fort improbable.

Reste le cas des biais d'items, qui mesurent la propension d'une question ou d'un thème à avantager un groupe culturel ou plus simplement sociodémographique⁸. De nombreuses contributions (Camilli & Shepard, 1994; Sireci, 1997; Laveault et Grégoire, 2002) analysent précisément les principaux enjeux de ces biais dans un cadre international, incluant les problèmes linguistiques liés notamment aux traductions des items dans les différents pays. Mais ce questionnement semble dépasser le cadre national français, d'un point de vue intérieur, en ce sens que l'homogénéisation structurelle et l'uniformité linguistique du système annihilent d'emblée des tentatives d'exploration scientifique dans cette direction. Il est ainsi frappant de constater que l'on se focalise exclusivement sur l'analyse des conséquences de la participation des élèves français aux évaluations internationales (Bottani & Vrignaud, 2005).

Pourtant, le cadre de fonctionnement décrit précédemment place objectivement les enseignants (et les élèves) des académies ultramarines telles que La Réunion dans une réalité d'exercice relativement éloignée des standards nationaux. De ce point de vue, le positionnement des enseignants sur les évaluations nationales apparaît potentiellement comme une source d'information intéressante.

L'utilité et la pertinence des évaluations nationales selon les enseignants

Nous avons rencontré, entre 2009 et 2011, près de 200 enseignants des écoles dans le cadre des travaux instaurés par l'Observatoire des pratiques pédagogiques de l'Académie de La Réunion. L'objectif initial de ces entretiens individuels consistait à identifier des pratiques pédagogiques efficaces et à analyser le rôle joué par différents facteurs déterminants de la

réussite scolaire. La présence des évaluations nationales, comme celle d'autres dispositifs mis en place par l'Éducation nationale, a ainsi fait partie des thèmes abordés par les enseignants.

Le discours des enseignants a en partie été influencé par des éléments conjoncturels liés à l'expérimentation très discutée médiatiquement des nouvelles évaluations nationales de CM2 placées en milieu d'année scolaire. Il en a résulté nombre de commentaires du type «Si c'est fait correctement, je suis pour, mais, pour moi, celles qui arrivent sont aberrantes». Au-delà de cette contingence particulière, le principe général des évaluations nationales a bien été discuté, notamment dans leur intérêt perçu pour les élèves, et bien entendu sur le plan local. Nous proposons ici une synthèse de différents extraits d'entretiens traduisant des opinions et des réflexions susceptibles d'éclairer notre questionnement, et illustrant les dilemmes qui peuvent caractériser un positionnement peut-être partagé vis-à-vis des logiques divergentes précédemment évoquées pour La Réunion.

L'adhésion au cadre de référence national et au principe de situation des apprentissages

Le discours des enseignants se centre tout d'abord sur une référence explicite au caractère national des évaluations pratiquées ainsi que sur la nécessité de s'y référer et de s'y situer, impératif ou qualité justifiant parfois exclusivement le dispositif :

- «Ce sont des évaluations nationales, un repère sur ce que l'enfant doit savoir. On se situe au niveau national. Il y a des choses à apprendre ; il faut se mesurer au niveau national. J'aime bien travailler sur leurs documents, voir où les enfants se situent. C'est utile ; c'est une façon de se repérer. Sinon, on est livrés à nos programmes.»
- «Ça permet de comparer, pour chaque classe, de faire une moyenne nationale, par académie.»
- «Au niveau national, c'est intéressant pour réfléchir sur sa pratique. Donc, les évaluations nationales, c'est OK pour mettre en relief les carences et les réussites des élèves.»
- «C'est une évaluation des enfants du pays pour voir comment ça fonctionne, les pratiques pédagogiques... Des fois, il faut faire le point pour faire avancer les choses, réajuster.»

- « Je suis plutôt favorable au principe pour que les enfants puissent se situer au niveau national. »
- « C'est bien, car l'élève sait où il en est. Nous aussi, car on sait que ça a été fait partout. »

Cependant, il arrive que les enseignants relativisent le crédit accordé aux évaluations nationales par manque de conviction sur le principe et son utilité générale, sur son apport au plan local ou encore sur le fait même d'évaluer les élèves :

- « Ça peut servir au niveau national. Nous-mêmes, on n'attend pas ces évaluations pour connaître les difficultés des élèves. C'est plus un outil pour le système, pas pour nous ; on sait déjà qui va réussir ou pas. »
- « Je ne sais pas comment ils vont exploiter ça, à quoi ça va servir. Le fait de pouvoir comparer nationalement, changer des choses, oui, pourquoi pas. »
- « Le principe général est de situer l'école par rapport au niveau national, comme point de repère. Pourquoi pas ? En même temps, c'est donner un classement. C'est bien en un sens, mais pas... L'évaluation en classe, je ne suis pas trop pour. En même temps, il faut passer par là pour savoir. »

L'argument comparatif et la possibilité offerte de situer les élèves sont également proposés de façon relativement générale, sans référence explicite à l'échelle nationale :

- « C'est nécessaire comme outil pour savoir où en sont les élèves, car ça permet de situer nos élèves. C'est important de faire cela. »
- « C'est un point de repère pour les familles, les élèves. »
- « Sur le principe, c'est bien de savoir le niveau des élèves, de comparer. »
- « Ça nous permet de voir le niveau en fin de primaire. C'est quand même un moyen intéressant pour connaître le niveau scolaire. »
- « J'aime bien pour me situer un peu. Je suis assez pour. Ça ne me rebute pas, mais il faut aussi que ça soit fait correctement. »

Les discours les plus négatifs développent l'idée principale que ces évaluations sont inutiles pour les enseignants, qui savent déjà ce qu'ils ont besoin de savoir concernant notamment les difficultés des élèves grâce à leurs propres évaluations :

- «Je crois moins aux vertus des évaluations nationales. En dehors de la connaissance de l'enfant, ça ne peut pas nous dire quel élève est en difficulté. Certains élèves le sont, mais les évaluations ne le disent pas; nous, on les connaît. Notre travail, notre connaissance de l'enfant peuvent nous aider à voir qui est en difficulté.»
- «Pour moi, dans ma classe, je ne vois pas l'intérêt de l'évaluation nationale. Je connais le niveau; c'est une perte de temps de corriger. Au niveau national, les résultats sont-ils fiables? Après les vacances? À quoi ça va leur servir? Savoir quelles sont les capacités des élèves, je ne sais pas, c'est un petit peu négatif. On peut comparer établissement par établissement, La Réunion/métropole... Il y a école et école; d'autres écoles ont des résultats meilleurs. Je ne sais pas si, dans le principe, c'est une bonne chose. Il fallait conserver peut-être en début de 6^e, mais je ne vois pas l'intérêt.»

Finalement, même en se plaçant du point de vue des élèves, les enseignants déplorent les modalités pratiques des évaluations nationales, et non leur principe de base. En revanche, une difficulté de fond est susceptible d'apparaître sur le plan des contenus, de par la référence à des réalités étrangères aux élèves de La Réunion.

Entre contextualisation et référence à un biais culturel

Dans cette perspective apparaît tout d'abord une contextualisation précise du discours, avec une référence à la spécificité réunionnaise, et en filigrane à ses difficultés :

- «Et puis, la situation scolaire par rapport à la métropole... Est-ce qu'on s'en rapproche?»
- «Pour situer les enfants de La Réunion, c'est intéressant. C'est mieux qu'une évaluation de chacun dans sa classe.»

Le problème de l'adaptation des évaluations peut être posé de façon relativement générale par les enseignants, avec éventuellement l'idée de protocoles départementaux ou institutionnels :

- «Pourquoi pas une évaluation pour chaque académie? Une pour La Réunion. Chaque académie aurait son projet.»
- «Dans le cas de La Réunion, il faudrait des évaluations départementales.»

D'autres réflexions sont plus précises, faisant référence à des contenus ou à des situations pour lesquels les élèves de La Réunion ne sont pas avantagés :

- «Ce n'est pas adapté à tous les publics ; c'est un peu faussé : le vocabulaire, les spécificités à La Réunion.»
- «Bien souvent, ce n'est pas adapté au milieu local. L'élève métropolitain est plus avantagé, par exemple sur les distances entre villes de métropole. Ici, l'élève réunionnais n'arrive pas à se les représenter.»
- «Ce n'est pas toujours adapté aux DOM. Le contexte est particulier, difficile pour le vécu, à imaginer.»
- «Ça devrait être régional, car il y a des particularités. Avec un texte en français sur des notions inconnues des élèves réunionnais, ceux-ci ont automatiquement moins de points que les autres. Il faut adapter les textes. Je n'étais pas la seule à avoir fait la remarque, avec l'exemple des giboulées.»

Le discours des enseignants peut aussi dépasser ce type de réflexion, voire décontextualiser en partie les difficultés éprouvées, et reconnaître l'utilité des tests :

- «Depuis des années, avec l'évolution de la société, les évaluations sont toujours bien adaptées pour les bons élèves, mais pas pour ceux en difficulté. Même si on est dans une école en difficulté, pour 15% des élèves, pas de problèmes aux évaluations. C'est le même niveau qu'en métropole pour ces élèves. Pour les élèves en difficulté, ces évaluations sont de bons outils par rapport à leur niveau réel. Mais ce sont des difficultés qu'on retrouve tous les jours. Le côté utilité de ces évaluations donne un seuil minimal d'acquisitions à obtenir. Toutes mes pratiques pédagogiques étaient toujours basées par rapport à ces résultats.»

Il reste que la proportion évoquée ici, seulement 15% d'élèves sans difficulté, replace la réflexion devant une situation bien particulière sur le plan local. L'évolution des résultats scolaires à La Réunion présentée par la suite en donne une illustration tangible.

Résultats aux évaluations nationales : évolution et variables prééminentes

Il est en effet intéressant de vérifier dans quelle mesure les résultats des évaluations nationales peuvent susciter des interrogations, sans doute légitimes, et peut-être présenter un enjeu déterminant pour la progression du sousystème éducatif local. Nous avons tenté de construire, sur une base la plus comparative possible, l'évolution des résultats de La Réunion par rapport à la moyenne nationale, en prenant en considération les deux types de protocoles qui se sont succédé (évaluations diagnostiques et bilans) et leur situation en fin d'école primaire (entrée en 6^e et CM2). Dans les deux cas, les résultats portent sur un échantillon stratifié en métropole (environ 5 000 élèves) et sur une population quasi exhaustive pour La Réunion (population de 15 000 élèves).

Dans le tableau 1 ci-dessous figurent les proportions d'élèves associés à un certain niveau de performance en français et en mathématiques, selon quatre seuils de réussite, à l'évaluation diagnostique organisée à l'entrée au collège jusqu'en 2008. Si un score de 50 sur 100 est considéré comme un seuil moyen, la proportion d'élèves se situant au moins à ce niveau s'établit, en français, à 75% pour La Réunion en début de décennie, puis décroît régulièrement pour se stabiliser à 42% en 2007-2008. Dans le même temps, cette proportion passe de 92 à 65% sur le plan national. Au-delà de l'explication de cette baisse tendancielle commune, l'information essentielle réside dans le maintien d'un écart de 20 points entre La Réunion et la métropole. En mathématiques, toujours sur la même période, les scores supérieurs à 50 points rassemblent un peu plus de 50% des élèves à La Réunion et un peu moins de 80% des élèves sur le plan national, soit un écart se situant autour de 30 points.

Pour le groupe des élèves les plus en difficulté au test (score inférieur à 30 sur 100), la proportion relevée pour La Réunion en français varie entre 16 et 14% entre 2001 et 2004, pour doubler ensuite entre 27 et 30% pour la période 2004-2008. Le même type de tendance est observé sur le plan national, mais sur une échelle sans commune mesure, entre 2 et 3% dans un premier temps, puis autour de 8% par la suite. Les résultats sont moins fluctuants en mathématiques, mais confirment un écart très important entre La Réunion et la moyenne nationale, avec une proportion d'élèves en grande difficulté autour de 15% dans le premier cas et de 4% dans le second.

Tableau 1
Résultats aux évaluations nationales en 6^e

Entrée en 6 ^e année	2001	2002	2003	2004	2005	2006	2007	2008
% d'élèves	R (M)							
Français								
Score > 70	38 (61)	23 (44)	23 (44)	25 (43)	17 (30)	16 (28)	17 (29)	17 (28)
50 < Score < 70	37 (31)	37 (38)	38 (40)	33 (40)	26 (39)	24 (36)	25 (37)	25 (37)
30 < Score > 50	19 (6)	28 (16)	28 (14)	28 (14)	30 (25)	30 (27)	29 (26)	31 (27)
Score < 30	16 (2)	12 (2)	11 (2)	14 (3)	27 (6)	30 (9)	29 (8)	27 (8)
Mathématiques								
Score > 70	25 (49)	21 (47)	19 (39)	22 (44)	19 (40)	21 (40)	22 (41)	19 (42)
50 < Score < 70	32 (32)	27 (30)	24 (35)	27 (34)	31 (36)	29 (36)	31 (36)	32 (36)
30 < Score > 50	28 (15)	29 (18)	30 (19)	32 (16)	34 (20)	34 (20)	32 (20)	37 (19)
Score < 30	15 (4)	23 (5)	27 (7)	19 (6)	16 (4)	16 (4)	15 (3)	12 (3)

Source des données initiales: Académie de La Réunion
 R : La Réunion (M) : Métropole

Il n'est pas inapproprié d'évoquer à ce niveau des tendances lourdes, à bien des égards, et il est sans doute instructif de vérifier si ces tendances sont corroborées par les résultats obtenus avec les évaluations nationales administrées en cours de CM2 à partir de 2009. Le tableau 2 ci-dessous reprend les résultats nationaux avec une répartition associée également à quatre seuils de réussite ou de mesure des acquisitions, toujours basés sur le nombre de réponses exactes au test.

En matière de tendance régulière des résultats, force est de constater l'extrême stabilité des proportions relevées d'une année à l'autre, peut-être en lien avec la répétition d'items à l'identique, protocole critiqué par le HCE, comme il a été expliqué précédemment. En français, la proportion d'élèves en réussite (Q1+Q2) s'établit à environ 50% pour La Réunion, contre près de 75% pour la moyenne métropolitaine. En mathématiques, les résultats sont quasi identiques pour La Réunion, toujours autour de 50%, et se situent entre 65 et 70% pour la métropole. Globalement, un écart de 20 points semble donc se maintenir.

Le groupe des élèves les plus en difficulté (Q4) est deux fois plus important à La Réunion qu'en métropole: 15 à 16% contre 7% en français, et une fourchette de 23 à 30% contre 10 à 15% en mathématiques, où la tendance générale semble dessiner une diminution de ce groupe ou, si l'on préfère, une moindre proportion d'élèves en échec important. Ces observations sont corroborées par une récente analyse des mêmes données centrée sur les moindres performances en mathématiques des élèves des DOM par rapport aux élèves métropolitains (Arneton, Bocéran, & Flieller, 2013).

Tableau 2
Résultats aux évaluations nationales en CM2

CM2		Janvier 2009	Janvier 2010	Janvier 2011
	% d'élèves	R (M)	R (M)	R (M)
Français				
	Q1	32 (45)	31 (43)	33 (43)
	Q2	28 (30)	27 (30)	28 (31)
	Q3	24 (18)	26 (20)	24 (19)
	Q4	16 (7)	16 (7)	15 (7)
Mathématiques				
	Q1	23 (35)	22 (35)	25 (38)
	Q2	24 (30)	26 (32)	27 (32)
	Q3	23 (20)	24 (20)	25 (20)
	Q4	30 (15)	28 (13)	23 (10)

Source des données initiales: ministère de l'Éducation nationale

Q1: Élèves ayant des acquis très solides

Q2: Élèves ayant de bons acquis qui seront développés dans les mois à venir

Q3: Élèves ayant des acquis encore fragiles qui seront à consolider dans les mois à venir

Q4: Élèves n'ayant pas les acquis suffisants et qui bénéficieront d'une aide spécifique

Ces premières analyses de résultats concernent les données de base des évaluations nationales selon les deux protocoles historiquement mis en œuvre, à l'entrée en 6^e (1989-2008), puis en CM2 (2009-2012). Nous avons eu l'occasion de combiner les résultats de ces deux protocoles d'évaluation dans le cadre de notre contribution pour l'Observatoire des pratiques pédagogiques en cycle 3 (Si Moussa, 2010); un de nos rapports d'étude concernait l'exploitation des résultats des évaluations nationales, en particulier ceux de septembre 2006 en CE2 et de janvier 2009 en CM2. Le choix de ces dates n'est pas anodin : il s'agit de la même promotion ou génération d'élèves, ce qui a permis de constituer un suivi partiel d'une cohorte d'environ 5 300 élèves ayant passé les mêmes tests, à trois années d'intervalle.

Une modélisation multivariée proposée ci-après (voir Tableau 3) concerne le score global obtenu en français et en mathématiques. Dans le cadre de ce type de traitement statistique, l'intérêt principal consiste à pouvoir raisonner à autres caractéristiques, scolaires ou sociales⁹, comparables. Dans le tableau 3, la valeur du coefficient indiqué donne une estimation du gain ou, au contraire, de la perte de points associé à chaque variable : par exemple, en français, les filles ont un avantage de 1,9 point sur les garçons, à autres caractéristiques comparables (score initial en français, origine sociale, retard scolaire); les élèves les plus défavorisés socialement perdent 7,4 points par rapport aux autres, toujours à autres caractéristiques données¹⁰.

Tableau 3
*Régression multiple sur le score en CE2/CM2
aux évaluations nationales 2006-2009*

Variables explicatives Variables expliquées	Score en CE2	Filles	Défavorisés socialement	Retard scolaire
Score en français en CM2	+0,8*** (+62,4)	+1,9*** (+4,9)	-7,4*** (-13,2)	-1,7*** (-4,2)
Score en mathématiques en CM2	+0,7*** (+51,8)	–	-8,4*** (-12,6)	-1,6*** (-3,3)

Source: Si Moussa (2010)

Ces résultats apportent deux dimensions supplémentaires par rapport aux précédentes analyses : la prise en considération des caractéristiques individuelles des élèves et de leur progression en matière de score d'un test à un autre. Sur le premier point, la confirmation de l'impact du facteur social est essentielle, de même que celui du retard scolaire, beaucoup plus important lorsqu'on se situe à un niveau scolaire donné (CE2 ou CM2) et qu'on ne prend pas en compte le niveau initial (ici, CE2). Sur le second point, on s'aperçoit du caractère prédictif de l'échec scolaire et de la difficulté à y remédier : le poids du niveau scolaire trois années auparavant est déterminant, voire irréversible, en cohérence avec les tendances observées à l'échelle nationale (Suchaut, 2007). Notons également que ces résultats, décomposés à l'échelon des circonscriptions, révèlent de fortes disparités. Celles-ci existent entre les circonscriptions, mais aussi entre les écoles et entre les classes d'une même école, interrogeant les modes de répartition des élèves et l'efficacité des pratiques pédagogiques.

Discussion conclusive

Si ces résultats restent à encadrer de marges d'appréciation et de précaution, les mouvements de fond qu'ils suggèrent et leur apport global doivent faire l'objet d'une attention particulière. De façon simple, il se confirme en effet que les résultats de l'école primaire à La Réunion demeurent à la fois largement insuffisants et, surtout, sans commune mesure avec ceux obtenus par les élèves des écoles métropolitaines, dont les résultats sont eux-mêmes qualifiés d'inquiétants dans le débat national. D'un point de vue politique, au regard d'un objectif prioritaire d'une réduction des inégalités géographiques de réussite scolaire, ici amplifiées à l'extrême, ce type de résultats revêt une importance stratégique incontournable. Sur la base d'un modèle éducatif partagé nationalement, dans ses curriculums et ses modes de fonctionnement, la dimension comparative inhérente à l'utilisation d'évaluations standardisées semble se justifier d'autant plus. En d'autres termes, dans les conditions actuelles de réussite éducative à La Réunion, l'absence d'outils d'évaluation à l'échelle nationale semble préjudiciable, sauf à entériner l'idée d'une évolution nécessairement différente se rapprochant par exemple de modèles régionaux décentralisés rattachés à des structures politiques fédératives, harmonisées mais pas toujours homogènes.

Ce scénario est en partie suggéré par certains enseignants rejetant la référence nationale, jugée inadaptée au niveau scolaire des élèves à La Réunion et, surtout, à leurs caractéristiques. S'exprime également un rejet plus fondamental de la standardisation des évaluations et, essentiellement, de ce qu'elle permet de faire, c'est-à-dire comparer et « classer » les écoles selon leurs résultats. Se dessine derrière cette argumentation un troisième rejet, lié à la vision d'une administration de l'éducation dont la volonté serait marquée par le souhait de contrôler et de sanctionner, le cas échéant, le travail effectué par les enseignants. Dans ce schéma, la notion de redevabilité et les moyens de la mettre en œuvre restent en suspens : le système éducatif doit-il rendre des comptes ? Comment peut-il le faire ?

Nous comprenons qu'il existe finalement deux niveaux de discussion, le premier étant relativement universel ou en tout cas national, tandis que le second est plus spécifique au contexte dans lequel les enseignants réunionnais travaillent.

Concernant le premier niveau de discussion, la pertinence même des évaluations standardisées est discutée, selon une perspective propre aux représentations particulières de l'évaluation en France ou, de manière plus générale, au regard des nombreuses réflexions suscitées par l'avènement des évaluations internationales comme instrument majoré pour l'orientation des politiques éducatives. L'image actuellement dégradée des évaluations nationales en France et les résultats décevants des élèves français aux différents tests internationaux construisent de ce point de vue une toile de fond négative, influençant sans nul doute les termes du débat, loin de considérations purement scientifiques et pédagogiques. À l'échelle internationale, la progression des protocoles d'évaluation, sans cesse améliorés et élargis, tant sur la rigueur méthodologique que sur la prise en considération d'items d'apprentissage diversifiés, en renforce la crédibilité. En témoigne le rythme d'adhésion croissant des pays participants, au-delà des divergences politiques, économiques et culturelles.

Concernant le second niveau de discussion, plus modeste mais aussi plus inexploré, la clé d'entrée essentielle paraît se situer dans les résultats des élèves dans le contexte ultramarin. Bien plus qu'ailleurs en France, il est question d'efficacité pédagogique, mais, surtout et simultanément, d'efficacité sociale (Si Moussa & Tupin, 2005) : la démocratisation de la réussite scolaire figure au premier plan des priorités, de façon d'autant plus cruciale au regard des situations d'illettrisme et des difficultés d'insertion

sociale. Les évaluations nationales peuvent-elles jouer un rôle constructif dans ce contexte? Indéniablement, si l'on considère qu'il est simplement préférable de privilégier la connaissance: se priver d'outils de mesure des inégalités sociales de réussite scolaire semble à la fois peu cohérent et peu constructif vis-à-vis d'une démarche visant précisément à les réduire. Se fixer des objectifs pédagogiques et didactiques différents du cadre national auquel on souhaite par ailleurs se référer semble également contradictoire. La validation des compétences pour les élèves de La Réunion doit nécessairement emprunter les chemins définis à partir des standards nationaux, fussent-ils difficiles à appréhender pour certains, comme le suggère la tension provoquée par l'articulation de pratiques pédagogiques adaptées aux possibilités avérées et à l'univers de référence des élèves à l'échelle locale, et par le suivi de programmes basés sur des séquences d'enseignement standardisées selon des critères d'évaluation nationaux (Si Moussa, 2012). Sans doute plus encore que, dans un cadre moins marqué socialement et économiquement, la référence à une démarche comparative constructive semble incontournable. Il suffit, pour en donner une illustration tangible, de remarquer la participation de certains pays en développement ou émergents aux évaluations internationales, anticipant au-delà de résultats de prime abord peu satisfaisants le bénéfice d'une évaluation essentiellement instructive et indicative pour se situer et mieux orienter leurs priorités en matière de politiques éducatives.

NOTES

1. L'essentiel des aménagements concerne l'enseignement de l'histoire et de la géographie, dont les dernières adaptations figurent au *Bulletin officiel* du ministère de l'Éducation nationale du 14 novembre 2013.
2. Des travaux sur la décentralisation du système éducatif français (Hirschhorn, 1991) se limitent ainsi au contexte métropolitain.
3. Dans un article de *Dossier d'actualité* de l'Institut français de l'éducation intitulé *PISA : ce que l'on en sait et ce que l'on en fait*, Rey (2011) propose une riche bibliographie, non exhaustive, sur ce point.
4. Les résultats de la Journée Défense et Citoyenneté 2013 montrent que la part des jeunes en difficulté de lecture varie de 5 à 18% dans les départements métropolitains, autour de 30% en Guadeloupe, Martinique et à La Réunion, et s'élève à 45% en Guyane et à 75% à Mayotte (Vourc'h, Rivière, De La Haye, & Gombert, 2014).
5. Chef du Service de l'informatique et des études statistiques (SIES) du ministère de l'Éducation nationale.
6. Instauré en 2005 dans le cadre de la Loi d'orientation et de programme pour l'avenir de l'école.
7. On pourra par exemple se référer à la synthèse de Mons (2009) pour un aperçu international de ces travaux et réflexions. Dans le contexte français, le débat, qui reste vif comme il sera expliqué plus loin, fut sans doute entamé avec Thélot (1993), puis alimenté avec les contributions de Duru-Bellat et Jarousse (2001), Perrenoud (2004) et Belair (2005). Plus récemment, dans un dossier de la revue *Sciences humaines*, Troger (2009) interroge Jean-Yves Rochex, dont les critiques à l'encontre du programme PISA témoignent de divergences de point de vue et d'appréciation qui confirment sans doute que la conception de l'évaluation en France, voire de l'éducation en général, reste éloignée de celle d'autres pays souvent cités en référence, par exemple la Finlande.
8. Par exemple, les garçons par rapport aux filles, avec un item portant sur du vocabulaire usité en pêche à la ligne, activité plus pratiquée par les garçons.
9. Les données sociales particulières de La Réunion ont mené à une distinction agréée de type « favorisé/défavorisé » reflétant de manière plus fiable un écart sociologique marqué.
10. Le nombre d'astérisques indiqué après chaque coefficient renseigne sur la significativité ou, si l'on préfère, sur la fiabilité de l'estimation. Le seuil de confiance accordé est de 99% pour ***, puis 95% pour ** et 90% pour *. S'il n'y a pas de coefficient indiqué, cela signifie que l'influence n'est pas significative. Figure entre parenthèses le test de Student, qui mesure la puissance de l'estimation en matière de probabilité : plus sa valeur absolue est élevée, plus la qualité de l'impact est forte. La valeur de significativité du test t s'interprète à partir d'un seuil estimé à environ 1,6 en valeur absolue.

RÉFÉRENCES

- Arneton, M., Bocéran, C., & Flieller, A. (2013). Les performances en mathématiques des élèves des départements d'outre-mer. *L'orientation scolaire et professionnelle*, 42(1), 33-66.
- Ascaride, P., & Vitale, P. (2008). L'île de La Réunion : regards contemporains, *Faire Savoirs*, 7(4), 34-45.
- Belair, L. M. (2005). Les dérives de l'obligation de résultats ou l'art de surfer sans planche. In C. Lessard & P. Meirieu (dir.). *L'obligation de résultats en éducation* (p. 179-188). Bruxelles : De Boeck Université.
- Bottani, N., & Vrignaud, P. (2005). *La France et les évaluations internationales*. Rapport pour le Haut Conseil de l'évaluation de l'école.
- Broadfoot, P. (2000). Un nouveau mode de régulation dans un système décentralisé : l'État évaluateur. *Revue française de pédagogie*, 130, 43-55. doi: 10.3406/rfp.2000.1052
- Camilli, L., & Shepard, L. (1994). *Methods for identifying biased test items*. Thousand Oaks, CA: Sage.
- Demailly, L. (2001). Enjeux de l'évaluation et régulation des systèmes scolaires. In L. Demailly (dir.). *Évaluer les politiques éducatives* (pp 13-30). Bruxelles : De Boeck Université.
- Duru-Bellat, M., & Jarousse, J.-P. (2001). Portée et limites d'une évaluation des politiques et des pratiques éducatives par les résultats. *Éducation et société*, 8, 97-134.
- Haut Conseil de l'Éducation (2011). *Les indicateurs relatifs aux acquis des élèves*, Bilan des résultats de l'école, Paris : HCE. Récupéré de http://www.hce.education.fr/gallery_files/site/21/114.pdf
- Hirschhorn, M. (1991). La décentralisation du système éducatif français. *Annuaire des collectivités locales*, 11, 397-419.
- Hong, W.-P., & Youngs, P. (2008). Does high-stakes testing increase cultural capital among low-income and racial minority students? *Education Policy Analysis Archive*, 16(6), 2-18.
- Inspection générale de l'éducation nationale & Inspection générale de l'administration de l'éducation nationale et de la recherche [IGEN & IGAENR] (2005). *Les acquis des élèves, pierre de touche de la valeur de l'école?* Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche, rapport n° 2005-079. Récupéré de http://www.education.gouv.fr/archives/2012/refondonslecole/wp-content/uploads/2012/07/rapport_igen_igaenr_les_acquis_des_eleves_pierre_de_touche_de_la_valeur_de_l_ecole_2005.pdf
- Laveault, D., & Grégoire, J. (2002). *Introduction aux théories des tests en sciences humaines*. Bruxelles : De Boeck Université.
- Le Gall, D., & Roinsard, N. (2010). *Chroniques d'une autre France : La Réunion*, Collection Sociologies et environnement, Paris : L'Harmattan.
- Maroy, C. (2008). Vers une régulation post-bureaucratique des systèmes d'enseignement en Europe? *Sociologie et société*, 40(1), 31-55. doi: 10.7202/019471ar

- Ministère de l'éducation nationale (2010). *L'évolution des compétences générales des élèves en fin de collège de 2003 à 2009*, Direction de l'évaluation, de la prospective et de la performance, Note d'information DEPP 10.22. Récupéré de <http://www.education.gouv.fr/cid54177/l-evolution-des-competences-generales-des-eleves-en-fin-de-college-de-2003-a-2009.html>
- Mons, N. (2009). *Les effets théoriques et réels de l'évaluation standardisée*. Complément à l'étude de l'EACEA et d'Eurydice, *Les évaluations standardisées des élèves en Europe: objectifs, organisation et utilisation des résultats*.
- Mons, N., & Pons, X. (2006). *Les standards en éducation dans le monde francophone: une analyse comparative*, Neuchâtel: IRDP.
- Peillon, V., & Pau-Langevin, G. (2012). Lettre à tous les personnels de l'éducation nationale, *Bulletin officiel*, 26 juin 2012, ministère de l'Éducation nationale. Récupéré de http://www.lemonde.fr/ecole-primaire-et-secondaire/article/2012/06/26/vincent-peillon-adresse-une-lettre-aux-personnels-de-l-education-nationale_1724558_1473688.html
- Perrenoud, P. (2004). Évaluer des compétences. *L'Éducateur*, n° spécial «La note en pleine évaluation», 8-11. Récupéré de http://www.ibe.unesco.org/fileadmin/user_upload/Poverty_alleviation/PresentationExperts/PresentationExperts_Madrid05_Perrenoud_EvaluerComp%C3%A9tences_FR.pdf
- Rey, O. (2011). PISA: ce que l'on en sait et ce que l'on en fait, *Dossier d'actualité*, 66, Institut français de l'éducation. Récupéré de <http://ife.ens-lyon.fr/vst/DA-Veille/66-octobre-2011.pdf>
- Seibel, C. (1984). Genèse et conséquences de l'échec scolaire: vers une politique de prévention, *Revue française de pédagogie*, 67, 7-28. doi: 10.3406/rfp.1984.1571
- Si Moussa, A. (2003, septembre). De la pertinence des évaluations nationales dans les DOM. In *L'efficacité sociale du système éducatif français dans les régions ultrapériphériques*, Colloque de l'ADMEE Europe, Liège.
- Si Moussa, A. (dir.) (2005). *L'école à La Réunion: approches plurielles*. Paris: Karthala.
- Si Moussa, A. (2010). *Analyse des pratiques pédagogiques et des déterminants de la réussite scolaire*. Rapport pour l'Observatoire des pratiques pédagogiques en cycle 3. Académie de La Réunion.
- Si Moussa, A. (2012, janvier). Évaluations nationales à l'école en France et contexte ultramarin: pour quelle validation des compétences des élèves? In *L'évaluation des compétences en milieu scolaire et en milieu professionnel*, Colloque de l'ADMEE – Europe, Luxembourg.
- Si Moussa, A., & Tupin, F. (2005). Efficacité sociale du système éducatif français dans les régions ultrapériphériques. In M. Demeuse et al. (dir.). *Vers une école juste et efficace* (pp. 115-132). Bruxelles: De Boeck.
- Sireci, S. G. (1997). Problems and issues in linking assessments across languages. *Educational Measurement: Issues and practice*, 16, 12-19.
- Suchaut, B. (2007). *Éléments d'évaluation de l'école primaire française*. Rapport pour le Haut Conseil de l'Éducation, IREDU-CNRS.
- Thélot, C. (1993). *L'évaluation du système éducatif: coûts, fonctionnement, résultats*. Paris: Nathan.
- Troger, V. (2009). École, des évaluations qui fâchent: Trois questions à Jean-Yves Rochex. *Sciences humaines*, 202, 10-11.

- Van de Vijver, F. & Tanzer, N. (1997). Bias and equivalence in cross cultural assessment: An overview, *European Review of Applied Psychology*, 47, 26-33.
- Vourc'h, R., Rivière, J.-P., De La Haye, F. & Gombert, J.-É. (2014). *Journée Défense et Citoyenneté 2013 : des difficultés en lecture pour un jeune Français sur dix*, Note d'information DEPP 12, ministère de l'Éducation nationale.
- Vrignaud, P. (2002). Les biais de mesure : savoir les identifier pour y remédier. *Bulletin de psychologie*, 55(6), 625-634.
- Woessmann, L. (2007). International evidence on school, competition, autonomy and accountability: A review. *Peabody Journal of Education*, 82(2-3), 473-497.

Réception: 13/08/13

Dernière version: 15/05/14

Acceptation: 21/08/14