Éducation comparée / nouvelle série, n°24-25, pp 41 - 64

# L'ÉVALUATION DE LA LECTURE DANS LE PISA ET LE SAEB : COMPARAISONS ET RÉFLEXIONS SUR LES INSTRUMENTS EN JEU

**Ada Magaly Matias Brasileiro**
**Univ. Fed. de Ouro Preto**
**(UFOP)**
**Brasil**

**Daniela de Faria Prado**
**Instituto Fed. de Educação,**
**Ciência e Tecnologia do**
**Triangulo Mineiro (IFTM)**
**Brasil**

**Lillian Gonçalves de Melo**
**Instituto Federal do Nordeste**
**de Minas Gerais (IFNMG)**
**Brasil**

RÉSUMÉ

Afin de mieux connaitre les propositions sous-jacentes aux évaluations à grande échelle des compétences en lecture, cet article vise à présenter les résultats d'une analyse comparative de deux instruments d'évaluation différents : le SAEB qui concerne les élèves brésiliens, et le PISA, appliqué à 80 pays, dont le Brésil. Pour ce faire, des comparaisons ont été effectuées entre quelques conceptions et critères adoptés par ces deux évaluations, puis plus spécifiquement, celles du PISA, en 2000 et 2009. Les résultats montrent une proximité entre les deux évaluations en ce qui concerne la méthodologie d'élaboration des éléments, et une distance entre les conceptions de la lecture adoptées, ce qui affecte la manière dont cette compétence est évaluée dans les deux processus.

MOTS CLÉS

PISA – Évaluation à grande échelle – Lecture – Instrument d'évaluation – Système d'évaluation de l'éducation de base au Brésil (SAEB)

ABSTRACT
Based on the need to know the propositions underlying the reading ability assessments in large-scale tests, this article aims to present the results of a comparative analysis performed between two instruments: SAEB, applied to Brazilian students, and PISA, in 80 countries, including Brazil. For this, comparisons were made between the conceptions and criteria adopted in both tests, as well as, specifically, in the instruments applied by PISA, in the years 2000 and 2009. The results showed an approximation between the two evaluations, regarding the elaboration methodology of items, and a departure from the reading conceptions adopted, which interferes at the way this ability is evaluated in both processes.

KEYWORDS
PISA – Evaluation in large scale – Reading – Evaluation tool – SAEB

## Introduction

Le Programme international pour le suivi des acquis des élèves (PISA), mené depuis 2000 par l'Organisation de coopération et de développement économiques (OCDE), dans le but de produire des indicateurs de qualité sur l'éducation mondiale, évalue maintenant 80 pays/économies du monde entier. Ce Programme est ainsi devenu la plus importante évaluation internationale en éducation. Selon l'Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira[1] (INEP), responsable du contrôle de la qualité de l'éducation au Brésil, les indicateurs PISA contribueraient aux politiques d'amélioration de l'éducation de base, car les résultats permettraient de déterminer si les pays participants préparent ou non leurs jeunes à exercer leur citoyenneté dans la société contemporaine (INEP, 2012). L'insertion du Brésil dans ce Programme s'explique par l'intérêt d'obtenir des informations permettant de situer la performance des ses élèves par rapport au contexte international. En outre, les données collectées par le PISA, discutées lors de forums ou de colloques internationaux, pourraient favoriser des discussions sur l'adoption de politiques publiques et de démarches visant le perfectionnement du système éducatif brésilien. Les analyses de l'OCDE (2014) et de l'INEP (2016) sur la situation de l'éducation au Brésil, depuis les années 1990,

1 Institut national d'études et de recherche sur l'éducation Anísio Teixeira.

sont remarquables et révèlent les progrès réalisés dans le pays grâce aux politiques éducatives. De telles études associent positivement ces résultats aux données obtenues par le PISA et le Système d'évaluation de l'éducation de base brésilien[2] (SAEB).

L'évaluation PISA a une périodicité triennale et est appliquée aux élèves âgés de 15 ans[3]. Chaque édition est axée sur un domaine de compétences – culture mathématique, culture scientifique et compréhension de l'écrit. Celle-ci a été la principale compétence évaluée en 2000, 2009 et 2018.

Au regard de l'ampleur du Programme, de l'importance internationale accordée aux indicateurs qu'il produit, de la répercussion médiatique des mauvaises performances du Brésil en lecture, puisqu'il n'a atteint que 407 points, dans PISA 2015, tandis que la moyenne globale des pays de l'OCDE correspond à 493 points (INEP, 2016, p. 128), un certain nombre de travaux ont été réalisés afin de porter un regard critique sur les paramètres pris en compte pour l'évaluation de la compréhension de l'écrit du PISA. C'est dans cette dynamique que s'est construit le projet de recherche collectif à l'origine des articles de ce dossier[4].

L'objectif spécifique de ce travail consiste à connaitre les propositions sous-jacentes aux évaluations des compétences en lecture, lors de tests à grande échelle réalisés au Brésil, concernant les critères adoptés, la composition des instruments et les compétences évaluées. Une analyse

2 Le SAEB est un ensemble d'évaluations externes à grande échelle, créées en 1990, visant à diagnostiquer l'éducation de base au Brésil, ainsi que les facteurs susceptibles d'interférer dans les résultats des élèves, créant ainsi des indicateurs sur la qualité de l'enseignement. http://portal.inep.gov.br/web/guest/educacao-basica/saeb.

3 Les élèves testés par le PISA sont ceux « qui avaient entre 15 ans et 3 mois révolus et 16 ans et 2 mois révolus au moment de l'évaluation, et avaient derrière eux au moins 6 ans de scolarité obligatoire, quels que soient leur année d'études, le type de leur établissement (établissement public, privé ou étranger), leur mode de scolarisation (à temps plein ou à temps partiel) et leur filière d'enseignement (générale ou professionnelle) » (OCDE, 2011, p. 25).

4 Ces travaux participent au projet Enjeux interculturels des évaluations et des comparaisons internationales en éducation: regards croisés Brésil-France PUC Minas (Brésil) et l'Université de Lille (France), 2017, comme indiqué dans l'introduction de ce dossier. Ils s'appuient également sur les travaux du Núcleo de Estudos em Linguagem, Letramentos e Formação de Educadores (Centre d'études en langue, Littératie et Formation d'Enseignants, programme de formation post-universitaire en Lettres de l'université PUC Minas de Belo Horizonte).

comparative a été menée, à cette fin, entre le test du SAEB, appliqué aux élèves brésiliens depuis 1990, et celui du PISA, appliqué aux élèves de 80 pays, dont le Brésil, depuis 2000. En outre, une analyse comparative a été réalisée entre les tests du PISA appliqués en 2000 et 2009[5], afin de permettre aux éducateurs brésiliens de mieux comprendre ce test qui évalue leurs élèves et dont les résultats peuvent parfois interroger ces professionnels.

Les objectifs descriptifs de cette recherche se concentrent sur les deux évaluations (PISA et SAEB) en ce qui concerne leurs modalités, leurs sujets abordés, leur nombre d'éléments ou la durée des épreuves. Cette description contribuera, ainsi, aux réflexions autour des critères d'évaluation de la compétence en lecture des élèves, et de la possibilité que se donne le PISA d'utiliser des textes identiques pour évaluer la lecture d'élèves de pays appartenant à des contextes distincts.

Le point de départ a été une recherche documentaire, destinée à savoir quelle approche de la lecture était sous-jacente aux deux évaluations - PISA et SAEB -, et à comprendre leurs catégories d'analyse respectives. Puis, une recherche bibliographique a porté sur la structuration des conceptions en jeu. Nous sommes revenues ensuite à la recherche documentaire, en utilisant les documents suivants: : les épreuves des années 2000 et 2009 (extrait des rapports PISA de 2000, 2001, 2009 et 2012), les Cadres d'évaluation de la lecture du PISA et du SAEB, les rapports de l'OCDE (2001, 2011, 2012, 2013 et 2014), OECD (2009 et 2016), et ceux de l'INEP/SAEB sur les résultats du Brésil (2001, 2012 et 2016).

Nous avons ensuite procédé à la cartographie des données, à partir des tableaux descriptifs des éléments composant les tests et des échelles d'évaluation respectives, en considérant les catégories d'analyse du cadre de référence de la lecture du PISA, par rapport au cadre de référence de la Prova Brasil (Épreuve Brésil) du SAEB, dans le but de montrer les proximités et/ou les écarts des compétences évaluées dans les deux tests et leurs critères d'évaluation respectifs. L'éclaircissement de ces aspects nous a permis de mieux comprendre à la fois les outils d'évaluation et les raisons possibles du mauvais résultat des élèves

5 En considérant que, jusqu'au moment de la réalisation de cette recherche, les seuls résultats divulgués étaient ceux des deux éditions ayant mis l'accent sur la lecture, nous avons décidé d'analyser les instruments appliqués dans les années 2000 et 2009. La présentation des résultats de l'édition 2018 s'est déroulée le 3 décembre 2019.

brésiliens. Nous avons ensuite comparé les tests du PISA organisés en 2000 et 2009, en utilisant les données fournies dans les rapports de l'OCDE. Cette stratégie nous a permis de proposer quelques réflexions qui pourraient contribuer à la compréhension de l'outil et de ses implications, en contrepoint des orientations adoptées au Brésil, concernant l'action même d'évaluer et le choix des compétences qui démontreraient la maitrise de la lecture par les élèves.

## Conceptions et critères d'évaluation de la lecture adoptés par le PISA et par le SAEB

L'une des questions récurrentes des recherches qui s'intéressent aux résultats des tests de lecture du PISA est la conception de la littératie adoptée. Ce concept est explicité dans les rapports de l'OCDE, qui montrent des changements subtils au cours des années 2000, 2012 et 2015, ainsi que dans le projet de Cadre d'évaluation proposé pour 2018, dans un rapport publié en 2016.

> Comprendre l'écrit, c'est non seulement comprendre et utiliser des textes écrits, mais aussi réfléchir à leur propos. Cette capacité devrait permettre à chacun(e) de réaliser ses objectifs, de développer ses connaissances et son potentiel, et de prendre une part active dans la société. [Définition adoptée dans le PISA 2000] (OCDE, 2013, p.23).

> Comprendre l'écrit, c'est non seulement comprendre et utiliser des textes écrits, mais aussi réfléchir à leur propos et s'y engager. Cette capacité devrait permettre à chacun de réaliser ses objectifs, de développer ses connaissances et son potentiel, et de prendre une part active dans la société. [Définition adoptée dans le PISA 2009 et 2015] (OECD, 2016, p.55).

Bien que cela ne soit pas toujours explicite, la définition adoptée par le PISA est liée à la conception de la lecture en tant qu'activité cognitive (la compréhension, l'utilisation, la réflexion et la connaissance), d'où ressort la vision selon laquelle le sens est immanent au texte (Jou & Sperbe, 2003). Cela diffère de la conception adoptée dans le SAEB, qui considère la lecture comme une activité socioculturelle, dont le sens est construit au moment de l'interaction auteur-texte-lecteur (Leffa, 1996). Une telle différence peut interférer à la fois dans le choix des textes pour la composition du test, ainsi que dans les critères pris en compte pour l'évaluation de la compétence en lecture des élèves.

Au Brésil, le SAEB, coordonné par l'INEP, guidé depuis 1996 par les Paramètres nationaux des programmes de la langue portugaise[6] (Brasil, 1997), a adopté une perspective de lecture socioculturelle, dont les aspects évaluatifs sont précisés dans des cadres de référence, adaptés aux programmes et aux contenus des écoles brésiliennes (en langue portugaise et en mathématiques).

Si nous examinons le cadre de référence pour la langue portugaise, et plus précisément la lecture, nous constatons que les compétences sont abordées sous forme de descripteurs (nom attribué à chacune des compétences à évaluer), construits à partir d'une association entre les contenus du programme et les opérations mentales développées par les élèves, en les traduisant en compétences et capacités (Gomes, 2006). Les documents de paramétrage du SAEB ne définissent pas explicitement la lecture, mais plutôt les connaissances linguistiques opérationnelles impliquant des actions avec et sur le langage, d'après ses critères d'évaluation.

Pour le SAEB, le genre du discours est vu à partir d'une conception du texte comme objet d'étude englobant les compétences discursives, considérées comme essentielles à la situation de lecture car elles visent à réaliser, entre autres, des opérations d'identification, d'inférences, de relation de cause à effet, de correction, d'implication et d'interprétation. À ce stade, on voit une proximité et une divergence avec la proposition du PISA : la proximité porte sur les opérations mentales requises et la distance réside dans le fait que, dans le travail du texte, le PISA, bien qu'il y fasse implicitement référence, n'utilise pas l'expression « genres du discours » dans ses directives, comme c'est le cas dans les documents brésiliens.

En ce qui concerne la méthodologie d'élaboration des instruments d'évaluation, les évaluations à grande échelle du SAEB, depuis 1995, sont produites sur la base d'une technique connue sous le nom de Blocos incompletos balanceados (BIB ; Blocs incomplets équilibrés), avec une distribution en spirale. Cette méthodologie permet d'appliquer 169 items de capacités matricielles différentes dans chaque niveau et discipline. Les items sont répartis en 13 blocs de 13 items

6 Paramètres nationaux des programmes de la langue portugaise (PCN). Il s'agit de directives élaborées pour guider les enseignants par une normalisation de certains aspects fondamentaux concernant chaque discipline de l'éducation de base brésilienne.

chacun, sous la forme de 26 cahiers différents. L'élève qui participe à l'évaluation répondra à 26 items de langue portugaise et 26 items de mathématiques, portant sur l'ensemble des compétences à tester. De même, dans le PISA,

> Les « unités d'évaluation » des épreuves papier-crayon du cycle PISA 2009 sont réparties en 13 groupes de 30 minutes chacun. Ces groupes sont au nombre de sept en compréhension de l'écrit, de trois en mathématiques et de trois en sciences. Ces groupes d'unités sont répartis dans 13 cahiers d'évaluation selon un modèle tournant, à raison de quatre groupes par cahier. Chaque cahier contient quatre groupes d'unités, dont un au moins en compréhension de l'écrit. Les élèves se voient attribuer l'un des 13 cahiers d'évaluation, dont la durée individuelle est de deux heures. (OCDE, 2012, p.15)

Une différence s'observe dans le fait que les élèves qui participent à une épreuve du PISA ne sont pas censés répondre à tous les items de l'évaluation, mais seulement à ceux auxquels ils parviennent en deux heures. L'élaboration des items et l'analyse des données s'inscrivent, dans les deux tests, dans la Théorie de la réponse aux items (TRI). Celle-ci permet de comparer les performances des élèves entre différentes populations au fil des années et entre les niveaux évalués dans chaque discipline, à condition que les épreuves comportent des exercices communs ou comparables. Cette méthodologie permet d'estimer la probabilité d'avoir une certaine réponse à un item (question), notamment pour les questions à choix multiple dont les options doivent proposer une réponse correcte et des réponses distractrices (réponses érronées) (Andrade, Tavares & Valle, 2000). Au fil des ans, les tests du SAEB et du PISA conservent certains items appliqués dans les épreuves précédentes, ainsi que le classement des items sur une échelle de performance, ce qui contribue à l'évaluation comparée du niveau de performance des élèves (Queiroz, 2013).

Dans notre recherche documentaire visant la comparaison entre le PISA et le SAEB de langue portugaise, nous avons eu du mal à accéder aux épreuves. En règle générale, seuls les rapports, qui ne contiennent qu'un certain nombre de questions, sont disponibles et, en plus, les outils employés ne sont pas accessibles aux chercheurs. Par conséquent, les études sur cet objet d'évaluation de l'éducation en sont affectées. Néanmoins, les documents sont dans le même temps de plus en plus

nombreux et complexes, de telle sorte que leur compréhension exige des connaissances spécialisées en statistiques.

Toujours axés sur les aspects statistiques qui décrivent ces deux tests, une autre question se pose concernant le nombre d'élèves brésiliens considéré comme cible. Selon le Recensement scolaire de l'éducation de base (2017), le Brésil compte 62 000 écoles en charge des dernières années de l'école primaire, ce qui correspond à 12 millions d'élèves. Sur ce total, 22,16%, soit 13 700 institutions, sont privées et 77,84% sont publiques, soit 48 300 institutions. Le SAEB est appliqué à tous les élèves des institutions publiques et à un échantillon d'au moins 10 élèves de chaque classe des écoles privées. Quant au PISA, il est réalisé sur un échantillon représentatif de l'ensemble de la population visée, c'est-à-dire, les élèves de 15 ans. Au Brésil, en 2018, il a été administré à 19 000 élèves, provenant de 661 écoles représentant les cinq régions du Brésil, et à au moins 50% des élèves présents le jour du test (INEP, 2019, p. 6).

Il convient de souligner que, contrairement au PISA, les questions du SAEB sont, comme nous l'avons vu, élaborées à l'aide de descripteurs (depuis 1991). Bien qu'il n'y ait pas de thématique qui mette en relation les sujets diversifiés des deux évaluations, les thèmes des évaluations du SAEB essaient d'illustrer la réalité brésilienne et de s'approcher du vécu social, régional et culturel des élèves brésiliens, même si ces réalités sont diverses à l'échelle nationale. En comparant les thématiques proposées par le SAEB et le PISA, il est possible d'observer que ce dernier tend à poser plus de difficultés aux élèves, en termes de construction d'inférences, car les textes choisis thématisent souvent des sujets sans rapport avec les réalités des élèves brésiliens.

En revenant à la conception adoptée par le PISA et aux critères d'évaluation, on remarque que les compétences en lecture sont évaluées au moyen d'items qui en déclinent sous différents *aspects*, supposés de décrire l'objectif des lecteurs d'un texte : *localiser et extraire des informations* (localisation de données provenant essentiellement du texte) ; *intégrer et interpréter* (utilisation de connaissances antérieures et externes au texte pour le comprendre globalement et pour en construire le sens) ; *réfléchir et évaluer* (capacité d'analyse critique du contenu et de la forme du texte). Dans les versions électroniques du test, l'aspect « *complexe* », applicable au texte numérique, a été ajouté en raison de son mode d'organisation plus fluide que le texte imprimé, ce qui

interfère également dans l'ordre de lecture construit par le lecteur (OCDE, 2013).

Du côté du SAEB, les compétences en lecture énumérées pour l'épreuve de 9e année sont résumées en 21 descripteurs répartis en six objets : *procédures de lecture ; implications du support, du genre et/ou de l'énonciateur dans la compréhension du texte ; relation entre les textes ; cohérence et cohésion dans le traitement du texte ; relations entre les ressources expressives et les effets de sens ; variation linguistique*. En comparaison avec les critères du PISA, nous pouvons voir que le troisième aspect de la compétence de compréhension du PISA, visant à *évaluer un texte*, n'est pas pris en compte par le cadre de référence du SAEB[7].

En ce qui concerne l'évaluation de ces différentes dimensions, le SAEB et le PISA utilisent une échelle de compétences, divisée en sept niveaux de compétence : 1b, 1a, 2, 3, 4, 5 et 6 (du plus bas au plus élevé). Le même texte peut donc, selon le PISA, être concerné par des questions exigeant des capacités différentes et évaluer différents niveaux de compétence.

Dans le PISA, ces compétences sont évaluées à travers différents textes dont les formats peuvent être continus (en prose), non continus (listes, formulaires, graphiques ou schémas), mixtes et multiples (informations provenant de textes de format et support différents), et de type varié (descriptif, narratif, informatif, argumentatif, prescriptif ou instructif). Comme l'explique Marcuschi (2002), il s'agit d'une classification traditionnelle des modes d'organisation du discours, différente du concept de genre du discours (adopté par le SAEB), qui envisage, dans sa composition, les types que nous avons mentionnés ci-dessus.

Dans le PISA, les textes à analyser par les élèves sont également caractérisés selon la situation[8] et leur support (papier ou électronique).

---

7 Ce facteur pourrait expliquer la faible performance que les élèves brésiliens ont eue à la question 5 de l'unité « Macondo », objet d'enquête de ce dossier, et qui visait notamment à tester la capacité des élèves à évaluer l'extrait·

8 Elles sont définies en fonction de l'usage pour lequel les textes ont été rédigés : romans, lettres personnelles ou biographies, sont considérés écrits en vue d'une lecture à usage « privé » ; les documents ou les avis officiels sont destinés à un usage « public » ; les manuels d'entretien ou les rapports sont destinés à une lecture à des fins « professionnelles » ; et les manuels scolaires ou les fiches d'exercices sont destinés à une utilisation « scolaire » (PISA, 2012, p. 13).

Dans le Tableau 1 ci-dessous, nous avons compilé à partir du rapport de l'OCDE de 2012 les caractéristiques prises en compte par le PISA dans la classification des textes.

**Tableau 1.**
**Caractérisation des textes - données extraites du rapport PISA (2012)**

| Caractéristiques | Catégories |
|---|---|
| **Situation** | personnelle, publique, éducative, professionnelle |
| **Support** | papier, électronique |
| **Format** | continu, non-continu, mixte, multiple |
| **Type** | narratif, descriptif, expositif, argumentatif, prescriptif ou instructif, interactif |
| **Aspect** | localiser et récupérer, intégrer et interpréter, réfléchir et analyser, complexe |
| **Structure des questions** | fermée (choix multiple), ouverte |

Parmi les six caractéristiques (situation, support, format, type de textes, aspect et structure des questions), les cinq premières se réfèrent à des critères à respecter lors de la conception de l'épreuve, servant ainsi de paramètre au choix diversifié des textes. La dernière caractéristique (structure des questions) renvoie plus à l'évaluation des réponses des élèves. Malgré cela, il est clair que pour que l'élève développe les compétences en lecture *localiser et extraire, intégrer et interpréter et réfléchir et évaluer des informations*, visées dans le PISA, il devra être un lecteur des genres du discours, que ce soit dans les sphères de communication décrites par Bakhtine (1984) ou dans les quatre situations de lecture auxquelles les textes sont destinés selon la classification du PISA (personnelle, publique, éducative ou professionnelle). Il faut d'ailleurs souligner que, depuis 2018, la Base nacional comum curricular (Nouvelle base commune des programmes nationaux brésiliens, 2018, p. 69), peut-être influencée par le PISA, a inclus cette terminologie, considérant les pratiques linguistiques contemporaines essentielles à la participation dans les sphères de la *vie publique, professionnelle et personnelle*. Bien qu'elle ne mentionne pas la sphère éducative, le vocabulaire choisi est similaire à celui utilisé dans le PISA, ce qui n'était pas le cas dans les documents précédents.

**Tableau 2.**
**Le test de lecture du SAEB – données extraites du site du SAEB/INEP (2019)**

| Caractéristiques | Catégories |
|---|---|
| **Support** | Imprimé |
| **Format** | Multimodaux |
| **Genres du discours** | Qui circulent dans des situations sociales les plus diverses |
| **Structure des questions** | Fermée (choix multiples) |
| **Objectif des questions** | Compétences réparties en six objets : procédures de lecture ; les implications du support, du genre et / ou de l'énonciateur dans la compréhension du texte ; relation entre les textes ; cohérence et cohésion dans le traitement de texte ; relations entre ressources et effets de sens ; et variation linguistique (traduites en 21 descripteurs pour l'évaluation) |

Dans une étude comparative des critères d'évaluation du SAEB et du PISA, Lopes-Rossi et Paula (2012) ont conclu que, dans le cadre d'évaluation brésilien, bien qu'il n'existe pas de descripteur faisant référence à la position critique du lecteur (contrairement au PISA), les compétences en lecture vont au-delà du décodage, en visant la compréhension de textes de genres différents, construits dans certains contextes de production et de circulation sociale, dans une perspective fonctionnaliste du langage.

Pour conclure cette partie, nous pouvons synthétiser les similitudes et différences que nous avons soulignées en comparant les conceptions et les critères utilisés dans les deux évaluations à grande échelle. Le PISA et le SAEB sont donc proches sur le plan de la méthodologie choisie pour l'élaboration de l'instrument, de l'échelle de compétences (à l'exception de la compétence à évaluer un texte) et dans les niveaux de compétences. En revanche, ces deux évaluations se distinguent dans la conception de la lecture, dans l'approche du texte (comme genre ou type de texte) et dans la contextualisation des thématiques choisies pour

les textes des examens, en ce qui concerne la réalité socioculturelle des élèves brésiliens.

## Les instruments d'évaluation de la lecture administrés lors des PISA 2000 et 2009 : Une tentative de cartographie

Après l'analyse comparative des aspects méthodologiques et conceptuels qui guident la structuration des tests de lecture appliqués dans le SAEB et le PISA, notre analyse va porter sur l'organisation globale des instruments d'évaluation du PISA 2000 et 2009, avec comme sources principales les rapports de l'OCDE parus en 2001, 2011, 2012, 2013 et 2014 et OECD, 2009 et 2016. Cette stratégie est justifiée, comme indiqué ci-dessus, par la nécessité de connaitre plus clairement ce test qui évalue les élèves et qui, par sa diffusion fragmentée, dans ses rapports volumineux, dans différentes langues, finit par être peu connu des éducateurs, un public pourtant très intéressé par les causes et les conséquences des résultats négatifs qui se répètent à chaque édition. Il est donc nécessaire d'éclairer ce qui semble encore être mal compris. Pour l'analyse quantitative, nous avons utilisé les items publiés dans les rapports de l'OCDE consacrés aux éditions de 2000 et 2009 de la manière suivante : 19 items du PISA 2000 relatifs à 6 unités ; 40 items du PISA 2009 relatifs à 10 unités ; et 8 items de la version électronique du test de 2009, relatifs à 3 unités, sur un total de 67 items. Ces données nous ont également permis de comparer les deux éditions du test, ce qui nous a conduit à décrire la logique d'élaboration du PISA.

Une caractéristique importante, quand on s'intéresse à une évaluation, concerne le score maximal qu'elle permet d'obtenir. La méthodologie de la TRI suppose qu'un nombre de points plus élevés est attribué aux questions avec un taux de réussite inférieur et réciproquement. Il n'est donc pas possible de fixer le nombre de points le plus élevé pour l'instrument. Ainsi, parmi les documents consultés, on ne voit pas quel serait le score maximal qui pourrait servir pour une comparaison.

Comme le montre le tableau suivant, l'échelle de points correspondant à chaque niveau de compétence élaborée par le PISA permet néanmoins de comparer les performances des pays et des élèves participants.

**Tableau 3.**
**Score correspondant à chaque niveau de compétence (OCDE, 2013)**

| Niveau | Score |
|---|---|
| 6 | 698 |
| 5 | 626 |
| 4 | 553 |
| 3 | 480 |
| 2 | 407 |
| 1a | 335 |

En ce qui concerne le poids relatif des trois aspects des compétences en lecture évalués dans les différentes questions en 2000 et 2009, le PISA a publié des données quantitatives qui sont rassemblées dans le Tableau 4 ci-dessous.

**Tableau 4.**
**Distribution en pourcentage des compétences en lecture dans les items de 2000 et 2009 (d'après OCDE, 2001 et 2012)**

| Objectifs de la question | % d'items en 2000 | % d'items en 2009 (papier) | % d'items en 2009 (électronique) |
|---|---|---|---|
| Localiser et extraire | 40 | 25 | 25 |
| Intégrer et interpréter | 30 | 50 | 35 |
| Réfléchir et évaluer | 30 | 25 | 20 |
| Complexe | - | - | 20 |
| Total | 100 | 100 | 100 |

Ces données montrent que, dans l'édition de 2000, l'accent était mis sur les compétences en lecture littérale dans la mesure où l'aspect de *localisation et d'extraction* d'informations représentait 40% des items ; dans l'édition papier-crayon de 2009, les compétences inhérentes à la compréhension du texte *intégrée* avec d'autres connaissances étaient majoritairement évaluées avec 50% des questions. Fait intéressant, cet aspect relatif à *l'intégration et l'interprétation* ne représentait que 35% des items dans la version électronique du test appliqué, pourtant, la même

année. Lors des deux éditions, on note également que l'aspect le moins représenté dans les trois épreuves était celui de *réfléchir et d'évaluer* le contenu et la forme des textes (respectivement 30, 25 et 20% des items). La variation de ce paramètre pose alors le problème de la comparaison puisque différents niveaux de difficulté semblent testés d'une épreuve à l'autre. Dans le tableau suivant, une autre facette de ces tests est présentée : les formats des textes sélectionnés.

**Tableau 5.**
**Distribution en pourcentage du format des textes dans les items de 2001 et 2012 (d'après OCDE 2001 et 2012)**

| Formats des textes | % d'items en 2000 | % d'items en 2009 (papier) | % de items en 2009 (électronique) |
|---|---|---|---|
| Continu | 66 | 60 | 10 |
| Non-continu | 34 | 30 | 10 |
| Composé | | 05 | 10 |
| Multiple | | 05 | 70 |
| Total | 100 | 100 | 100 |

Dans la classification du PISA, un texte continu est organisé en paragraphes tandis que les textes non continus peuvent prendre diverses formes comme des graphiques, des cartes, des vers, des tableaux, des dessins, etc. Les textes de format mixte sont constitués de parties continues et non continues et les textes multiples combinent différents textes juxtaposés. Ces deux derniers formats ne figuraient pas dans l'édition 2000. Par contre, les textes multiples étaient les plus utilisés dans la version électronique de 2009, sans aucune mention du genre de ces textes.

Là encore, il existe une disparité quantitative dans le choix des textes quant à leur format qui interroge également la comparabilité des résultats du PISA, puisque les ressources et les textes utilisés dans les différentes épreuves sont variables.

On remarque un phénomène similaire dans le Tableau 6 ci-dessous qui présente les pourcentages d'items portant sur chaque situation de lecture.

**Tableau 6 – Distribution en pourcentage des situations de lecture dans les items de 2000 et 2009 (d'après OCDE, 2000 et 2012)**

| Situations de lecture | % d'items en 2000 | % d'items en 2009 – écrite | % d'items en 2009 – électronique |
|---|---|---|---|
| Personnelle | 28 | 30 | 30 |
| Éducative | 28 | 25 | 15 |
| Professionnelle | 16 | 15 | 15 |
| Publique | 28 | 30 | 40 |
| Total | 100 | 100 | 100 |

Les données révèlent également une asymétrie dans le choix des textes en fonction de la situation de lecture[9] ; toutefois, cela est justifié dans le rapport 2012 de l'OCDE (p. 27), selon lequel les textes personnels et publics se chevauchent avec les textes professionnels, parce qu'il s'agit d'un thème plus spécifique et donc moins accessible aux élèves. Bien que le PISA n'ait pas ignoré ce fait, la différence de textes entre le test papier et le test électronique de la même année 2009 est frappante. Dans les textes relatifs au domaine de l'éducation, entre les épreuves sur support papier et électronique, par exemple, il existe une différence de 10 points de pourcentage (respectivement 25% et 15%). Cette situation se répète lorsque l'on compare les lectures à usage *public*, puisque la version électronique présente 10 points de pourcentage d'items en plus que le test imprimé. On peut en déduire qu'il n'est pas possible d'établir un critère commun relatif à la quantité d'items pour chaque situation de lecture lors des différentes épreuves.

En considérant que ces choix jouent directement sur le degré de difficulté des épreuves de compréhension, nous pensons que le fait d'avoir des contenus différents d'un instrument à l'autre (et de la version papier à l'électronique dans une même édition) affecte les résultats des élèves évalués. Cette constatation affaiblit la comparaison faite entre les performances d'un pays d'une édition à l'autre, ainsi que le classement établi entre les différents pays participants.

9 Les tableaux de distribution des types de textes et des structures des items ne sont pas présentés dans cette recherche car seules les informations du PISA 2000 sont disponibles en la matière.

## Une évaluation unique pour différents publics et contextes

En continuant l'étude comparative des instruments PISA, appliqués en 2000 et 2009, et leur contrepoint avec le SAEB, nous poursuivons l'analyse avec un regard plus qualitatif, en nous concentrant sur les unités rendues publiques par l'OCDE qui constituent également le corpus documentaire de cette recherche. À partir du cadre d'évaluation de la lecture du PISA, nous tenterons de décrire comment les questions posées cherchent à mesurer les compétences des élèves. Nous avons tout d'abord vérifié que les textes qui composent l'évaluation sont représentatifs de divers genres du discours, ce qui est similaire aux orientations brésiliennes. Concernant l'édition du PISA 2000, l'échantillon considéré dans cette étude comprenait : un graphique, une brochure, une carte, un diagramme, un rapport, un article, un formulaire, une facture, une chronique, un texte théâtral et un éditorial. Pour l'édition 2009, on a utilisé un graphique, un plan (d'une bibliothèque), un roman, une carte, un article, un conte, une chronique, un avertissement public et un roman. Cependant, les questions portaient sur des concepts liés à l'engagement en lecture et à la métacognition, sans prendre en compte la conception de la littératie en tant que pratique sociale, comme c'est le cas dans l'approche brésilienne.

Il faut souligner que, bien que l'évaluation du PISA n'utilise pas explicitement les terminologies de la théorie des genres discursifs, certaines questions explorent la fonctionnalité du texte (à quoi sert-il ?) et demandent, en plus, d'identifier des informations explicites et implicites dans le texte, de lire des graphiques et des images, d'identifier le thème, de connaitre le sens des mots ou des expressions, de construire et de distinguer les faits de l'opinion. Ces compétences sont également évaluées dans le SAEB.

Les divergences entre le SAEB et le PISA concernent des questions ouvertes qui n'existent que dans le PISA. Les thématiques présentées par le SAEB sont construites sous l'angle des genres qui circulent au Brésil et l'épreuve de lecture est basée sur une perspective discursive, comme activité sociale, explorant des questions liées au genre (support, but, compréhension de l'énoncé, ressources expressives, cohésion et cohérence etc.).

L'unité « Macondo », par exemple, rendue publique lors de l'édition 2009 et ayant fait l'objet de recherches présentées dans ce dossier, est construite à partir d'un extrait d'un roman plus vaste qui, à partir d'un langage littéraire, plutôt allégorique, évoque un fait (l'arrivée du cinéma dans un village) qui n'est pas familier pour de très nombreux contextes sociaux à travers le monde. On peut faire un commentaire identique à propos de l'unité « Métrotransit » (OCDE, 2012), basée sur un plan, pouvant provoquer de l'étrangeté et interférer négativement dans la construction de sens des élèves qui ne connaissent pas le métro, un fait assez ordinaire dans la réalité brésilienne. Cette constatation interroge l'utilisation d'un même texte pour évaluer 80 pays aux réalités sociales, économiques et culturelles aussi diverses. Même si cela n'est pas explicite dans le guide de correction du PISA, l'existence de connaissances contextuelles antérieures est une condition pour que le lecteur puisse *interpréter, intégrer, réfléchir et évaluer* des informations dans un texte.

Les différents sujets abordés dans les unités du PISA (technologies, changements sociaux, environnement, modèles gouvernementaux, mythologie, etc.), ainsi que dans les évaluations du SAEB, sont indépendants et ne font l'objet d'aucun dialogue entre eux. Les instruments appliqués dans les PISA 2000 et 2009 portent sur des thématiques qui ne correspondent pas aux pratiques de lecture existantes dans de nombreuses écoles brésiliennes, dans la mesure où ils fournissent des exemples et des informations sur la réalité d'autres pays, tels que le « Plan d'une bibliothèque », « Les plus hautes tours » (une publication du magazine *NoRwegian* sur les plus hautes tours du monde), « Destination : Buenos Aires » (un extrait de *Vol de Nuit* de Saint-Exupéry), « Metrotransit » (un plan de transport urbain) etc. (OCDE, 2012). Ces sujets s'éloignent des discussions présentes dans de nombreuses écoles brésiliennes et compliquent la tâche des élèves pour établir des liens et complexifient plus encore l'évaluation de la compétence en lecture testée par le PISA.

Cette distance entre le lecteur et la thématique interfère dans les inférences réalisées pour la construction du sens, facteur nécessaire à la compréhension en lecture, selon des spécialistes tels que Koch et Travaglia (2007). Paulo Freire (1989) aborde également cette discussion à travers l'expression « lecture du mot monde », où l'auteur considère

comme nécessaire l'interprétation, la réalisation d'inférences et le positionnement pour qu'il y ait de la lecture. Cela signifie que quelle que soit la perspective de lecture, la connaissance du contexte est essentielle à la construction de sens.

Il faut aussi dire que les évaluations du PISA (2000 et 2009) comportent des thématiques qui supposent un lecteur vivant dans les grands centres urbains comme on le voit avec les unités « Bibliothèque », « Les plus hautes tours », « Le Trekking africain », « Métrotransit » et « Macondo » (OCDE, 2012). Ce sont des textes qui, en fonction des contextes, peuvent s'éloigner de l'expérience des élèves, car ils témoignent de la réalité de ceux qui vivent dans les grandes villes, connaissent les grandes bibliothèques, voient des graffitis et tags sur les murs et bâtiments, connaissent le cinéma, les itinéraires de voyage, le métro, etc. Ils sont donc capables de relier ces contextes sociaux quotidiens au sujet traité dans les questions. Cela ne veut pas dire que l'élève en milieu rural n'est pas en mesure d'interpréter des textes de cet ordre, mais si cette évaluation des compétences en lecture comprenait des textes sur des réalités proches des échanges dans les écoles brésiliennes et des connaissances partagées (Koch & Travaglia, 2007), cela pourrait favoriser une augmentation du taux de réussite.

Il existe néanmoins, dans les questions du PISA 2000 et 2009, des textes dont les connaissances sont plus proches de la réalité des élèves brésiliens et qui traitent d'activités physiques, des chaussures appropriées, des annonces au supermarché, etc. Ces textes circulent dans la vie quotidienne de la population et sont présents dans les discussions en classe, ce qui favorise l'interprétation et la maitrise de la lecture.

Les données et réflexions présentées ici renforcent nos questions sur la possibilité réelle qu'un même instrument puisse évaluer la lecture de réalités aussi différentes, avec l'objectif clair de la mondialisation des modèles scolaires, piloté par des pays « développés ». Cela fait écho aux critiques formulées dans d'autres études sur le PISA, comme celles signalées par Bart et Daunay (2018), Carnoy *et al.* (2015) et Reigota (2009). En effet, évaluer des compétences en lecture au niveau global n'est pas une tâche facile. Toutefois, tel étant l'objectif affiché de l'OCDE, il est nécessaire de discuter son instrument d'évaluation,

principalement à partir des arguments présentés dans les approches critiques du PISA.

## Considérations finales

L'analyse comparative entre le PISA et le SAEB, ainsi qu'entre deux éditions du PISA dans lesquelles la lecture a été le domaine majeur évalué, a été la préoccupation première de cette étude, motivée par la perception que nombre d'enseignants n'étaient pas au courant de des démarches et des épreuves elles-mêmes. Ainsi, cet article pourrait clarifier les propositions sous-jacentes aux évaluations des compétences en lecture, en ce qui concerne les critères adoptés, la composition globale des instruments et les compétences évaluées.

La diffusion des données du test du PISA, dont les rapports écrits gagnent en volume et en complexité, semble faire apparaitre une transparence qui, en réalité, peut être discutée. La complexité des tableaux de lecture, destinés aux spécialistes, rend peu accessible ce test qui, au fil des ans, a été transformé en un instrument doté d'un pouvoir de classement et de distinction des bons et des mauvais.

Quand on analyse les tests du PISA, leurs cadres respectifs et les rapports disponibles, et les épreuves et les cadres du SAEB, on observe que les aspects qui déclinent les compétences dans le PISA (*localiser et extraire ; intégrer et interpréter ; réfléchir et évaluer*), sont similaires aux six objets des évaluations du SAEB (même si ces dernières n'incluent pas la compétence à évaluer les textes). Bien que ce constat conduise à une impression première de proximité, l'analyse de la matérialité des tests montre plutôt des différences dans leur forme et leur contenu. Plusieurs des textes choisis dans le PISA traitent ainsi de thématiques éloignées de la réalité des élèves brésiliens, ce qui nuit à la construction d'inférence et, par conséquent, à des interprétations appropriées.

En continuant l'analyse, nous avons identifié des aspects qui distinguent les deux épreuves, tels que la conception de lecture adoptée, l'approche du texte (marquée par la présence du genre du discours dans le SAEB et par le type de texte dans le PISA) et la composition des tests (en quantité de questions, dans le SAEB, et en heures, dans le PISA).

Concernant les différences, un autre point à considérer touche aux variations existantes entre les deux éditions du PISA 2000 et 2009, par

rapport aux critères de choix des textes. Les informations analysées ont indiqué l'absence d'un critère commun relatif à la quantité d'items composant les différentes épreuves, en soulignant la différence quantitative qui existe entre les versions imprimées et numériques de la même édition du PISA 2009. Si on considère que cette composition est déterminante pour le niveau de difficulté du test de lecture, on peut penser que les résultats du PISA sont fragiles tant pour comparer des performances d'un pays d'une édition à l'autre, que pour classer les performances des différents pays participants.

De ce fait, nous interrogeons la faisabilité de la méthodologie du PISA, et non les aspects liés aux statistiques ou à la démarche d'élaboration des items, ni même à la définition des compétences évaluées dans le test. Ce que nous interrogeons concerne la possibilité de s'appuyer sur un même test pour évaluer les élèves de 80 pays, qui ont des réalités sociales, économiques et culturelles très différentes, avec un instrument identique, supposant des connaissances préalables communes et standardisées. En partant de la conception du langage en tant qu'activité interactionnelle, dans laquelle le contexte est essentiel à la construction des sens, l'idée qu'une telle diversité d'élèves puissent *interpréter, intégrer, refléchir et évaluer* les informations contenues dans un texte devient discutable.

Pour conclure, il convient de réfléchir à l'utilisation des résultats du PISA. Le Ministre de l'Éducation du Brésil, Mendonça Filho, a ainsi utilisé les données de l'édition 2015, publiées par l'INEP (2016), comme argument en faveur de changements[10] dans l'enseignement au lycée. Or, si la moyenne globale du Brésil n'était pas satisfaisante, l'observation des résultats non agrégés permet de voir de grandes différences à l'intérieur du pays lui-même. Alors que le score moyen en lecture des écoles du réseau municipal était de 325 points, celui du réseau fédéral a atteint 528 points (INEP, 2016, p. 131), supérieur donc à celui de Hong Kong et du Canada (527 points) et de la Finlande (526 points). Ce même

10 Les propositions libérales du gouvernement pour l'enseignement au lycée au Brésil, entre autres, ont divisé le programme commun en cinq domaines, dont seulement deux composantes (langue portugaise et mathématiques) sont obligatoires ; en outre, ces propositions privilégient l'enseignement technique et permettent de s'acquitter à distance d'une partie de la charge de travail.

réseau, qui est maintenant menacé de disparition, est par conséquent capable de placer le Brésil au premier rang du classement en lecture, si telle est l'intention des dirigeants. Ces données qualitatives et quantitatives semblent mettre encore en évidence qu'il y a un enjeu majeur à questionner les évaluations à grande échelle, telles qu'elles sont construites aujourd'hui, en tant qu'instruments de pouvoir qui, ignorant les différentes réalités et contextes coexistants au niveau local ou international, visent à mesurer le savoir de manière standardisée.

## Références bibliographiques


ANDRADE, F.; TAVARES, H. R. & VALLE, R. C. (2000). *Teoria de Resposta ao Item: Conceitos e Aplicações*. En ligne : https://www.researchgate.net/publication/228593873_Teoria_da_Resposta_ao_Item_conceitos_e_aplicacoes.

BAKHTINE, M. (1984). *Esthétique de la création verbale*. Paris : Gallimard.

BART, D. & DAUNAY, B. (2018). *Pode-se Levar a Sério o PISA? O tratamento do texto literário em uma avaliação internacional*. Campinas SP: Mercado de Letras.

BRASIL, MEC/CEB. (1997). *Parâmetros Curriculares Nacionais*. Brasília: CEB.

BRASIL. MEC/CEB. (2018). *Base Nacional Comum Curricular*. Brasília: CEB.

BRASIL, MEC (2018). *Matriz de Referência da Prova Brasil*. En ligne : http://www.seed.se.gov.br/arquivos/Matriz_Mat_3o_ano_EM_PROF.pdf

CARNOY, M. et al. (2015). A Educação Brasileira Está Melhorando? Evidências do PISA e SAEB. *Cadernos de Pesquisa*, vol.45 n°157. En ligne : http://dx.doi.org/10.1590/198053143331

FREIRE, P. (1989). *A Importância do Ato de Ler: em Três Artigos que se Complementam.* (21. ed.). São Paulo: Autores Associados, Cortez.

GOMES, E. M. L. (2006). Avaliação de Língua Portuguesa do Saeb: da Leitura ao Letramento. Estudos em Avaliação Educacional, vol.17, n° 34.

KOCH, I. V. & TRAVAGLIA, L. C. (2007). *Texto e coerência*. São Paulo: Cortez.

INEP (2001). *PISA 2000: Relatório Nacional. Apresentação*. Brasília.

En ligne : http://portal.inep.gov.br/web/guest/acoes-internacionais/pisa/resultados

INEP (2012). *PISA: Resultados Nacionais – PISA 2009*. Brasília: O Instituto.

INEP (2016). *Brasil no PISA 2015: Análises e Reflexões sobre o Desempenho dos Estudantes Brasileiros*. Brasília. En ligne :

http://download.inep.gov.br/acoes_internacionais/pisa/resultados/2015/pisa2015_completo_final_baixa.pdf.

INEP (2017). *Censo Escolar da Educação Básica (MEC)*. En ligne : http://download.inep.gov.br/educacao_basica/censo_escolar/notas_estatisticas/2018/notas_estatisticas_Censo_Escolar_2017.pdf.

INEP (2019). Portaria 366, de 29 de abril de 2019. *Estabelece as diretrizes de realização do Sistema de Avaliação da Educação Básica* (SAEB).

INEP (2019). *Nota Técnica nº 10/2019/CGIM/DAEB, de 02 de julho de 2019. Detalhamento da população e resultados do SAEB 2019*. Brasil.

JOU, G. I. & SPERBE, T.M. (2003). Leitura Compreensiva: um Estudo de Caso. *Linguagem & Ensino*, nº 6, 13-54.

LEFFA, J. (1996). *Aspectos da Leitura: uma Perspectiva Psicolinguística*. Porto Alegre: Sagra Luzzatto.

LOPES-ROSSI, M. A. G. & PAULA, O. (2012). As Habilidades de Leitura Avaliadas pelo PISA e pela Prova Brasil: Reflexões para Subsidiar o Trabalho do Professor de Língua Portuguesa. *Fórum Linguístico*, Florianópolis, vol. 9, n° 1, p. 34-46. En ligne : https://periodicos.ufsc.br/index.php/forum/article/download/1984-8412.../22551

MARCUSCHI, L. A. (2002). Gêneros Textuais: Definição e Funcionalidade. In DIONÍSIO, Â. P., MACHADO, A. R., BEZERRA, M. A. (dir.). *Gêneros Textuais & Ensino*. Rio de Janeiro: Lucerna.

OCDE (2000). *Mesurer les Connaissances et les Compétences des Élèves. Lecture, Mathématiques et Science : l'évaluation de PISA 2000*. PISA, Éditions OCDE. En ligne : http://dx.doi.org/10.1787/9789264281561-fr

OCDE (2001). *Connaissances et Compétences : des Atouts pour la Vie. Premiers Résultats du Programme International de l'OCDE PISA 2000*. PISA, Éditions OCDE. En ligne : http://dx.doi.org/10.1787/9789264295902-fr

OCDE (2011). *Résultats du PISA 2009 : Savoirs et savoir-faire des élèves : performance des élèves en compréhension de l'écrit, en mathématiques et en sciences (Volume I)*. Paris : PISA, Éditions OCDE. En ligne : http://dx.doi.org/10.1787/9789264097643-fr

OCDE (2012). *Le Cadre d'Évaluation de PISA 2009. Les compétences clés en compréhension de l'écrit, en mathématiques et en sciences*. PISA : Éditions OCDE. En ligne : http://dx.doi.org/10.1787/9789264075474-fr

OCDE (2013). *Le Cadre d'Évaluation de PISA 2012*. PISA : OCDE Éditions. En ligne : http://dx.doi.org/10.1787/9789264190559-fr

OCDE (2014). *Résultats du PISA 2012: Savoirs et savoir-faire des élèves : Performance des élèves en mathématiques, en compréhension de l'écrit et en sciences, série I*. PISA : Éditions OCDE. En ligne : http://dx.doi.org/10.1787/9789264208827-fr.pdf

OECD (2009). *Learning Mathematics for Life: A Perspective from PISA*. PISA: OECD Publishing. En ligne : http://dx.doi.org/10.1787/9789264075009-en

OECD. (2016). *Draft Analytical Frameworks - Pisa 2018*. PISA: OECD Publishing. En ligne : https://www.oecd.org/pisa/data/PISA-2018-draft-frameworks.pdf

SAEB/INEP. (2019). *Matrizes e Escalas*. En ligne : http://portal.inep.gov.br/web/guest/educacao-basica/saeb/matrizes-e-escalas

QUEIROZ, P. A. A. (2013). *Concepções de Letramento que Respaldam as Avaliações do Saeb e Prova Brasil.* 2013. 204 f. Brasília: Université de Brasília. Mémoire de Maitrise.

REIGOTA, M. (2009). Les élèves brésiliens ne savent-ils donc rien ?, *Revue internationale d'éducation de Sèvres*, Actes du Colloque 2009. En ligne : http://journals.openedition.org/ries/5645