

## Évaluer le savoir lire

Jacques Aubret

---



### Édition électronique

URL : <http://journals.openedition.org/ries/4254>

DOI : 10.4000/ries.4254

ISSN : 2261-4265

### Éditeur

Centre international d'études pédagogiques

### Édition imprimée

Date de publication : 1 juin 1994

Pagination : 31-38

ISSN : 1254-4590

### Référence électronique

Jacques Aubret, « Évaluer le savoir lire », *Revue internationale d'éducation de Sèvres* [En ligne], 02 | 1994, mis en ligne le 17 avril 2015, consulté le 20 avril 2019. URL : <http://journals.openedition.org/ries/4254> ; DOI : 10.4000/ries.4254

---

Ce document a été généré automatiquement le 20 avril 2019.

© Tous droits réservés

---

# Évaluer le savoir lire

Jacques Aubret

---

- 1 De nombreux problèmes éducatifs se cristallisent autour de l'apprentissage de la lecture. Paradoxalement, les évaluations qui sont faites des acquis scolaires en matière de lecture sont parfois contestées par ceux-là mêmes qui s'appuient sur ces constats pour proposer des remédiations. Toutefois, personne ne rejette vraiment l'idée de l'utilité de l'évaluation des acquis en lecture, ni le politique, ni l'économiste, ni le pédagogue, ni le psychologue, ni le parent d'élève, ni l'élève lui-même. Chacun des acteurs concernés par ce problème a besoin, en effet, des points de repères que l'évaluation est censée lui fournir. Mais ces points de repères ne sont pas nécessairement les mêmes pour tous ces acteurs. De là naît souvent l'ambiguïté du discours sur l'évaluation de la lecture.
- 2 L'objectif de cet article est d'énoncer les raisons qui peuvent nous conduire à porter un regard critique sur les données résultant d'évaluations de la lecture. Pour ce faire, il nous a semblé nécessaire d'identifier au préalable les facteurs de diversification des pratiques. Le terme évaluation correspond déjà, dans les faits, à une diversité d'opérations : attribution d'une valeur à un « objet », reconnaissance d'un degré de conformité à des normes, identification de la position d'un objet dans un ensemble d'objets de référence. Pour aller plus loin, trois plans d'analyse peuvent être distingués : le plan des valeurs et des normes, celui des procédures qui permettent d'isoler des faits sur lesquels on peut porter des jugements de valeur, celui des processus par lesquels un acteur social énonce un jugement. Les facteurs de diversification des pratiques d'évaluation doivent être rapportés à chacun de ces trois niveaux. Nous commencerons par les aspects les plus « visibles », les procédures.

## La diversité des procédures d'évaluation de la lecture

- 3 En désignant le « savoir lire » comme objet à évaluer on évoque un objet dont les dimensions varient selon les contextes théoriques ou les définitions empiriques invoquées. Chaque outil ou pratique d'évaluation est une définition de cet objet sur la base d'une sélection d'« observables » et de l'application de procédés de mesure.

## La diversité des observables

- 4 La lecture est toujours appréhendée à partir d'observables interprétés comme indices, indicateurs, traces, témoignages d'un niveau, d'une caractéristique ou de la globalité d'un comportement ou de la performance d'un individu en matière de savoir lire. Qu'il s'agisse d'oral et d'écrit, ces observables sont nombreux et divers. On citera pêle-mêle : la durée, le rythme ou la vitesse de lecture, les pauses, le mouvement des yeux, les erreurs de prononciation, les hésitations, les reprises, l'intonation, le fait de restaurer un texte que l'on a préalablement amputé de certains mots ou dont les mots ont été soudés (segmentation) ou transformés (mots tordus), le fait de répondre à des questions portant sur un texte, de « catégoriser » un écrit en fonction d'une typologie des écrits. Des exemples sont donnés dans Aubret et Blanchard (1991) et Aubret et Dosnon (1993).
- 5 Ces différents observables ne renvoient pas aux mêmes dimensions de l'acte de lecture : les uns sont proches de caractéristiques perceptibles de l'activité de lecture, d'autres prennent en compte ses finalités (ce pour quoi il y a lecture, c'est-à-dire comprendre ou s'exprimer), d'autres enfin sont considérés comme relevant de compétences que l'on suppose être en amont de la capacité à lire. Ces observables ne sont pas également accessibles à l'observation, certains nécessitant la mise en œuvre de procédures d'observation plus sophistiquées que d'autres (il est techniquement plus difficile d'observer le mouvement des yeux que de repérer des erreurs de prononciation). Les conditions dans lesquelles ces observations sont faites peuvent varier (on peut tester la compréhension à travers des réponses libres ou un QCM). Ces observations ne sont pas également adaptées à tous les publics. Mais dans tous les cas les relations que l'on peut établir entre les observables et l'objet à évaluer ne se justifient que par des prises de position théoriques ou des vérifications empiriques qu'il sera nécessaire d'explicitier dans tout acte d'évaluation pour en comprendre le sens. Il n'y a par exemple aucun lien définitivement établi entre le fait de réussir une épreuve dite de *closure*<sup>1</sup> et le fait de savoir lire. Autrement dit, toute épreuve de lecture devrait être « éprouvée » (ou validée).
- 6 Mais le problème ne s'arrête pas à la variété des « observables ». Le plus souvent la pratique de l'évaluation conduit les évaluateurs à catégoriser et/ou combiner des observables : pour établir des scores, distinguer et hiérarchiser des niveaux de performance. Se pose alors le problème de la mesure et de sa signification.

## Les artifices de la mesure

- 7 Mesurer c'est attribuer les propriétés des nombres à des observables qui n'ont pas nécessairement ces propriétés. On compte le nombre de mots lus par minute pour définir une vitesse de lecture, on dénombre des fautes d'orthographe pour apprécier le degré de connaissance des règles orthographiques, on constitue des scores de performance en compréhension pour apprécier le savoir lire, on totalise des réussites à des exercices variés de nature différente (comme dans les enquêtes de la direction de l'évaluation et de la prospective du ministère de l'éducation nationale) pour évaluer des acquis. Des mesures portées sur des objets différents peuvent être combinées pour résumer plusieurs informations par une seule donnée. Dans une épreuve de lecture connue, on combine un indicateur de compréhension et un indicateur de vitesse de lecture pour construire une mesure de l'efficacité. Chaque fois se pose le problème de la signification des indications

quantitatives obtenues : scores équivalents obtenus de façon différente, signification des différences entre les scores (l'égalité des intervalles postulée par l'établissement d'une note ou d'un score n'est pas une propriété des objets mesurés).

## Au-delà de la mesure, la diversité des normes et des échelles de valeur

- 8 La mesure n'a pas sa raison d'être en elle-même. Elle est un intermédiaire parfois utile pour situer les données d'une observation dans des ensembles de caractéristiques plus ou moins organisés, structurés, hiérarchisés. En ce qui concerne la lecture, ces ensembles ont des noms et des formes variés : programmes d'apprentissage comme ensembles de contenus à maîtriser, populations de référence comme ensembles d'individus dont on connaît la distribution des scores sur des échelles de performances ou de comportements, modèles théoriques comme ensembles d'éléments permettant de décrire le développement, l'organisation et le fonctionnement du lecteur dans l'activité de lecture, ensembles de critères permettant de définir des niveaux de performance, ensembles de valeurs attribuables aux lecteurs en fonction des performances et des comportements manifestés, ensembles d'indicateurs quantitatifs ou qualitatifs servant à décrire les propriétés des textes à lire (indice de lisibilité par exemple), etc. Selon les cas, on parlera de « référentiel » (ensemble de normes, de contenus, de critères), d'étalonnage, d'échelle de valeur.
- 9 Ces ensembles ne sont pas produits par l'acte singulier d'évaluation, mais font l'objet d'appropriation par l'évaluateur. Il n'est donc pas indifférent pour comprendre la signification d'une évaluation de s'interroger sur ces ensembles qui servent de référence et sur les conditions dans lesquelles ils sont produits. En effet, aux questions du type : « Qu'est-ce qu'un bon lecteur ? » ou « Qu'est-ce qu'un illettré ? » diverses réponses sont possibles en raison même de la diversité des points de vue que l'on peut développer sur la lecture. Ainsi le pédagogue, le psychologue, le chercheur, le politique, le journaliste, le non-spécialiste situés différemment par rapport à la lecture ont-ils des raisons de ne pas invoquer les mêmes normes de jugements ou de valeur, de ne pas utiliser les mêmes référentiels.
- 10 Le pédagogue emprunte naturellement aux programmes d'apprentissage (ou aux séquences qui constituent sa progression) les points de repère essentiels par rapport auxquels il peut situer les acquis ou les progrès de ses élèves. Il puise également dans sa propre expérience les éléments modérateurs qui permettent de relativiser l'impact de telle ou telle observation. Toutefois la valeur informative de ces jugements demeure limitée à ceux qui suivent rigoureusement la même progression, qui ont la même conception des apprentissages ou qui, d'une manière ou d'une autre, sont impliqués par la relation pédagogique ou éducative (élève, parent d'élève).
- 11 Le psychologue à qui l'on confie le soin de comprendre les difficultés d'apprentissage ou de trouver des moyens de remédier à des dysfonctionnements langagiers cherchera à établir son diagnostic en se servant d'outils d'évaluation standardisés à l'aide desquels il peut produire des comparaisons entre ce qu'il observe et ce que l'on peut observer sur des populations de sujets dont il connaît les caractéristiques (principe des étalonnages) ou en empruntant aux chercheurs spécialisés dans l'étude des processus d'apprentissages ou de fonctionnement de la lecture les éléments susceptibles de rendre compte de ces

observations par rapport au modèle théorique invoqué. La valeur informative de l'évaluation dépend donc de la connaissance que l'on peut avoir de la population de référence, ou de l'intelligibilité des modèles théoriques élaborés par la recherche.

- 12 Ceux qui partagent des responsabilités régionales ou nationales impliquant des comparaisons entre élèves ou entre institutions, ou des mesures d'efficacité (responsables pédagogiques, responsables politiques, économistes, contribuables) cherchent à recueillir des informations comparables. Des échelles de mesure, le plus souvent standardisées dans leur mode de passation et de correction sont élaborées dans ce but. Ainsi en est-il des évaluations effectuées en France par la direction de l'évaluation et de la prospective (DEP) du ministère de l'éducation nationale ou par l'International Association for the Evaluation of Educational Achievement (IEA). Mais les résultats de ces enquêtes n'ont de signification que par référence aux choix qui ont présidé à la construction des épreuves.
- 13 Par exemple, la dernière enquête de l'IEA qui a concerné trente-deux systèmes scolaires différents et s'est déroulée, pendant l'année scolaire 1990-1991, sur deux échantillons représentatifs des élèves de 9 à 10 ans (en France, niveau CM1) et des élèves de 14 à 15 ans (en France, niveau troisième) a privilégié une forme de mesure de la compréhension en lecture parmi d'autres : le questionnement sur des textes lus silencieusement. En revanche, les enquêtes (ou les évaluations) de la DEP, depuis les évaluations effectuées au CP en 1978, ont souvent privilégié la constitution de scores établis sur des réussites à des items diversifiés dont les contenus sont empruntés aux programmes d'enseignement de la langue et les formes de questionnement aux modalités habituelles d'évaluation écrite des professeurs des écoles.
- 14 À ces évaluateurs « savants » on peut opposer en quelque sorte « le non-spécialiste » qui distingue le lecteur de l'illettré sur sa manière de se servir des écrits dans les situations de la vie sociale habituelle : une manière de voir simple, comparable à la notion d'« analphabétisme fonctionnel » introduite par l'UNESCO, en 1960, à la Conférence mondiale pour l'éducation des adultes de Montréal et qui désigne ainsi l'incapacité pour un individu de lire ou d'écrire, en le comprenant, un exposé simple et bref de faits en rapport avec sa vie quotidienne. Le critère est certes simple. Mais peut-on faire l'inventaire de ces situations de la vie quotidienne et doit-on considérer que la capacité à lire est la seule capacité en cause dans l'échec à l'écrit ?

## La variabilité du comportement des évaluateurs dans le processus d'évaluation

- 15 La compréhension des phénomènes attachés à l'évaluation nous mène jusqu'à l'observation des comportements de chaque évaluateur en situation d'évaluer. À ce niveau, chaque acteur a sa manière propre de concevoir l'acte de lecture, ses attentes face au lecteur qu'il observe, sa manière de définir et de prendre en compte la « difficulté », sa conception du rôle et des modalités de l'acte d'évaluer.
- 16 D'une manière générale, le comportement évaluatif peut se décrire dans le cadre des conduites adaptatives de l'individu fondées sur le recueil et le traitement d'informations de l'environnement et leur interprétation en vue de l'action. Les normes et les échelles de valeur fournissent le cadre interprétatif ; les observables constituent la base des données recherchées. Toutefois, sur ces deux aspects, prise d'informations et interprétation, une large part doit être faite aux influences que subit l'évaluateur dans l'élaboration de son

jugement. Les études conduites dans le cadre de la docimologie montrent la variabilité « inter » et « intra » individuelle des comportements en matière d'évaluation (Noizet, Caverni, 1978). Des correcteurs différents donnent des appréciations divergentes sur les mêmes copies. Un même correcteur peut varier dans ses manières de noter une même copie selon les conditions dans lesquelles la copie est appréciée et selon les informations dont il dispose au moment de la correction. On ne peut pas penser que l'évaluation du savoir lire échappe à cette variabilité.

- 17 Pour comprendre cette variabilité, et y remédier, des recherches sur la modélisation du comportement de l'évaluateur sont nécessaires. Dans le champ de la lecture, on peut être tenté d'établir un parallélisme entre le comportement d'un lecteur et celui de l'évaluateur. Dans la compréhension de l'écrit l'activité du lecteur est guidée par des informations d'origines diverses : celles qui émanent du texte que le lecteur analyse à différents niveaux (phonémique, graphémique, syntaxique, sémantique) et celles qui proviennent du contexte, qui oriente et détermine l'analyse du texte. L'élaboration du sens qui en résulte témoigne du type d'informations privilégié par le lecteur et parfois aussi de son mode d'entrée dans le texte (modèles *top-down* et *bottom-up*<sup>2</sup>). En ce qui concerne l'évaluation, les comportements des évaluateurs peuvent également se distinguer par l'importance relative accordée aux différentes opérations par lesquelles un jugement de valeur s'élabore : on peut ainsi se focaliser sur l'établissement des « faits » avant toute activité de jugement (construction d'observables, d'échelles de mesure) ou aborder les faits avec un cadre d'analyse prédéterminé par des attentes relatives au comportement ou aux performances du lecteur (projection sur les faits des valeurs véhiculées ou des normes dictées par le contexte social ou pédagogique). D'un côté on cherche à décrire le comportement du lecteur selon toutes les dimensions possibles, de l'autre on contrôle la correspondance entre ce qui est observé et ce qui définit le « bon lecteur » pour un milieu donné. Si l'évaluation est le produit de ces deux activités le comportement d'un évaluateur n'est pas dicté d'avance. Des variations sont possibles en fonction des prises de position et des savoir-faire de l'évaluateur ainsi que des influences qu'il subit. Mais comment alors comprendre le discours sur l'évaluation si l'on ne sait rien de l'évaluateur ?

## Du bon usage de l'ambiguïté

- 18 Qu'il s'agisse d'évaluation individualisée (l'enseignant dans sa classe) ou d'évaluation collective (les enquêtes nationales ou internationales), d'évaluateurs impliqués dans le processus d'apprentissage de la lecture ou d'acteurs non directement impliqués (le psychologue par exemple), chaque évaluation est toujours située par rapport à des objectifs, des intentions (formation, régulation, diagnostic, certification), des circonstances et relève donc d'une logique propre qu'il faut saisir pour comprendre le discours évaluatif qui en est issu.
- 19 Face à un apprentissage aussi complexe que celui de la lecture (que l'on ne réduit pas à la maîtrise d'une combinatoire de lettres, de mots, de phrases) et parce qu'il mobilise des compétences cognitives plus générales que celles qui sont impliquées dans la reconnaissance du code linguistique et nécessite un engagement personnel (une motivation à lire), on ne peut espérer pouvoir énoncer sans nuance qu'un élève « sait lire » alors que son voisin « ne sait pas lire ». Toute tentative visant à privilégier une dimension de la maîtrise de la lecture (exemple : on sait lire quand on « comprend », on

sait lire quand on lit à une vitesse déterminée) est réductrice. Mais l'addition pure et simple d'une pluralité de dimensions (par exemple pour constituer un score sur des épreuves ayant un rapport apparent avec la lecture) relève de l'artifice du calcul, davantage producteur de comparaisons que de sens.

- 20 Comment alors répondre aux besoins d'information du public ? Si informer le public c'est confirmer ses attentes, le problème ne relève pas d'un article sur l'évaluation de la lecture. Plusieurs enquêtes récentes sur la lecture (à condition de les sortir de leur contexte) ont été interprétées dans un sens favorable ou défavorable. Un jour la France est « bonne élève » (*Le Monde*, 16 octobre 1992). Un autre jour on s'interroge sur l'ampleur de l'« illettrisme » pour constater qu'une affirmation du genre : « Un tiers des élèves de sixième ne savent pas lire » n'est confirmée par aucune enquête sérieuse sur le sujet (*Le Monde*, 6 mai 1993) et l'on crie « Halte au trucage ! » (C. Garin, *Le Monde*, 15 octobre 1993).
- 21 De même, si l'information consiste à transformer un problème complexe en un problème simple pour répondre aux exigences supposées de la communication, alors il suffit de sortir les « chiffres » de leur contexte pour donner l'illusion que l'évaluation est objective et absolue (le chiffre évoque, aux yeux de beaucoup, le scientifique, le sérieux, le vrai).
- 22 En matière de lecture comme en matière d'acquis scolaires, il ne nous paraît pas possible de dissocier l'information sur les données recueillies d'une information rigoureuse sur les conditions de recueil de l'information (situations, intentions, acteurs, procédures, conceptions de la lecture). Des clarifications s'imposent. Toutefois celles-ci n'ont de sens que si les évaluations et leurs interprétations sont conduites selon la même rigueur. Les études de docimologie et de psychométrie (des épreuves de lecture ont été produites dans ce contexte) nous ont habitués à reconnaître l'importance des notions de fidélité et de validité. Ces concepts s'appliquent aux évaluations de la lecture comme des moyens nécessaires de contrôle des outils et des pratiques. Il faut aller au-delà en s'interrogeant sur leur cohérence et leur pertinence dans un processus éducatif, c'est-à-dire sur la nécessaire continuité de la signification entre le choix des observables, le recours à des procédés de mesure, et le champ interprétatif invoqué (cohérence) et sur l'adaptation des pratiques aux objectifs visés par l'évaluation (pertinence).

---

## BIBLIOGRAPHIE

AUBRET J., BLANCHARD S. (1991), *L'évaluation des compétences d'un lecteur*, Issy-les-Moulineaux, EAP.

AUBRET J., DOSNON O. (1993), *Les Illettrés et le bilan des compétences personnelles et professionnelles*, rapport présenté à la Délégation régionale à la formation professionnelle de la région Nord-Pas-de-Calais et au Groupe permanent de lutte contre l'illettrisme.

AUBRET J. (1993), *À propos d'une enquête internationale sur la lecture ; La lettre de la DFLM*, n° 12, 7-9.

NOIZET G., CAVERNI J.-P. (1978), *La psychologie de l'évaluation scolaire*, Paris, PUF.

## NOTES

1. Les tests de *closure* proposent des textes volontairement tronqués d'un certain nombre de mots. Ils sont utilisés pour évaluer le degré de performance des lecteurs.
  2. Le modèle descendant (*top-down*) mise essentiellement sur l'activité cognitive générale du sujet. Il s'appuie principalement sur l'anticipation du sens, l'élaboration et la vérification d'hypothèses. Le modèle ascendant (*bottom-up*) se centre essentiellement sur le traitement des mots. Il propose de partir des niveaux linguistiques les plus bas pour arriver à la compréhension.
- 

## RÉSUMÉS

Cet article pose le problème de l'hétérogénéité des méthodes d'évaluation du savoir lire en faisant apparaître les ambiguïtés qui s'y rattachent. Il montre comment une démarche de clarification cognitive permet de lever ces ambiguïtés.

## INDEX

**Index géographique :** France

**Mots-clés :** apprentissage, didactique, évaluation, lecture, apprentissage de la lecture

## AUTEUR

**JACQUES AUBRET**

Université Charles-de-Gaulle, Lille III