

LIRE LES RECHERCHES SUR LA LECTURE

Jean Foucambert

Les partis pédagogiques qui s'affrontent autour de la lecture tentent aussi de se présenter en hérauts scientifiques. «Que voulez-vous, les faits sont les faits, et on sait aujourd'hui que...». Pour autant, la robustesse d'un fait apparaît moins, actuellement, dans la lisibilité de la démarche qui l'établit que dans la renommée du support où paraît l'article qui le présente. Tout se passe comme si la publication dans une revue internationale à comité de lecture attestait que le dispositif de construction de la preuve avait été vérifié par un comité d'experts. Pourtant, lorsqu'il arrive à ceux-ci de se pencher sur la question de l'évaluation des recherches, ils découvrent ce que révèle l'article de Gary A. Troia (*Les recherches sur les interventions sur la conscience phonologique*, Bonnes Pages, p.61) dont nous donnons de larges extraits : peu de recherches pédagogiques échappent à des biais méthodologiques qui rendent leurs conclusions pour le moins incertaines.

Loin d'en déduire que tout serait affaire d'opinion, il faut que ce souci de construction de la preuve devienne une exigence partagée. En d'autres termes, le professionnel de la recherche doit livrer tous les éléments d'une analyse contradictoire de ses données au professionnel de la pédagogie qui doit (apprendre à) se livrer à une lecture experte de ce qu'on lui propose. Non qu'il faille «recompter» derrière les chercheurs, comme s'ils étaient étourdis ou malhonnêtes, mais parce que les mêmes données peuvent être «lues» de manière différente. Galilée et ses juges étaient au moins d'accord qu'il fait jour à midi.

1. POURTANT, ÇA SE VOIT...

Prenons un exemple banal. Il s'agit d'un tableau de résultats reproduit par L. Sprenger-Charolles dans l'ouvrage collectif (*Psychologie cognitive de la lecture*, PUF, dirigé par Michel Fayol) auquel elle participait. «*La recherche de Content & Leybaert (1989, 1990), y écrit-elle, se situe dans le prolongement de celle d'Alegria, Morais et d'Alimonte. Elle a comme point de départ la critique faite par Seymour (1986) au modèle développemental de Frith (1985). Ce modèle, d'après Seymour, ne tient pas compte de l'effet des méthodes sur l'acquisition de la lecture, en particulier en ce qui concerne le passage de la phase alphabétique à la phase orthographique.*» Le fameux passage de la voie indirecte à la voie directe. Content & Leybaert vont alors comparer des élèves ayant appris à lire dans des classes «phoniques» et des élèves ayant appris dans des classes «globales» en leur demandant de «dire» les mots qu'on leur présente, choisis pour être fréquents ou rares, réguliers ou irréguliers. Le tableau 1 donne le temps moyen en ms et le pourcentage de réponses correctes. On pourrait s'attendre à ce que les mots irréguliers (exemple : femme) soient mieux «dits» dans les classes «globales» (voie directe), surtout s'ils sont fréquents, et les mots réguliers, surtout s'ils sont rares (exemple : véranda), dans les classes «phoniques» (voie indirecte).

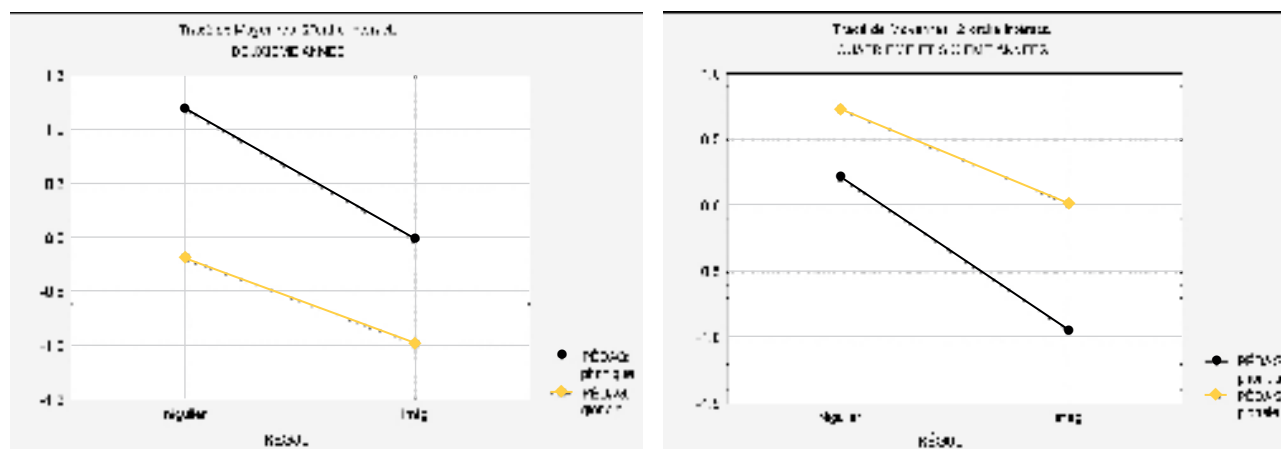
	FRÉQUENT				RARE			
	Régulier		Irrégulier		Régulier		Irrégulier	
Deuxième année phonique	1918	70%	2358	38%	2181	66%	2592	25%
Deuxième année globale	3541	47%	3168	19%	3867	49%	4686	29%
Quatrième année phonique	953	97%	1154	73%	1198	90%	1229	63%
Quatrième année globale	803	97%	930	83%	1015	95%	1020	68%
Sixième année phonique	769	96%	821	77%	872	94%	838	69%
Sixième année globale	613	99%	639	91%	657	93%	687	85%

Tableau 1 : la récapitulation des résultats de Content & Leybaert telle qu'elle apparaît dans la contribution de L. Sprenger Charolles à *Psychologie cognitive de la lecture*, PUF donnant le temps et le % de réussite d'élèves issus de classes «phoniques» et «globales» à la lecture de mots fréquents ou rares, réguliers ou irréguliers.

La démonstration (que je ne discute pas ici) se focalise sur la deuxième année. Les résultats y sont, de manière homogène, inférieurs pour les classes «globales». Content & Leybaert vont en conclure que la méthode «globale» ne semble pas conduire en deuxième année à un développement précoce du lexique orthographique et Sprenger-

Charolles que, quelles que soient les méthodes d'enseignement utilisées, la médiation phonologique occupe une place intermédiaire centrale dans l'acquisition de la lecture. On peut dire cela. Mais personne ne semble préoccupé de ce que le tableau 1 comporte encore 4 lignes au-delà de la seconde année. Dès la quatrième année, et de manière encore plus significative à la sixième, les élèves de la méthode «globale» ont des performances supérieures. Que les mots soient fréquents, rares, irréguliers ou réguliers, ils répondent plus vite (en moyenne 649 ms au lieu de 825) et commettent moins d'erreurs (92 % de réponses correctes au lieu de 84 %), notamment sur les mots irréguliers, fréquents ou rares. Ce temps de réponse qui témoigne d'un accès plus automatisé, plus économique en coût cognitif, quel que soit le mot, mais aussi plus sûr puisqu'il y a moins d'erreurs, devrait alerter les zéloteurs de l'approche «phonique» sur le fait qu'à moyen et/ou long terme, celle-ci ne semble pas conserver les avantages qu'on lui trouve à court terme. Ces résultats font écho, à leur manière, à ceux qu'on rencontre dans l'école où seulement 5 à 6 % des élèves de 6ème connaissent des difficultés du côté de la correspondance grapho-phonologique alors que plus de 80 % n'accèdent pas à l'exercice des compétences remarquables de lecture.

Les résultats massivement médiocres au niveau d'une classe d'âge ne trouveraient-ils pas à s'expliquer, non dans une maîtrise insuffisante du principe alphabétique mais, au contraire, dans l'entrée réussie en lecture sur cette base ? Comment s'en assurer ? Si ces données étaient accessibles autrement que sous la forme d'un tableau de moyennes, le lecteur aurait le loisir de les retravailler à partir de ses propres questions. En «calculant» une performance moyenne à partir de ces moyennes (après les avoir standardisées), on obtient, en effet, 2 courbes inverses selon qu'on considère les résultats des classes «phoniques» et des classes «globales» lors de la deuxième année ou lors des quatrième et sixième années.



Pour autant, cette visualisation ne constitue en rien une preuve, laquelle supposerait d'avoir accès de manière controversiale aux données. Et non seulement aux conclusions des chercheurs.

2. ET QUAND ÇA NE SE VOIT PAS ?

L'exemple précédent reste un cas relativement transparent dans la mesure où, malgré tout, des éléments sont donnés qui permettent au lecteur attentif d'y «trouver» autre chose que ce qui a retenu l'attention du chercheur. Mais, bien plus fréquemment, c'est l'investigation statistique insuffisante qui cache, ou fausse, ce qu'il y aurait d'autre à voir dans les données.

2.1. Un exemple sur des données fictives

Je me servirai de cet exemple uniquement pour rappeler quelques notions de statistique relatives à la «régression», régression simple et régression multiple. Ces gros mots ne doivent pas effrayer : il en va de la «raison statistique» comme de la «raison graphique». C'est affaire de statut, et d'abord dans sa tête.

Le tableau de chiffres que je vais utiliser ici a circulé, il y a plusieurs années, dans une réunion entre chercheurs de l'INRP. Les auteurs d'une recherche avaient extrait 15 observations dans une population expérimentale afin de fournir à leurs collègues un éventail des résultats, des meilleurs aux plus faibles. Il ne s'agit donc pas du tableau général de leurs résultats. En d'autres termes, ce que j'en dirai est «statistiquement vrai» par rapport à un tableau de chiffres dont la complexité, sur un sujet familier, est éclairante pour une présentation technique. Mais, pour autant, rien n'autorise à dire que ces conclusions se retrouveraient sur la population expérimentale totale, à laquelle je n'ai pas eu accès, malgré mes demandes.

Quel était le plan d'expérience ? Des élèves de cours préparatoire sont décrits en début d'année à partir de 9 capacités techniques. En fin d'année, on évalue leur performance en lecture. Les 9 épreuves de début d'année concernent respectivement :

- ◆ la connaissance des supports de l'écrit,
- ◆ la représentation du projet d'apprendre à lire,
- ◆ la familiarité avec le langage technique de la lecture,
- ◆ la possibilité de donner le nom des lettres,
- ◆ le niveau de conscience phonologique,
- ◆ l'habileté à résoudre des devinettes,
- ◆ la faculté d'isoler des mots dans la chaîne parlée,
- ◆ l'interprétation des rapports texte/image,
- ◆ la qualité de l'écriture inventée (cf. Émilie Ferreiro).

Il s'agit bien évidemment ici de la volonté d'explorer des relations entre compétences initiales et performance terminale, afin de comprendre la contribution des unes¹ à l'autre. On dit alors qu'il y a une variable à expliquer et 9 variables candidates pour apporter une explication ; une variable à prédire et 9 variables prédictives ; ou encore, une variable dépendante (la performance finale de lecture) et 9 variables indépendantes (des capacités techniques en début d'année).

	support	projet	langtec	lettres	conspho	devin_	motisolé	textimag	écritinv	perfect
1	4,5	4	25	2	1	5	5	2	4	20
2	7	4	23	9	15	5	7	3	5	21
3	9	5	28	9	15	6	8	3	5	22
4	8	4	25	8	1	6	8	3,5	4	22
5	8	5	26	10	5	6	8	3,5	4	22
6	8	5	22	10	5	6	9	3	3	22
7	5	2	21	5	4	5	6	2	3	18
8	1	2	23	2	5	1	7	2	3	17
9	8	2	22	9	0	5	3	2	3	17
10	3	0	10	0,5	0	4	1	1	2	2
11	8	1	18	2	10	3	2	1	2	2
12	6	1	12	2	0	4	2	1	2	4
13	5	1	8	2	0	4	2	1	2	6
14	3	0	18	1	4	4	3	1	2	9
15	3	1	2	2	1	2	0	1	2	2

Tableau 2 : les 9 premières colonnes donnent les résultats à des épreuves techniques passées en début de CP, la dixième colonne la performance de lecture en fin de CP.

¹ Dans le tableau initial, toutes les données étaient numériques, sauf pour 2 variables où elles étaient exprimées par des lettres (A,B, C, C+, D, E) mais manifestement de manière ordonnée et non comme dans une typologie, si bien que, pour la facilité d'exposition, nous les avons transformées en valeurs numériques après nous être assurés que cela ne changeait rien aux résultats de l'analyse.

2.1.1. La régression simple

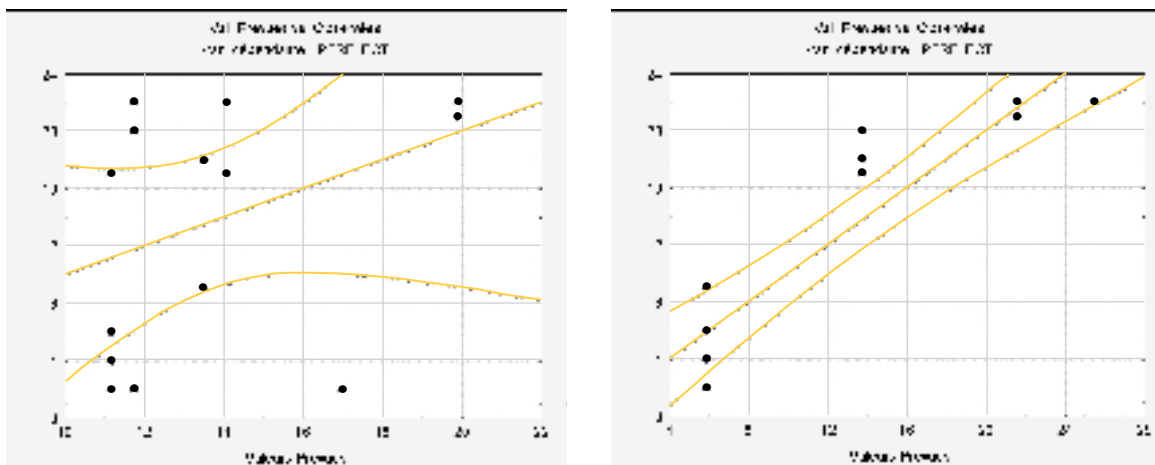
L'exploitation statistique la plus ordinaire, après s'être assuré de la normalité des distributions, consiste à rechercher quel lien (quelle corrélation) existe entre une capacité initiale et la performance finale. Cette relation est exprimée par un coefficient calculé (r) qui varie de -1 à +1, la valeur absolue indiquant l'intensité de la liaison, le signe indiquant le sens.

	support	projet	langtec	lettres	conspho	devi_	motisolé	textimag	écritinv
PERFLECT	,44	,88	,87	,79	,36	,64	,93	,91	,86

Tableau 3 : corrélations entre chaque capacité technique initiale prise séparément et performance finale (sur fond jaune, celles qui sont significatives)

Le tableau 3 donne la valeur du r calculé entre chaque compétence mesurée en début d'année et la performance de lecture de fin de CP. Ici, toutes les capacités initiales sont corrélées positivement avec la performance finale. On peut dire, par exemple, que meilleur on est en début de CP dans la connaissance des supports de l'écrit ou dans la maîtrise du langage technique de la lecture et meilleur on est à la fin du CP en lecture. Rien pour autant ne permet de dire que ceci est cause de cela. En outre, certaines corrélations sont significatives (sur fond jaune) et d'autres non, en fonction de leur valeur absolue et de la taille de la population sur laquelle ce coefficient est calculé. Dans le cadre du fichier de nombres utilisé ici, la connaissance des supports et la conscience phonologique, bien que corrélées positivement avec la performance finale de lecture, ne le sont pas suffisamment : il y aurait plus de 5 possibilités sur 100 de se tromper en affirmant que la relation est significative. Mais pour les 7 autres capacités techniques, la probabilité d'erreur est inférieure à 5% ($P < ,05$).

On se représente sur les 2 plans suivants en quoi une corrélation permet de prédire, à partir d'une donnée initiale, une performance finale et comment, plus la corrélation est forte, plus la prédiction est précise. Le plan de gauche compare les performances effectives de lecture en fin d'année avec celles qui peuvent être prédites à partir du niveau de conscience phonologique ($r = .36$) ; et le plan de droite lorsque le prédicteur est le niveau d'interprétation du rapport texte/image ($r = .91$) :

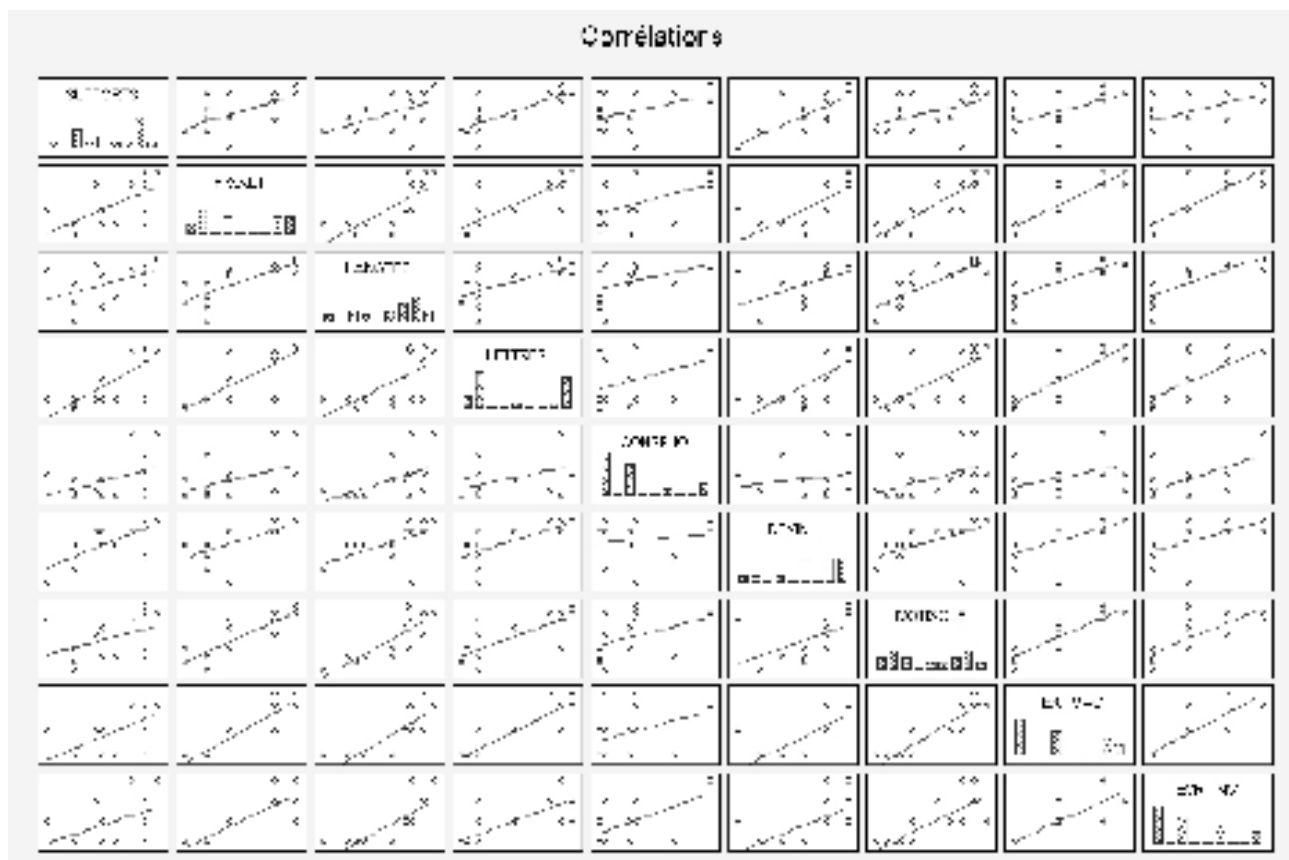


Pourtant, on se gardera bien, au vu des résultats d'une régression simple, même devant un graphique comme celui de droite, de prescrire un investissement pédagogique particulier². Les statisticiens n'arrêtent pas d'appeler à la prudence. Par exemple, disent-ils, il existe une corrélation certaine entre le montant des pertes dans un incendie et le nombre de pompiers présents pour éteindre le feu ; or, cette corrélation ne nous encourage pas à prétendre que, si le nombre de pompiers était plus faible, les dégâts seraient moindres. Une troisième variable (l'ampleur initiale du feu) vient certainement «influencer» à la fois le montant des dégâts et le nombre de pompiers. Si on

² et *a fortiori*, sur des données non représentatives !

«contrôlait» cette variable (par exemple en ne prenant en compte que les incendies de même taille), la corrélation disparaîtrait, ou même changerait de sens. Le principal problème d'interprétation des corrélations simples est qu'en général, nous ignorons ce qu'est l'agent «caché» (ou même s'il y en a un) qui fait que 2 variables sont liées entre elles sans pour autant que l'une explique l'autre. Dans n'importe quel phénomène un peu complexe, et pour la lecture autant que pour l'incendie, il est probable qu'interviennent une multitude de facteurs eux-mêmes en relation.

Par exemple, si nous en revenons aux données du tableau 2, les relations des 9 capacités initiales (les 9 premières colonnes) entre elles sont importantes, comme en témoigne ce regroupement des droites de régression simple, chaque rectangle visualisant le lien entre ligne et colonne :



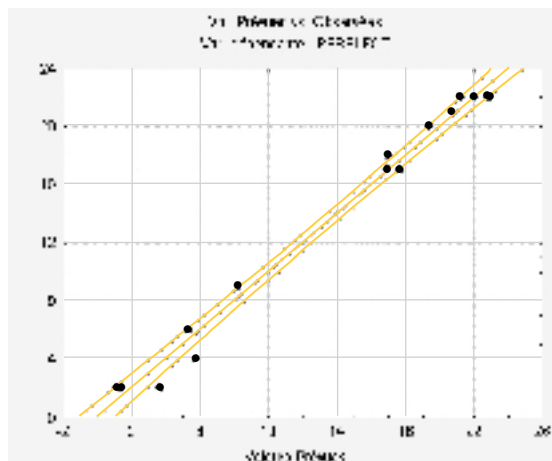
Y aurait-il parmi ces données un «agent caché» qui explique pourquoi les capacités techniques au début du CP sont, globalement, liées entre elles ? Comment s'assurer que telle capacité initiale n'est liée à la performance finale que parce qu'elle est sous la dépendance d'une autre variable et que, si on parvenait à «bloquer» cette autre variable (comme on le fait de la taille de l'incendie), la relation cesserait ou même s'inverserait ? Cette interrogation conduit le chercheur à utiliser un outil statistique plus puissant afin d'accroître le nombre d'éléments explicatifs pris en compte **simultanément**. Toutefois, si l'agent «caché»³ est à l'extérieur du modèle, parce qu'on n'a pas su ou pensé à le mesurer, il restera caché...

³ Cette métaphore de l'agent caché est dangereuse si elle fait penser à un facteur général. En fait, ce qui est d'abord caché, c'est que les variables distinctes dont on se sert pour en expliquer une autre sont déjà redondantes entre elles (elles varient ensemble). Or, pour connaître le rôle d'une variable, il faut l'observer comme si elle était seule à varier, ce qui est impossible dans la réalité et n'est possible qu'à travers un modèle mathématique.

2.1.2. La régression multiple

La régression multiple est cet outil qui permet de connaître la contribution de chaque variable prédictive à la variable à expliquer en neutralisant l'effet simultané de toutes les autres variables prédictives. On obtient ainsi un modèle « explicatif », certes limité aux seules variables introduites⁴ et pour la population sur laquelle il est construit⁵, mais qui décrit l'effet de chacune d'elle, toute chose restant égale par ailleurs.

Dans le cas de ces données fictives du tableau 2, la capacité prédictive du modèle est très forte, comme en témoignent le plan (à droite) et le tableau 4 où sont réunis les résultats réels (la 10ème colonne du tableau 2) des élèves en fin de CP et les résultats prédictibles à partir des 9 capacités initiales considérées simultanément :



observées	20,000	21,000	22,000	22,000	22,000	22,000	18,000	17,000	17,000	2,0000	2,0000	4,0000	6,0000	9,0000	2,0000
prédits	19,373	20,690	22,893	21,201	22,773	21,970	16,968	17,668	16,946	3,6862	1,3639	5,7764	5,3086	8,2381	1,1419

Tableau 4 : on voit combien le prédit est proche de l'observé, en moyenne une erreur de 0.77 point soit moins de 6 % de la performance observée

Il serait même possible ici de prévoir les performances en lecture à la fin de CP d'un 16ème élève dont on aurait connu les capacités initiales à condition, évidemment, qu'il ait vécu, pendant l'année, la même chose que les 15 premiers. Ce modèle puissamment prédictif ne prédit pourtant rien en dehors de ce qui a permis de l'établir, c'est-à-dire un lien multiple entre des données initiales et finales, lien qui se médiatise par une intervention pédagogique. Pour être rigoureux, il faudrait dire : « si on fait revivre à des enfants ce qui a été vécu par ces 15 enfants pendant un an, alors on peut prédire leur niveau de lecture à la fin du CP à partir de 9 capacités initiales en les faisant entrer dans cette équation. Mais on ne peut rien savoir pour un autre type d'intervention pédagogique puisque cette source probable de variation n'a pas été contrôlée par le modèle. »

À ces conditions près (qu'il est tellement facile d'oublier), le modèle va « dire » quelle est la contribution propre de chaque capacité initiale. On la lira aisément dans le tableau 5 en consultant la colonne des *paramètres du modèle*. La performance finale en lecture d'un élève est donnée par une équation qui ajoute à une constante le produit de chacune de ses capacités initiales (cf. tableau 2) par le paramètre qui lui correspond, par exemple :

pour le premier élève : $\text{Perflec} = -1,98698 + (4,5 \cdot 0,84723) + (4 \cdot 0,53542) + (25 \cdot 0,32873) + \text{etc.} = 19,373$ pour une performance effective de 20, ce qui crée, on le voit, une faible distorsion.

pour le dernier élève : $\text{Perflec} = -1,98698 + (3 \cdot 0,84723) + (1 \cdot 0,53542) + (2 \cdot 0,32873) + \text{etc.} = 1,142$ pour une performance effective de 2.

⁴ Je n'aborde pas ici le critère du R² qui indique le pourcentage de la variance totale expliqué par le modèle

⁵ Savoir quelles sont les limites de confiance du modèle est un autre aspect que je n'aborde pas ici et qui dépend de la manière dont la population (sur laquelle le modèle est construit) est un échantillon représentatif d'une population plus vaste à laquelle on souhaite étendre les conclusions.

	Paramètres du modèle	Corrélation partielle	P Probabilité
CONSTANTE	- 1,98698		.446332
SUPPORTS	- 0,84723	-.627	.131668
PROJET	0,53542	.286	.533275
LANGTEC	0,32873	.722	.066570
LETTRES	1,25862	.853	.014563
CONSPHO	- 0,45232	-.770	.042862
DEVIN_	0,16723	.130	.779724
MOTISOLÉ	1,55103	.709	.073933
TEXTIMAG	- 5,05267	-.673	.097538
ÉCRITINV	3,56548	.732	.061167

Tableau 5 : les résultats de la régression multiple. Sur fond jaune, les seules capacités initiales dont la contribution est significative à $P < .05$

Il faut distinguer ici entre une information mathématique (*Paramètres du modèle*) et une information statistique (*Corrélation partielle*). Grâce à la colonne « paramètres du modèle », on comprend donc que toute augmentation d'un point à l'épreuve mesurant la connaissance des supports de l'écrit (ce score variait de 1 à 9 dans le tableau 2) provoque à elle seule, les autres performances restant inchangées, une baisse (puisque le signe du paramètre est négatif) de la performance finale de lecture de 0.84 point (score qui varie de 2 à 22 avec une moyenne de 13.73) ; toute augmentation d'un point du score qui mesure la qualité de la représentation du projet d'apprendre à lire provo-

que, à elle seule, les autres performances restant inchangées, une élévation de la performance finale de lecture de 0.53 point (puisque le paramètre est positif) ; etc. On voit ainsi que 3 capacités initiales contribuent de manière négative à la construction du score de lecture finale : connaissance des supports, conscience phonologique, interprétation des rapports texte/image. Les 6 autres y contribuent positivement. Pour autant, cette valeur « mathématique », si elle informe sur le sens de la liaison entre une variable explicative et une variable à expliquer, ne dit pas s'il est légitime de parler de cette liaison, au sens où la contribution de l'une à l'autre est statistiquement significative.

Dans la colonne « *corrélation partielle* » est donnée la corrélation entre chaque variable initiale et la variable à expliquer. Par exemple, entre la compréhension du projet d'apprendre à lire et la performance finale de lecture, la corrélation est $r = .28$, ce qui est beaucoup plus faible que dans la régression simple $r = .88$ (tableau 3). Certaines mêmes, comme pour la connaissance des supports de l'écrit, s'inversent.

Dans la colonne « *P probabilité* » est donné le risque d'erreur, c'est-à-dire le degré de signification de la corrélation. Au seuil généralement admis de moins de 5% de risque d'erreur ($P < 0.05$), seulement 2 capacités initiales (contre 7 lorsqu'il s'agissait de corrélations simples) contribuent de manière significative à l'établissement de la performance finale de lecture :

- **positivement**, la possibilité d'appeler les lettres par leur nom : chaque point gagné dans cette capacité (dont le score varie de 0,5 à 10) produit, toutes choses égales par ailleurs, une **augmentation** de la performance de lecture de 1,25 point.
- **négativement**, le niveau de conscience phonologique : chaque point gagné dans cette capacité (dont le score varie de 0 à 10) produit, toutes choses égales par ailleurs, une **baisse** de la performance de lecture de 0,45 point.

Si ces données n'étaient pas fictives, il resterait à interpréter de tels résultats dont le constat ne se suffit pas à lui-même ; ce qui n'est jamais chose aisée. Dans l'immédiat, ces chiffres se prêtent seulement à la mise en évidence des risques de traitements statistiques insuffisants, risques d'autant plus grands que le lecteur n'a pas les moyens de s'assurer ni que des corrélations significatives dans une régression simple le resterait dans une régression multiple, ni que, si elles le restaient, elles garderaient le même sens de contribution.

3. MAINTENANT SUR DES DONNEES REELLES

Nous avons, à l'INRP, suivi 59 élèves de la Grande Section de maternelle à la fin du CM2. À cette date, ils ont passé des épreuves complexes qui ont permis d'établir une performance de lecture⁶ à la fin du cycle 3. Par ailleurs, nous avons mesuré 5 capacités techniques⁷ en relation supposée avec le traitement de l'écrit, grâce à des épreuves passées à la mi CE1, :

- une **épreuve lexicale** donnant un score d'identification de mots isolés (présentés ou non dans un contexte phrastique oral), score majoré par la vitesse d'identification
- une **épreuve grapho-phonologique** mesurant la qualité du déchiffrage de pseudo-mots afin de décrire la connaissance de la valeur sonore, hors contexte lexical, de quelques-uns des phonogrammes de 2 ou 3 lettres parmi les plus fréquents de la langue française
- une **épreuve graphématique** mesurant la qualité de l'écriture inventée, analysée sous l'angle de la présence des graphèmes, épreuve adaptée de Émilie Ferreiro.
- une **épreuve phono-graphique**, la même que précédemment, mais mesurant la qualité de la transcription phonologique, analysée sous l'angle de la reproduction des phonèmes.
- une **épreuve «graphique»** consistant à distinguer des mots possibles en français (mais inconnus) et des mots impossibles car utilisant des trigrammes n'apparaissant jamais en français à cette position dans un mot (d'après Content et Radeau, 1988).

Globalement donc, un plan de même type que le précédent, à la différence qu'il porte sur une authentique population expérimentale. À la différence également que la pédagogie de la lecture pratiquée au cycle 2 est contrôlée par une variable selon 3 modalités :

- une première où on enseigne directement la voie directe et exclut le recours à la correspondance grapho-phonologique pour entrer dans l'écrit (**directe**),
- une seconde où on fait explicitement de cette correspondance le principe sur lequel doit se construire le processus de lecture (**indirecte**),
- une troisième où on fait de ce principe un recours et/ou une vérification, ni préalable ni séparable du traitement sémantique (**mixte**).

En quoi des capacités techniques mesurées en fin de CE1 sont-elles prédictives des performances futures en lecture ? Pour aller rapidement à l'essentiel, lorsque la variable pédagogie est contrôlée, les résultats de la régression multiple pour «expliquer» à la fin du cycle 3 la performance de lecture par 5 capacités techniques se lisent dans le tableau 6 :

	PARAM	P	CORRÉLATION PARTIELLE	CORRÉLATION SIMPLE
Constante	40,782	.00001		
Péd directe	10,47978	.009	.335	
Péd mixte	- 3,52083	.421	-.112	
Péd indirecte	- 6,95894	.102	-.226	
Lexique	1,086	.017	.287	-.04
Décodage	- 3,780	.038	-.248	-.18
Graphème	0,331	.166	.164	-.06
Phonème	- 0,449	.007	-.327	-.16
Consgrap	1,468	.191	.154	.18

Tableau 6 : résultat de la régression multiple «expliquant» la performance de lecture en fin de CM2 par 5 capacités techniques évaluées au CE1 et par le type de pédagogie pratiquée au cycle 2. La colonne de droite donne la corrélation simple, c'est-à-dire le lien entre la performance finale et les capacités techniques prises isolément ; aucune est significative (le seuil serait ici de .25). L'avant-dernière colonne donne la corrélation de chacune de ces capacités techniques lorsque toutes les autres sont «contrôlées» dans une régression multiple ; 3 d'entre elles (sur fond jaune) contribuent de manière significative ($P < .04$).

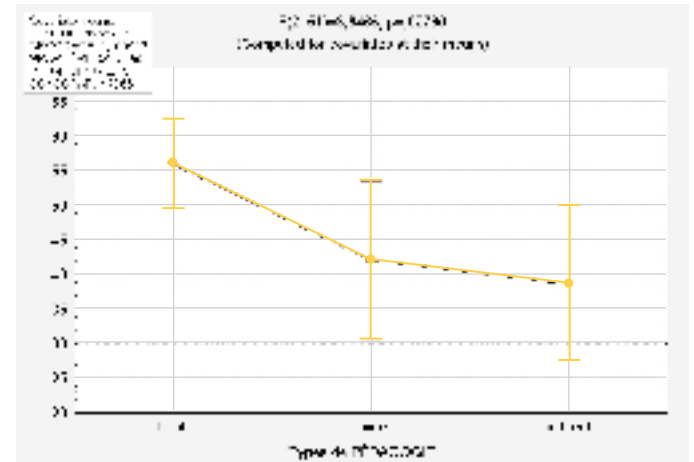
⁶ Ces épreuves ont été conçues et utilisées en 1995 dans le cadre de la recherche INRP : Lecture et voie directe

⁷ Ces épreuves ont été conçues et utilisées entre 1990 et 1993 dans le cadre de la recherche INRP : La lecture dans le cycle des 5-8 ans

On comprend désormais que le score final peut-être calculé en utilisant les informations de la colonne «paramètres du modèle». Chaque point dans la maîtrise du lexique augmente la performance finale en lecture de 1,086 point ; chaque point supplémentaire dans l'habileté à décoder fait baisser le score final de 3,78 points ; chaque point supplémentaire dans la qualité de l'écriture inventée observée sous l'angle de la présence des graphèmes augmente le score final de 0,331 point ; etc. Mais seulement 3 variables (sur fond jaune) contribuent au modèle, toutes les autres variables étant contrôlées, de manière significative (moins de 4 % de risques d'erreur) : l'identification de mots isolés (présentés ou non dans un contexte phrastique oral), la qualité du déchiffrage de pseudo-mots, la qualité de la transcription phonologique analysée sous l'angle de la reproduction des phonèmes. L'épreuve lexique en positif, le décodage et la transcription phonologique en négatif.

Dits brutalement, ces résultats suggèrent que, pour la population sur laquelle ils ont été établis, plus un enfant de CE1 est, toutes choses égales par ailleurs, habile dans le maniement de la correspondance grapho-phonologique et plus ses résultats en lecture seront faibles 3 ans plus tard. Cette proposition est inadmissible par les présupposés des prescriptions pédagogiques actuelles. Elle peut même passer pour une provocation qu'il faut combattre au nom de la science. Et si ces effets inconvenants étaient la conséquence de pratiques pédagogiques douteuses dont aurait souffert cette population et qui l'auraient détournée de bien utiliser le principe alphabétique ? Or, ces 59 élèves «proviennent» de 3 «pratiques» distinctes dont les effets sont eux-mêmes pris en compte dans le modèle : toutes choses égales par ailleurs, les différences entre types de pédagogie sont, avec moins de 3% de risques d'erreur, significatives en faveur de celle qui exclut de prendre appui sur le principe alphabétique. Voilà qui renforce encore les effets observés.

Comment expliquer cette convergence d'effets⁸ que révèle ici la régression multiple⁹ ? Sauf à laisser entendre que les données initiales ont été truquées, force est bien de consacrer quelque attention à ces résultats. Globalement, ils suggèrent que le passage de la phase alphabétique à la phase orthographique, de la voie indirecte enseignée pour apprendre à lire à la voie directe exercée pour lire, ne se fait pas aussi bien qu'on le pense. Non seulement ces phases ne seraient pas en continuité mais il se pourrait qu'elles soient antinomiques. L'habitude inculquée initialement de régler les questions dans l'écrit en le référant à l'oral, non seulement ne préparerait pas à les régler sans l'oral mais ferait obstacle au développement des processus directs



⁸ On peut avoir une autre confirmation de la contribution négative de ce recours initial à la correspondance grapho-phonologique en prenant en compte le moment où se produirait le fameux «déclic» qui fait dire que l'élève est désormais capable de lire n'importe quel mot, y compris ceux qu'il n'aurait jamais vus. Plus cette maîtrise est précoce au cycle 2 et plus les résultats seront faibles en CM2. On pourrait encore espérer que c'est la manière dont a été calculée la performance de lecture en fin de CM2 qui est en cause. Mais nous avons déjà rendu compte de résultats semblables mis en évidence en fin de CE1 à partir d'une évaluation de la performance de lecture construite évidemment sur d'autres épreuves (Rapport INRP : la lecture dans le cycle des 5-8 ans, 1994, 1995). Incontestablement, sur cette population, les faits sont établis solidement. Pour autant, ils ne témoignent que pour eux-mêmes et devraient être retrouvés dans d'autres circonstances pour devenir «vrais».

⁹ Rappelons au passage qu'une régression simple (colonne de droite du tableau 6) donne des corrélations non significatives, d'où la nécessité de ne pas en rester à des analyses statistiques de surface et/ou de déposer ses fichiers de données en même temps que la publication qui en rend compte.

nécessaires puisqu'elle offrirait à l'apprenti une réponse provisoirement suffisante pour satisfaire des attentes et des besoins rudimentaires. C'est en quelque sorte l'avantage des systèmes alphabétiques de rendre possible l'alphabétisation, c'est-à-dire un mode d'accès non spécialisé à l'écrit, bien différent du mode d'accès expert. Aujourd'hui que les pays qui, depuis deux siècles, se sont simultanément industrialisés et alphabétisés envisagent de généraliser l'usage expert de l'écrit, on est peut-être en train de découvrir qu'un alphabétisé, même vélocé, n'est pas un lecteur, même médiocre, et que la différence, outre la fondamentale question politique du rapport à l'écrit¹⁰, s'observe dans la nature technique des opérations selon que l'écrit est abordé indirectement comme un système de notation de l'oral ou directement comme un système linguistique.

4. DE MANIERE PLUS GENERALE ?

L'état des recherches et le niveau des résultats des élèves devraient inciter à considérer un instant une question qui ne se pose pas que pour la lecture : **quel est le rapport entre ce qu'on fait pour apprendre et ce qu'il faut apprendre à faire ?** Si, sur d'autres données, il se répétait que la compétence alphabétique contribue de manière négative à l'explication de la performance experte de lecture, sans doute devrait-on s'interroger sur la pertinence d'un détour pédagogique initial qui détourne à terme l'apprentissage de son but.

Il serait aisé de lever le doute sur la pertinence d'une telle interrogation (sans parler encore de s'assurer ensuite de la validité d'une réponse) en reprenant autrement les données de nombreuses recherches achevées, à commencer par les 2 que j'ai citées au début. Pourquoi ne pas revisiter, comme cela se pratique dans d'autres disciplines¹¹, des recherches dont les conclusions semblables sont sempiternellement citées dans les articles spécialisés comme si le lecteur les ignorait et dont les modalités de construction de la preuve sont passées sous silence comme si le lecteur les connaissait. Trois préoccupations pourraient « guider nos pas ».

1. Se donner les moyens de distinguer entre recherche sur la lecture et recherche sur la pédagogie de la lecture et refuser de passer, sauf en terme d'hypothèse, de l'une à l'autre¹². Le non respect de cette distinction conduit à la prolifération d'articles construits, d'une part, sur une compilation sélective de recherches psychologiques faisant délibérément l'impasse sur tout ce qui est problématique, de l'autre sur des prescriptions pédagogiques. Entre les deux, rien, sauf la volonté d'induire que ces prescriptions ont été élaborées scientifiquement. Le plus souvent, on chercherait en vain sur quelles références à des recherches proprement **pédagogiques** se développe l'engagement des auteurs. Le rôle de ces publications est de dénier à la recherche pédagogique un statut de discipline scientifique afin de tirer quelque profit de son illégitimité¹³. Les professionnels de l'enseignement n'ont pas à être dupes de ces manœuvres. Le témoignage de leur vigilance quant à l'exigence d'une méthodologie scientifique dans un domaine de recherche qui les concerne directement serait un facteur déterminant d'assainissement des mœurs.

¹⁰ On peut, justement du fait de l'enjeu politique d'une lecturisation réussie, douter de cette intention de généraliser l'usage expert de l'écrit. D'où la surenchère actuelle pour une sur-alphabétisation qui fonctionne sur l'espoir qu'on inverserait les effets de ce qu'on fait aujourd'hui simplement en le faisant davantage.

¹¹ Cf. notamment Damasio *L'erreur de Descartes, la raison des émotions* (Odile Jacob)

¹² Dans les résultats que j'ai rapidement présentés ici, il est important de ne pas confondre 2 informations de nature différente. Mettre en évidence que le niveau de compétence dans la correspondance grapho-phonologique est inversement corrélé avec la performance de lecture n'autorise pas à conclure qu'il ne faut pas enseigner la correspondance grapho-phonologique. Mettre en évidence que les élèves issus de classes où ne s'enseigne pas la correspondance grapho-phonologique ont de meilleurs résultats en lecture peut y autoriser davantage. La première information est de nature psychologique et ne préjuge pas de la manière dont s'est constituée la compétence de lecture. La seconde est issue d'un contrôle d'une variable pédagogique et s'y prête plus facilement. Mais gare alors à « l'agent oublié » !

¹³ Le champ de la dyslexie semble à nouveau un excellent exemple d'une conjonction provisoire et toujours renaissante d'intérêts de psychologues et de médecins, contre la pédagogie.

2. Cette vigilance passe par un effort d'information et de formation des enseignants. Il urge qu'ils ne délèguent à personne le soin de «savoir lire» un article afin d'être en mesure de juger de sa validité. Gageons que si les dispositifs scientifiques de construction de la preuve étaient convaincants, les chercheurs, qui partagent équitablement le bon sens avec le reste du monde, parviendraient aisément à se mettre d'accord. Or, ce qui se dit aujourd'hui avec aplomb sur la lecture est tellement contradictoire, ne serait-ce qu'avec les résultats que tout le monde s'accorde à déplorer, que les enseignants ne peuvent faire l'économie d'y aller voir de plus près. On est en droit d'attendre d'un médecin qu'il sache «lire» la recherche médicale. Et d'un enseignant la recherche pédagogique. L'exemple que j'ai développé sur des données fictives voulait montrer le risque d'analyses statistiques insuffisantes ; celui comparant classes phoniques et classes globales voulait montrer qu'il y a souvent plus à voir dans des données que la raison qui les a fait réunir. Sans devenir pour autant des professionnels de la statistique, les enseignants sont heureusement capables de juger d'une méthodologie de recherche. C'est aussi ce qui justifie qu'ils se recrutent aujourd'hui après des études universitaires où enseignement et recherche sont liés.

3. Réciproquement, le chercheur a la responsabilité de montrer (et/ou d'offrir à la communauté la possibilité de s'assurer) qu'il n'y a pas d'autres effets dans les données qu'il traite que ceux qu'il a mis en évidence. Pendant longtemps, les données (dont le recueil est toujours une opération coûteuse et délicate) ont permis au chercheur de valider une hypothèse. Depuis la généralisation des méthodes d'analyse des données, celles-ci sont une source d'informations dont l'exploration méthodique est susceptible de faire surgir des hypothèses nouvelles. Au chercheur de ne les abandonner qu'après s'être assuré qu'elles n'ont plus rien à lui apprendre. Mais ces données font alors partie du patrimoine de la communauté scientifique, au même titre que le document que l'historien a découvert, et avec le même statut. Aucune communication ne devrait se faire sans que le matériau sur lequel elle a été élaborée ne soit accessible par d'autres aux fins d'investigation alternative et d'analyse contradictoire. Ce contrôle réciproque, souhaité et nécessaire¹⁴, implique que les chercheurs considèrent que leurs fichiers de données entrent, après qu'ils y aient travaillé, dans le domaine public, et d'autant plus légitimement qu'ils ont le plus souvent été constitués sur des fonds publics.

Jean FOUCAMBERT

¹⁴ Les données qui ont permis la régression multiple présentée dans le tableau 6 sont disponibles sur le site AFL (<http://www.lecture.org>)