

UNIVERSITÉ DE GENÈVE
FACULTÉ DE PSYCHOLOGIE
ET DES SCIENCES DE L'ÉDUCATION

Un bilan des recherches
processus-produit
L'enseignement peut-il contribuer à
l'apprentissage des élèves et, si oui, comment ?

Marcel Crahay

Carnets des sciences de l'éducation
2^e édition
2014

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT
L'ENSEIGNEMENT PEUT-IL CONTRIBUER À
L'APPRENTISSAGE DES ÉLÈVES
ET, SI OUI, COMMENT ?

Table des matières

Préambule	3
I. L'analyse de l'enseignement :	
un paradigme de recherche	7
Enseignement et apprentissage : des objets de recherche différents mais complémentaires	7
De l'évaluation des enseignants à l'étude de l'enseignement	15
L'incontournable question de l'efficacité	18
II. Le projet <i>Follow Through</i>	
et les recherches de Stallings	31
<i>The Follow Through Evaluation Study</i>	32
<i>The California ECE Study</i>	40
La troisième étude de Stallings et le <i>Descriptive-Correlational-Experimental Loop</i>	42
III. Polémique autour du <i>leadership</i>	
de l'enseignant	47
Les recherches de Flanders : influence directe <i>versus</i> influence indirecte	47

Soar et Soar : des clarifications conceptuelles aux vérifications empiriques	59
IV. À la recherche de l'enseignant efficace	69
De la stabilité de l'effet comme préalable à la recherche de l'enseignant efficace	69
Les recherches de Brophy et Evertson	73
Le programme de recherche de Good et Grouws	87
V. Deux pistes d'approfondissement de la méthodologie processus-produit	97
La <i>Beginning Teacher Evaluation Study (BTES)</i> ou l'articulation des approches quantitatives et qualitatives	98
Retour sur les études quasi-expérimentales du groupe de Canterbury	115
VI. Penser le dépassement du paradigme processus-produit	129
L'enseignant efficace existe-t-il ?	129
L'efficacité en matière d'enseignement	132
Comment l'enseignement contribue-t-il à l'apprentissage des élèves ?	138
VII. Bibliographie	147
Liste des revues portant sur les recherches processus-produit	147
Références bibliographiques	150

PRÉAMBULE

Aux alentours des années 60, s'est développé, dans le monde anglo-saxon, un courant de recherche fort qui s'est donné pour objectif l'analyse de l'enseignement (*the study of teaching*) et de ses effets sur l'apprentissage des élèves. Un paradigme de recherche était né. Il a connu son apogée dans les années 80 et un déclin certain dans les années 90. Pour l'essentiel, le monde francophone est resté étranger à ce courant : c'est seulement à la fin des années 80 et dans les années 90 que sont apparus des travaux francophones centrés sur l'efficacité des enseignants ou, dit autrement, sur l'*effet-maître*. Pourquoi cette extériorité ? Question difficile qu'il serait intéressant que des sociologues et historiens des sciences prennent en charge, et ce d'autant plus que ce courant de recherche – fréquemment critiqué de ce côté-ci de l'Atlantique – paraît encore largement méconnu. Ce sera l'ambition de cet opuscule que de faire connaître ce paradigme de recherche.

De façon précise, l'objectif principal de ce petit ouvrage est d'expliquer en quoi consistait ce paradigme d'analyse de l'enseignement et, donc, quels étaient les présupposés épistémologiques sous-jacents aux études *processus-produit*. Les plus importantes de celles-ci seront ensuite passées en revue. La présentation a été organisée de façon à faire émerger les controverses méthodologiques principales qui ont jalonné l'évolution de ce courant de recherche. Enfin, nous clôturerons par une discussion critique afin de souligner les points forts de ces travaux et d'indiquer les problèmes conceptuels et/ou méthodologiques qu'il conviendrait de tenter de résoudre. Car, pour nous, ce courant de recherche ne peut rester lettre morte. À cet égard, nous rejoignons Gauthier (2005) qui a, en quelque sorte, ressuscité les travaux de Stallings et Kaskowitz (1974). Cependant, nous ne suivons pas Gauthier lorsqu'il tire argument de ces recherches en faveur d'un enseignement bien différent de ce que préconisent actuellement ceux qui se réclament du socioconstructivisme. Notre positionnement face aux recherches processus-produit est différent. Pas question de vénérer ce courant de recherche, ni de le bannir ou de le honnir. Pas question, non plus, d'y chercher référence pour l'une ou l'autre croisade pédagogique. Plus justement, il nous paraît essentiel que les chercheurs en sciences de l'éducation tirent les leçons épistémologiques, conceptuelles et méthodologiques d'un courant qui a été dominant au siècle précédent et ce, trois décennies durant.

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

Comme indiqué ci-dessus, cet opuscule concerne exclusivement les recherches processus-produit menées dans le monde anglo-saxon. Pour les études conduites dans le monde francophone, le lecteur est invité à se référer aux excellentes synthèses de Bressoux (1994, 2003). Par ailleurs, le lecteur ne s'étonnera pas de trouver des termes ou expressions en langue anglaise, ceci afin de faciliter le travail de l'étudiant qui souhaite retourner aux textes originaux.

CHAPITRE I

L'ANALYSE DE L'ENSEIGNEMENT : UN PARADIGME DE RECHERCHE

Enseignement et apprentissage : des objets de recherche différents mais complémentaires

Des écoles existent. Des enseignants sont recrutés et rétribués ; ils ont pour mission explicite d'accueillir des élèves selon un horaire journalier, de les occuper en les mobilisant autour d'activités reconnues comme éducatives et de promouvoir chez eux certains apprentissages définis dans le cadre de programmes ou de plans d'études. Pour des auteurs comme Dunkin et Biddle (1974), Gage (1963), Rosenshine (1971), Rosenshine et Furst (1973), Turner (1971), ces affirmations reposent sur des faits institutionnels qu'il ne paraît pas possible de contester. Ces mêmes affirmations permettent de définir les contours d'une activité humaine importante : l'enseignement.

Pour les pères fondateurs de ce paradigme de recherche (Smith, 1960 ; Gage, 1963 ; Turner, 1971 ;

Dunkin & Biddle, 1974), l'activité d'enseignement est considérée comme conceptuellement distincte du processus d'apprentissage. D'une part, il convient de reconnaître que l'enseignement est une activité humaine fondamentale : l'espèce humaine se caractérise par son souci de transmettre ses conquêtes culturelles (p. ex. techniques, cognitives) à ses descendants. Dans toute société, quel que soit son degré de développement, on observe des pratiques par lesquelles les adultes s'attachent à infléchir les conduites des plus jeunes, de sorte que ceux-ci respectent certaines normes sociales, maîtrisent des compétences considérées comme utiles (voire nécessaires) au sein d'un groupe social donné. Il y a donc là un fait de nature psychosociale qui mérite d'être investigué. Les pratiques mises en place dans cette perspective sont plus ou moins codifiées et, dans tous les cas, reflètent les croyances, les convictions et les valeurs de la société au sein de laquelle elles se développent. À cet égard, il faut remarquer que le concept d'enseignement est passible d'une définition large et spécifique. Ainsi, Gage (1963) définit l'enseignement comme « toute forme d'influence interpersonnelle qui a pour but de changer les façons qu'ont d'autres personnes de se comporter » (p. 96). De façon plus générale encore, on pourrait dire que l'enseignement constitue un processus interactif au cours duquel un individu, porteur de normes ou de valeurs et/ou maîtrisant des compétences reconnues socialement, essaie d'infléchir la façon dont une ou plusieurs autres personnes pensent, agissent ou

ressentent les choses. Dans cette acception large du terme « enseignement », la mère enseigne lorsqu'elle s'efforce d'influencer la conduite de son enfant. Il en va de même de l'enfant qui essaie d'influencer son camarade, comme du contremaître qui s'efforce d'améliorer les performances d'un ouvrier. En ce sens, l'enseignement ne diffère pas de tous les processus d'influence sociale et relève, par voie de conséquence, du domaine de la psychologie sociale. On peut quelque peu circonscrire le champ d'investigation en précisant que l'effet visé correspond à la maîtrise de compétences, ou à l'adhésion à des normes ou valeurs, considérées comme socialement souhaitables ; le ou les sujet(s) ciblé(s) par l'enseignement se caractérise(nt), au point de départ, par le fait qu'il(s) ne maîtrise(nt) pas ces compétences et/ou qu'il(s) n'adhère(nt) pas *ipso facto* à ces normes ou valeurs.

Par ailleurs, il convient de tenir compte de nos sociétés. En leur sein, des formes institutionnalisées de transmission des connaissances, compétences, normes et valeurs se sont implantées. Avec les créations des écoles de cathédrale au Moyen Age, des collèges de Jésuites aux XVI^e-XVII^e siècles, de l'école obligatoire pour tous au XIX^e ou au XX^e siècle, l'enseignement est devenu une forme d'influence sociale pratiquée par une catégorie professionnelle déterminée – les instituteurs, maîtres, enseignants et/ou professeurs – dans des lieux institutionnellement définis – les écoles, collèges, lycées ou universités. À ce titre, on doit donc considérer que

l'enseignement correspond à une réalité socio-historique, marquée par l'évolution culturelle et organisationnelle de la société au sein de laquelle elle est étudiée. C'est dans ce sens que le terme *enseignement* est entendu dans le cadre du paradigme de recherche qui prend forme aux Etats-Unis dans les années 60. Plus précisément, il nous semble que le concept d'enseignement tel qu'il est compris dans ce paradigme renvoie aux caractéristiques suivantes :

- L'enseignement est une activité, un processus ou encore une pratique. Ceci signifie qu'il est – au moins partiellement – observable.
- L'enseignement est, par essence, de nature interpersonnelle, interactive, ou encore, sociale. Ceci a pour conséquence que l'enseignement met nécessairement en présence deux ou plusieurs individus. Cette interaction peut être directe, comme dans les classes, ou médiatisée, lorsqu'il s'agit d'enseignement assisté par ordinateur, d'éducation à distance ou encore, de formations véhiculées par Internet. Toutefois, le paradigme de l'analyse de l'enseignement restreint son champ d'investigation aux formes directes d'enseignement¹. Par voie de conséquence, le carac-

¹ Sans doute, pour des raisons liées aux conditions de l'enseignement au moment de son émergence : dans les années 60, les moyens technologiques n'étaient pas ce qu'ils sont aujourd'hui.

tère interactif de l'enseignement prend un sens précis. Il s'agit d'un processus bidirectionnel en ce sens que celui qui enseigne s'efforce d'influencer celui (ou ceux) qui est (ou sont) enseigné(s). En retour, ces derniers influencent d'une manière ou d'une autre le premier. Les tentatives d'influence des uns sur les autres, et réciproquement, se réalisent essentiellement par le canal verbal ou gestuel (p. ex. mimiques, regards, sourires, gestes, déplacements) ; ceci n'exclut pas le recours au code écrit ou à d'autres moyens technologiques (p. ex. vidéo, ordinateur).

- L'enseignement est intentionnel². Ceci signifie d'abord que l'enseignement tend vers un but ou, autrement dit, veut produire des effets spécifiques. Comme souligné ci-dessus, l'enseignement a pour fonction de susciter des apprentissages, ce terme étant compris dans un sens large (c'est-à-dire que les

² Ceci signifie ensuite que l'enseignement comporte un aspect réfléchi ; celui-ci peut correspondre à une phase de planification et/ou à une phase de réflexion, voire de régulation postérieure à l'activité d'enseignement proprement dite. Cette dimension essentielle apparaît dans les intérêts des chercheurs anglo-saxons aux alentours des années 80. Ici, elle ne sera pas au cœur de l'analyse.

effets visés peuvent être d'ordre cognitif ou affectif, attendus à court comme à long terme)³.

- L'enseignement se pratique dans un contexte institutionnel. Ceci signifie qu'il s'observe dans des lieux réservés à son exercice : classe, salle de cours, auditoire. Ceci signifie également qu'il correspond à l'activité d'une catégorie professionnelle reconnue socialement. Ainsi son exercice s'inscrit dans le cadre de contraintes institutionnelles auxquelles correspondent des règlements, des législations, des programmes ou encore des plans d'études. Cette dernière caractéristique implique, enfin, que toute forme d'enseignement s'inscrit dans une histoire sociale, qui l'influence au moins partiellement.

Eu égard à cette définition de l'enseignement, il apparaît évident que sa conceptualisation ne se ramène pas à celle de l'apprentissage. Gage (1963) fait remarquer que, par essence, une théorie de l'apprentissage n'explicite pas les processus par lesquels les enseignants s'efforcent d'offrir aux élèves les conditions optimales pour leurs apprentissages. Par ailleurs, il va de soi que les maîtres véhiculent une théorie d'apprentissage lorsqu'ils enseignent.

³ Les trois premières caractéristiques ont été discutées par différents auteurs : Gage (1963), Amidon et Hunter (1967), Jackson (1986) et enfin Anderson et Burnes (1989).

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

En matière de formation des enseignants, ceci a pour conséquence qu'il ne suffit pas de former les maîtres aux théories d'apprentissage les plus avancées : ceux-ci ont besoin de connaissances spécifiques sur la nature de l'enseignement, et sur la façon de le pratiquer. Bien plus, dans une perspective pragmatique, les pratiques d'enseignement les plus performantes ne s'infèrent pas directement d'une théorie d'apprentissage, aussi validée soit-elle. Une théorie spécifique de l'enseignement efficace est à construire ; elle ne peut l'être que par l'intermédiaire de recherches portant spécifiquement sur cet objet.

Par ailleurs, toujours selon Gage (1963), la transposition de l'une ou l'autre théorie d'apprentissage à l'enseignement est confrontée à deux obstacles majeurs.

D'une part, les expériences menées dans le cadre de la psychologie expérimentale ont pour ambition légitime de comprendre les processus généraux mobilisés lorsqu'un individu apprend. Le regard du scientifique se concentre, dès lors, sur le fonctionnement de la mémoire, sur les facteurs de *feedbacks* et/ou de renforcement, etc. Ces analyses correspondent indéniablement à une dimension importante de l'enseignement, mais n'épuisent pas la totalité du phénomène. Les enseignants doivent nécessairement prendre en compte les aspects propres aux compétences disciplinaires qu'il leur faut faire maîtriser par les élèves et ceci en vertu des contraintes institutionnelles évoquées ci-dessus.

D'autre part, puisque l'enseignement s'exerce dans des lieux spécifiquement dévolus à cet effet, et qu'il comporte le plus souvent une dimension groupale, sa pratique implique nécessairement la gestion de relations interindividuelles et exige, par voie de conséquence, l'intégration de savoirs relevant à la fois de la psychologie cognitive et de la psychologie sociale. Ce dernier argument a été particulièrement développé par Nuthall (1968). Pour ce dernier auteur, dans le meilleur des cas, les théories d'apprentissage pourraient être considérées comme directement applicables à la situation de préceptorat. Or, celle-ci est exceptionnelle dans le contexte contemporain des écoles et autres institutions d'enseignement. À ses yeux, l'interrogation-clé à laquelle sont confrontés quotidiennement les enseignants est la suivante : quelle est la meilleure façon de gérer un groupe d'élèves et/ou d'étudiants de telle sorte que la plupart d'entre eux (idéalement tous) réalisent un apprentissage de type disciplinaire tel que défini dans un programme particulier ? À cet égard, la psychologie de l'apprentissage se révèle trop générale pour être suffisamment instructive⁴.

⁴ Précisons que pour Nuthall comme pour Gage, tout étant insuffisants, les apports de la psychologie de l'apprentissage ou de la psychologie cognitive sont nécessaires.

De l'évaluation des enseignants à l'étude de l'enseignement

À l'origine, la recherche sur l'enseignement trouve sa raison d'être dans la pratique. Plus précisément, elle naît de l'obligation des inspecteurs, directeurs d'école et autres responsables pédagogiques d'évaluer les enseignants (de Landsheere, 1978). Rosenshine (1986) rappelle qu'en une première période, les recherches ont porté sur la personnalité et les caractéristiques des enseignants. Les travaux de Barr et Dreeben (1977) portent sur ce que l'on a nommé les *variables présomptives* : les attitudes, intérêts, personnalités des enseignants, mais aussi sur d'autres paramètres comme leur degré de maîtrise de la matière à enseigner ou la durée de leur formation initiale. Leur objectif est de déterminer lesquelles de ces variables vont significativement de pair avec une augmentation du rendement des élèves ou, le plus souvent, avec une meilleure appréciation des enseignants par leur directeur ou inspecteur. À propos de ces recherches, Gage (1963) dresse le bilan suivant :

Ces études ont produit des résultats décevants : les corrélations soit ne sont pas significatives, soit ne se retrouvent pas d'une étude à l'autre. D'une manière générale, elles manquent de signification psychologique ou pédagogique. (p. 118).

De leur côté, Getzels et Jackson (1963) écrivent que, « en dépit d'un demi-siècle de considérables efforts de

recherche, on ne sait rien de solide sur les liens entre la personnalité des enseignants et leur efficacité pédagogique » (p. 574).

Une seconde période débute dans les années cinquante, avec notamment les travaux de Flanders, Medley et Mitzel. Ces chercheurs s'attachent principalement à comptabiliser, de manière systématique, des comportements spécifiques du maître et, parfois, des élèves. L'objectif ultime est de mener des études processus-produit dont le principe méthodologique consiste à corréler la fréquence de certains comportements d'enseignement avec des mesures de rendement. Mais le but premier est purement pragmatique. Comme le souligne Bayer (1986), il s'agit avant tout de rationaliser les pratiques des inspecteurs ou superviseurs scolaires, « dont le défaut majeur était – et reste d'ailleurs – l'inconstance flagrante des jugements portés sur les enseignants » (p. 485). Bien que ce mouvement de recherches d'*application* pure n'ait pas abouti⁵, il a fortement contribué à enraciner dans le champ de l'analyse de l'enseignement deux idées de base qu'il ne fût pas facile de dépasser. Bayer (1986) les a dénoncées avec lucidité :

Tout d'abord, ces recherches postulèrent, ne serait-ce qu'implicitement, comme une égalité de nature entre comportements et processus d'enseignement, de sorte que

⁵ Notamment, « parce que la méthodologie d'observation recommandée, trop lourde, ne pouvait pas être réalistement adoptée par les inspecteurs » (Bayer, 1986, p. 486).

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

l'étude de l'enseignement se réduisit alors à la description aussi exhaustive que possible des interactions maître-élève en classe. Ce fut l'époque de la floraison des grilles d'observation, dominée par l'obsession de posséder une grille suffisamment complexe pour rendre compte à elle seule de toutes les dimensions pertinentes du phénomène étudié (...) ⁶ (p. 486).

D'autre part, largement conditionnée par la nécessité d'évaluer des personnes, l'adoption d'un modèle d'analyse des données dérivé de la psychologie différentielle ancre l'idée que la variance des comportements observés en classe reproduisait fidèlement celle des qualités des enseignants, conduisant à penser que l'enseignement est, pour l'essentiel, déterminé par les caractéristiques personnelles – innées ou acquises – des enseignants (Bayer, 1986, pp. 486-487).

Il a fallu attendre les années 60 et, plus particulièrement, la publication du premier *Handbook of Research on Teaching* pour assister à une révolution paradigmatique qui, dépassant l'idée d'évaluer les enseignants,

⁶ En Belgique, le travail doctoral de Bayer (1972), intitulé « Analyse multidimensionnelle de la communication verbale d'enseignement » est illustratif de cette tendance. L'ouvrage de Landsheere, (avec la collaboration de Delchambre), daté de 1979 et intitulé *Comment les maîtres enseignent* participe du même souci de produire une description exhaustive de ce qui se passe en classe et, par voie de conséquence, de compléter l'analyse des comportements verbaux par celle des comportements non verbaux.

pousse les chercheurs à appréhender les réalités d'enseignement avec le souci premier d'en saisir la logique propre et, en principe, de reléguer à l'arrière-plan toute velléité prescriptive ou réformatrice. Ainsi, dans un texte de synthèse, Dunkin (1986) insiste sur l'importance d'analyser les processus d'enseignement tels qu'ils sont et non tels qu'ils devraient être. Cet axiome le conduit naturellement à s'interroger sur la nature de l'objet à étudier. Il pose dès lors la question fondamentale : qu'est-ce qu'enseigner ? La définition en quatre points proposée ci-dessus doit beaucoup aux réflexions de cet auteur australien, devenu célèbre pour un modèle conceptuel qui, distinguant les variables de contexte, de présage, de processus et de produit, a structuré pendant plusieurs décennies les recherches dans le domaine.

L'incontournable question de l'efficacité

En éducation, il est difficile – voire impossible – d'échapper durablement à toute ambition réformatrice. Dans le courant de l'analyse de l'enseignement, elle ne fut, dans les faits, jamais réellement évacuée. Dans le troisième *Handbook of Research on Teaching*, Shulman (1986) soutient que tout l'investissement scientifique réalisé dans le domaine de la recherche sur l'enseignement peut être considéré comme un faisceau de tentatives concertées pour, à la fois comprendre le phénomène de l'enseignement, en améliorer l'efficacité et découvrir les procédures de formation des enseignants les plus rentables. Meddley (1988) défend la même idée

lorsqu'il considère la variabilité de la qualité de l'enseignement comme le fait de base de la recherche sur l'enseignement, celle-ci consistant d'une part à trouver quelles variables différencient les prestations très efficaces de celles qui le sont peu et, d'autre part, à expliquer le pourquoi et le comment de ces différences.

Le souci de contribuer à l'amélioration de l'efficacité de l'enseignement s'est manifesté de diverses façons.

Il est implicitement présent dans ce qu'il est convenu d'appeler les études descriptives. Un exemple connu nous est donné par l'étude réalisée par Adams (1969) sur la direction des échanges engagés par les maîtres. L'auteur y montre que les maîtres s'adressent principalement (approximativement dans 70 % des cas) aux élèves situés dans un espace précis et relativement étroit de la classe situé le long de la rangée centrale. De surcroît, plus un élève se trouve au centre, plus il est probable qu'il sera impliqué activement dans les échanges. Adams et Biddle (1970) soulignent combien l'existence de ce phénomène n'est pas en soi intéressante. Mais il prend de l'importance par toutes les questions qu'il engendre : s'il existe une relation entre la participation verbale des élèves et le rendement, il importe de comprendre en détails les déterminants de la relation entre situation dans l'espace et participation verbale ; mieux même, il importe de contrôler le phénomène. Un second exemple peut être pris dans les recherches sur la gestion du temps. Pour l'essentiel, ces recherches révèlent la grande variabilité, inter-classes, du temps réellement

dévolu aux activités d'enseignement et, à l'inverse, des temps morts et des moments consacrés à la gestion de la discipline (Crahay, 2000). En soi, la variabilité d'un phénomène n'est ni positive, ni négative. En revanche, si l'on considère – ce qui est courant – que le temps dévolu à l'enseignement des diverses disciplines constitue une mesure indirecte des opportunités d'apprentissage offertes aux élèves, la mesure de la variabilité inter-classes devient une estimation de l'inégalité des chances.

Dans ces deux exemples, la variable dont on décrit les fluctuations n'a de pertinence qu'en fonction d'une même prémisse : la variable étudiée constitue un facteur d'efficacité pédagogique. Dans l'un et l'autre cas, le bon sens tend à faire croire à la réalité d'une relation positive. Le devoir du chercheur le pousse, à l'inverse, au doute : il se doit de vérifier, dans l'un et l'autre cas, le bien-fondé du présupposé. Or, comme on le verra plus loin, celui-ci n'est pas vérifié dans le cas de la participation verbale. En revanche, il l'est – avec certaines nuances – dans le cas du temps alloué aux divers objets d'apprentissage (Crahay, 2000).

La question de l'efficacité est explicitement présente dans les recherches qui ont porté sur la comparaison de méthodes : comparaison de méthodes de lecture (globale *versus* analytique) et comparaison « pédagogie active » *versus* « enseignement traditionnel ». Dans les deux cas, il a fallu conclure à l'impasse et ceci pour des raisons identiques : la variable indépendante est à la fois trop massive, trop générale et trop imprécise.

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

L'échec des études comparant les effets des pédagogies actives (*open classrooms*) à ceux générés par un enseignement traditionnel a influencé Gage (1963) dans son effort de conceptualisation de la méthodologie processus-produit. Rappelons d'abord que le bilan de ces études comparatives se solde par un non-lieu. Il existe autant de preuves à charge et à décharge des pédagogies inspirées de *l'Education Nouvelle*. Plus justement, la supériorité des effets de ce type de pédagogie est éminemment labile : elle apparaît dans certaines études alors que, dans d'autres, la tendance des résultats lui est défavorable. Les analyses sophistiquées menées par Hedges, Giaconia et Gage (1981) pour comprendre les raisons de résultats aussi discordants renvoient à une question fondamentale : concrètement parlant, que sont les pédagogies actives ? Dit autrement, cette question – cruciale du point de vue de la méthodologie de recherches – peut être formulée comme suit : que compare-t-on lorsqu'on oppose, d'une part, les pédagogies actives et, d'autre part, l'enseignement traditionnel ? Philosophiquement et conceptuellement les deux approches sont distinctes mais, sur le plan des pratiques, qu'est-ce qui fait véritablement différence ? Car, si l'enseignement dit traditionnel renvoie à un modèle comportemental relativement homogène, qu'en est-il des pédagogies actives ? Bien plus, qu'il s'agisse de l'enseignement traditionnel ou des pédagogies actives, le modèle comportemental renvoie à une multiplicité de facettes relationnelles et de pratiques dont les fluctuations composent une infinité de possibili-

tés. Or, si la variable indépendante est mal définie, la relation que l'on mettrait éventuellement en évidence avec une variable dépendante quelconque serait, en définitive, difficile à interpréter.

Explicitons cette objection fondatrice du paradigme de l'analyse de l'enseignement. Pour ce faire, supposons que pédagogie active et enseignement traditionnel se distinguent et s'opposent sur cinq dimensions : initiatives laissées aux élèves, climat affectif en classe, diversité du matériel didactique, questionnement du maître et types d'évaluation pratiqués par le maître. Supposons, de façon complémentaire, six classes caractérisées comme indiqué dans le tableau ci-après.

Tableau 1. Classes et caractéristiques

	Classe 1	Classe 2	Classe 3	Classe 4	Classe 5	Classe 6
Initiatives laissées aux élèves	<i>Nombreuses</i>	<i>Nombreuses</i>	<i>Moyenne</i>	<i>Moyenne</i>	<i>Faible</i>	<i>Faible</i>
Climat affectif	<i>Chaleureux</i>	<i>Froid</i>	<i>Chaleureux</i>	<i>Froid</i>	<i>Chaleureux</i>	<i>Froid</i>
Diversité du matériel pédagogique	<i>Pauvre</i>	<i>Riche</i>	<i>Pauvre</i>	<i>Riche</i>	<i>Pauvre</i>	<i>Riche</i>
Questionnement du maître	<i>Socratique</i>	<i>Ouvert</i>	<i>Ouvert</i>	<i>Socratique</i>	<i>Socratique</i>	<i>Ouvert</i>
Évaluation pratiquée par le maître	<i>Formative</i>	<i>Normative</i>	<i>Normative</i>	<i>Formative</i>	<i>Formative</i>	<i>Normative</i>

Comment dichotomiser cet ensemble de classes pour composer deux groupes au style pédagogique opposé ? Autrement dit, quelles classes correspondent au modèle de l'Éducation Nouvelle et lesquelles reflètent celui de l'enseignement traditionnel ? Aucune parfaitement, est-il raisonnable de répondre. La faille de ce genre de recherches provient, sans doute, de ce que les chercheurs ont apparemment supposé l'existence de deux catégories de pratiques s'observant dans leur pureté conceptuelle au niveau du terrain. D'un côté, il y aurait des classes où les initiatives laissées aux élèves sont nombreuses, le climat affectif chaleureux, la diversité du matériel riche, le questionnement du maître ouvert et l'évaluation pratiquée de type formatif. De l'autre côté, on trouverait des classes présentant les caractéristiques inverses. Mais la réalité pédagogique ne se conforme pas à cette vision manichéenne. Car, en définitive, la variabilité des paramètres permettant de décrire le fonctionnement des classes est plus grande qu'on le suppose généralement et les combinaisons de ces paramètres bien trop nombreuses pour se conformer à une simple dichotomie.

Une critique fondamentale peut, en définitive, être adressée à la comparaison de méthodes d'enseignement : cette approche comparative suppose l'existence de modèles d'enseignement *a priori* dont l'un serait supérieur à l'autre (ou aux autres). Ce faisant elle conduit à ramener la diversité des modalités d'enseignement à une simple dichotomie, présupposant en quelque sorte la colinéarité d'un ensemble de paramètres : le climat af-

fectif chaleureux irait toujours de pair avec de nombreuses initiatives laissées aux élèves ainsi qu'avec une grande diversité de matériel pédagogique, un questionnement ouvert et une évaluation de type formatif. Or, comme on le lira par la suite, ces paramètres sont largement indépendants.

L'échec des comparaisons de méthodes renvoie évidemment à la nécessité d'étudier l'enseignement tel qu'il est et non tel que les doctrines pédagogiques supposent qu'il est ou veulent qu'il soit. Même si la finalité reste l'amélioration des pratiques, le paradigme processus-produit repose sur cet axiome fondateur : étudier l'enseignement sans modèle pédagogique *a priori*. Cette position épistémologique constitue – pensons-nous – une rupture dans le champ des recherches pédagogiques. Elle oblige à passer par une première étape essentielle : celle de la description. Cette volonté a été clairement exprimée par le titre de l'ouvrage de Smith (1956) *On the Anatomy of Teaching*. En bref, il s'agit d'approcher l'enseignement avec l'attitude de l'entomologiste face aux insectes, de l'éthologiste face au comportement animal, de l'ethnologue face aux autres cultures.

Toutefois, l'analyse descriptive de l'enseignement n'exclut nullement l'étude de ses effets. Cette problématique doit simplement être abordée autrement, notamment en prenant en compte la multidimensionalité de l'enseignement et l'indépendance potentielle des paramètres qui le composent. Assez logiquement, c'est à l'approche corrélacionnelle que vont avoir recours les

chercheurs mobilisés par le projet de mettre en relation les processus d'enseignement et leurs produits. Plus précisément, le canevas des recherches processus-produit suppose, d'une part, des mesures ciblées sur les pratiques des enseignants (et éventuellement sur le comportement des élèves) pendant les activités en classe (*processus*) et, d'autre part, des mesures de résultat chez l'élève, rendement des tests cognitifs et éventuellement des tests d'attitude (*produit*). Dans les grandes lignes, le déroulement des recherches se ramène aux étapes qui suivent :

1. On observe dans un échantillon de classes, et à plusieurs reprises, les interactions maître-élèves et/ou le type d'activités organisées et le temps qui leur est dévolu. Pour ce faire, on utilise des grilles plus ou moins sophistiquées, organisées en catégories.
2. On comptabilise la fréquence des observations pour chaque catégorie retenue, ou on calcule le temps alloué aux différentes activités. On calcule la moyenne par classe, par leçon ou activité observée, ou encore par journée.
3. On met en relation les paramètres calculés à l'étape 2 avec le rendement des élèves appréhendé à travers des tests (le plus souvent, des tests de connaissance ; dans certains cas, des échelles d'attitude ou des questionnaires visant à évaluer les effets socio-affectifs de l'enseignement). Pour ce faire, on recourt au calcul de

corrélation, parfois à l'équation de régression. Le plus souvent, l'unité de mesure étant la classe, on corrèle la moyenne des paramètres de processus (mesurés pendant plusieurs leçons) avec la moyenne des scores des élèves de la classe. Dans le meilleur des cas, les élèves ont été soumis à un prétest. Celui-ci porte soit sur les connaissances qui vont être enseignées, soit sur le niveau intellectuel atteint (mesure de QI). Il est alors possible de calculer une mesure de gain ; celle-ci est généralement obtenue à l'aide de l'équation de régression : on calcule un score résiduel en soustrayant le score réellement obtenu par chaque élève du score que lui prédit la régression.

Il est utile de souligner d'emblée trois limites principales inhérentes à ce dispositif de recherche.

La plus importante consiste à rappeler qu'une corrélation n'est pas une relation de cause à effet, et, comme le soulignent Dunkin et Biddle dès 1974, certaines corrélations interprétées dans le sens suivant : « les comportements de l'enseignant influencent la performance des élèves » pourraient être renversées en : « la compétence des élèves (dont la mesure de produit est le reflet) influence la conduite du maître ». C'est dire si les chercheurs qui ont mené des recherches processus-produit étaient conscients des problèmes méthodologiques posés par leur approche de la relation entre l'enseignement et ses effets. On verra tout au long de notre revue comment

ils se sont efforcés de résoudre et/ou de contourner ces problèmes.

Le calcul de la corrélation repose sur la variance et la covariance des paramètres mis en relation. L'importance du coefficient dépend donc en partie de la variabilité des paramètres. Dit autrement, le calcul d'une corrélation implique que les deux variables aient une variance suffisante, ce qui suppose indirectement une fréquence moyenne relativement élevée. On verra que ceci a son importance en ce qui concerne les corrélations liées aux *feedbacks* et, plus encore, aux critiques.

Les différentes catégories retenues pour caractériser le processus d'enseignement sont théoriquement indépendantes. Dans la pratique, ce n'est pas nécessairement le cas et, par conséquent, le chercheur doit se méfier de phénomènes de colinéarité : si les processus A et B sont non seulement corrélés au rendement, mais aussi corrélés entre eux, il importe de s'interroger sur la signification de la relation du processus au produit. Peut-on considérer que les deux variables de processus influent sur le rendement ou, au contraire, faut-il supposer qu'une seule de ces variables est déterminante, la corrélation de la seconde avec le rendement étant positive par effet de colinéarité ?

Dans les pages qui suivent, nous débiterons notre revue par les travaux de Stallings qui reposent sur le célèbre et gigantesque *Follow Through Project*. Ensuite, nous reviendrons sur les recherches et les polémiques tournant autour de la question du *leadership* des ensei-

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

gnants. Pour ce faire, nous discuterons les travaux de Flanders et ceux, complémentaires et contradictoires, de Soar et Soar ; cette discussion sera l'occasion de clarifications conceptuelles déterminantes. Dans un chapitre ultérieur, nous aborderons deux programmes de recherche au sein desquels l'articulation des travaux constitue un point fort : avec les études de Brophy et Evertson, d'une part, et celles de Good et Grouws, d'autre part, le courant des recherches processus-produit atteint à la fois son apogée et ses limites. Des pistes de dépassement ont été explorées au sein même du paradigme. Dans l'avant-dernier chapitre, nous examinerons deux programmes de recherche. La *Beginning Teacher Evaluation Study* ou BTES présente la particularité de combiner approche quantitative et qualitative. Nous discuterons cette importante entreprise avant de revenir sur les études expérimentales de ce qui fut appelé le groupe de Canterbury (Nuthall, 1968). Enfin, nous clôturerons cet opuscule par une mise en perspective des recherches anglo-saxonnes qui ont tenté de mettre en relation les pratiques d'enseignement et leurs produits.

CHAPITRE II

LE PROJET *FOLLOW THROUGH* ET LES RECHERCHES DE STALLINGS

L'apport de Stallings à l'étude des relations entre les processus d'enseignement et les produits d'apprentissage est principalement lié au projet *Follow Through*. Celui-ci s'inscrit dans le mouvement d'éducation compensatoire qui a animé le monde de l'enseignement aux États-Unis pendant les années 60 et 70. L'objectif de tous ces programmes consiste à lutter contre les inégalités d'origine socioculturelle. L'idée fondatrice est simple : donner aux enfants issus de minorité ethnique et/ou culturelle des opportunités éducatives plus grandes. Dans un premier temps, ce mouvement d'éducation compensatoire a porté ses efforts sur le niveau préscolaire. Ceci a abouti au programme *Head Start*. L'évaluation de ces programmes préscolaires ayant montré que les effets engendrés ne persistaient pas au niveau primaire, le Congrès fédéral a pris la décision de promouvoir un effort comparable au

niveau primaire. Ceci a conduit au programme *Follow Through*. Comme pour le programme *Head Start*, le Congrès américain a lancé un appel d'offre et a retenu une série de projets expérimentaux menés soit par des centres universitaires, soit par des associations d'enseignants dont la finalité était la réduction des inégalités sociales en matière d'éducation. Les appellations *Head Start* et *Follow Through* regroupent donc des programmes éducatifs variables, mais qui, tous, ont l'objectif d'apporter un plus à l'éducation des élèves dits défavorisés. Afin d'établir le profil des programmes efficaces, Stallings et Kaskowitz (1974) ont mené une étude de type processus-produit sur 108 classes de première année primaire (1P) et 58 classes de troisième année (3P). Toutes les classes observées sont prises en charge par des enseignants chevronnés, qui ont accepté de participer à l'un des sept modèles expérimentaux repris dans le projet *Follow Through*.

Par la suite, Stallings a également conduit une étude corrélationnelle au niveau de la troisième année primaire suivie d'une recherche expérimentale sur l'enseignement de la lecture au niveau secondaire.

The Follow Through Evaluation Study

Dans cette importante étude, chaque classe de 1P et de 3P a été observée pendant trois jours consécutifs. Pendant deux journées les observateurs se centrent sur le maître et, la troisième journée, sur un petit échantillon d'élèves. La grille d'observation a été construite de ma-

nière à prendre en considération les aspects distinctifs des différents programmes. Pour l'essentiel, on s'attache à décrire les caractéristiques de l'environnement physique, à recueillir des données sur le temps alloué aux différentes activités et à enregistrer la fréquence de différents types d'interactions maître-élèves. Parmi les programmes concernés, on en trouve certains qui privilégient un haut degré de structuration et une forte centration sur les compétences de base (lecture, écriture, calcul). À l'inverse, d'autres modèles privilégient la pédagogie active ainsi que les objectifs affectifs et l'auto-apprentissage.

Les résultats sont clairs : les deux programmes qui enregistrent les gains les plus élevés en lecture et en mathématiques sont également ceux qui privilégient les objectifs correspondants et s'efforcent de les atteindre par des activités d'enseignement structurées. Il est intéressant de remarquer que ces deux programmes étaient implantés dans des milieux particulièrement défavorisés : les élèves de ces classes se caractérisaient au départ par les scores les plus bas au *Progressive Matrix* de Raven, ainsi qu'aux échelles d'attitude portant sur l'intérêt manifesté vis-à-vis de l'école. D'une manière générale, Stallings et Kaskowitz montrent que les paramètres qui maximisent les gains au test standardisé de connaissances ne sont pas nécessairement les mêmes que ceux qui maximisent les progrès relatifs à d'autres critères.

L'observation des classes montre que la plupart des enseignants ont suivi scrupuleusement les lignes direc-

trices données par les responsables des programmes auxquels ils adhéraient. Ceci confère un grand intérêt à l'étude de Stallings et Kaskowitz. En effet, la diversité des programmes sous étude est grande. On se trouve donc face à un corpus de pratiques d'enseignement extrêmement riche par sa diversité. Ainsi on observe des combinaisons de facettes que l'on ne trouve quasiment jamais dans des classes tout venant. Un bon exemple de ceci est donné par le *Kansas Program*. Ce programme, conçu pour des classes de première année primaire, recommande (1) l'organisation fréquente de petits groupes d'enseignement centrés sur les compétences de base et animés par un enseignant, un assistant d'enseignement et deux parents volontaires ; (2) l'usage intensif de matériaux programmés d'apprentissage individualisé ; (3) le tout, soutenu par un dispositif de louanges et de récompenses (matérielle ou symbolique) (*tokens*) organisé selon les principes behavioristes de l'apprentissage. On se trouve donc en présence de classes où les élèves sont renforcés selon le principe des *tokens* pour leurs progrès et attitudes scolaires. Mais ces classes recourent également à des procédures d'enseignement préconisées par l'Education Nouvelle : l'apprentissage individualisé et l'enseignement en petits groupes.

On retrouve, dans l'étude de Stallings et Kaskowitz, le problème classique des études menées à large échelle et utilisant de nombreuses mesures du processus d'enseignement : une multitude de résultats impliquant plus de variables que de classes. Ainsi, par exemple, Stallings et

Kaskowitz ont calculé 340 corrélations portant sur les variables relevées dans les 108 classes de 1P ; 108 de ces corrélations sont significatives au niveau 0,05 pour les mathématiques et 118 en ce qui concerne la lecture. Cela fait 226 corrélations significatives au seuil de P.0,05. Bref, on est face à un corpus imposant de relations processus-produit significatives, mais le coefficient de probabilité ne peut être interprété au pied de la lettre : les 340 variables de processus retenues ne sont pas mutuellement exclusives sur le plan conceptuel et, par conséquent, la probabilité d'obtenir des corrélations significatives par effet de colinéarité est élevé. On ne peut donc donner de réelle signification au seuil de 0,05 : il peut cependant aider à distinguer les corrélations qui méritent d'être prises en considération des autres.

Le *pattern* de corrélations le plus évident et le plus solide a trait aux gains enregistrés aux tests de connaissance. Ceux-ci sont corrélés positivement avec toutes les variables de processus qui indiquent qu'un temps de classe important a été consacré aux activités scolaires correspondant aux compétences reprises dans le test : temps alloué aux activités d'enseignement centrées sur les compétences de base, fréquence des leçons focalisées sur ces compétences, que celles-ci soient organisées sur un mode frontal, en petits groupes ou encore sous la forme de travaux individuels supervisés par l'enseignant. De façon complémentaire, on observe des corrélations négatives entre les mesures de gain et les paramètres indiquant qu'un temps important a été dévolu aux activi-

tés non académiques (musique, danse, bricolage). Sont également corrélés par exemple négativement avec l'apprentissage scolaire les paramètres qui reflètent principalement les options de l'Education Nouvelle (organisation en aire ouverte, périodes de temps durant lesquelles l'enseignant travaille avec un ou deux élèves plutôt que de faire la leçon à l'ensemble du groupe) ainsi que les interactions maître-élèves ne portant pas directement sur le contenu d'enseignement. D'une manière générale, toutes les variables apparentées au schéma d'enseignement traditionnel sont corrélées positivement avec le rendement scolaire : enseignement frontal suivi de questions précises d'ordre factuel, avec *feedbacks* correctif précis.

La tendance générale des résultats est claire : pour ce qui est des compétences de base les élèves progressent d'autant plus que l'enseignement porte de façon directe et explicite sur ces objets d'apprentissage. À l'opposé, les progrès sont pauvres lorsque l'enseignant mise sur l'auto-apprentissage ou lorsqu'il organise de fréquentes activités non académiques espérant, en quelque sorte, un effet indirect sur les compétences de base. Ces constats sont d'autant plus forts que les élèves impliqués sont d'origine modeste. On se rappellera que toutes les classes sont fréquentées par des élèves dits défavorisés, donc issus de familles provenant de minorités ethniques. Au sein de cet échantillon d'élèves défavorisés, certains le sont plus que d'autres. L'utilité d'une approche structurée et orientée très explicitement vers les apprentissages de base est

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

d'autant plus manifeste que les élèves sont d'origine modeste.

D'autres résultats de seconde importance méritent également d'être mentionnés.

1. Au niveau 1P, le rendement est corrélé positivement avec le temps passé en petit groupe (moins de huit élèves), tandis qu'au niveau 3P, il est corrélé positivement avec le temps passé en grand groupe (enseignement frontal).
2. Aux niveaux 1P et 3P, On observe un faisceau de corrélations ayant trait aux encouragements, aux critiques et aux *feedbacks*. Dans ce cas, les coefficients de corrélation sont inférieurs à ceux observés lorsqu'on prend en considération le temps consacré aux différents types d'activités. Par ailleurs, le *pattern* de résultats observé est relativement hétérogène, ce qui rend l'interprétation des tendances délicates. Épinglons quelques-unes des corrélations observées avant de les discuter.

L'utilisation de jetons comme renforcement est corrélée positivement avec les gains au niveau 1P, mais pas au niveau 3P.

D'une manière générale, les encouragements ou louanges consécutifs à une bonne réponse (ou motivés par un bon travail scolaire) tendent à être liés positivement aux gains d'apprentissage. L'intensité de cette relation varie en fonction de plusieurs paramètres. La

corrélation observée en 1P est plus forte que celle observée en 3P. Elle est plus forte lorsque l'on prend pour critère le test de mathématiques comparativement à celui de lecture. Enfin, la corrélation est plus forte lorsqu'on la calcule pour les élèves faibles (c'est-à-dire ceux caractérisés au départ par un faible niveau d'aptitude) plutôt que pour les élèves forts.

Les *feedbacks* correctifs exprimés de façon neutre, c'est-à-dire sans émettre d'approbation ni de désapprobation, corrélent le plus souvent positivement avec le rendement. De façon plus surprenante, une tendance analogue apparaît en ce qui concerne les *feedbacks* correctifs précédés d'une désapprobation (critique académique ou *feedbacks* correctif négatif). Notons que cette tendance ne dépasse pas toujours le seuil de signification requis (0,05).

L'interprétation des résultats portant sur les *feedbacks* appelle deux commentaires méthodologiques :

1. Par essence, tous les types de *feedbacks* sont fonctionnellement liés aux réponses des élèves. Logiquement, on observe donc davantage de *feedbacks* lorsque la participation verbale des élèves est abondamment sollicitée. On se trouve dès lors face à un nouveau risque de colinéarité : une corrélation positive entre les gains d'apprentissage et les *feedbacks* peut se superposer à une liaison positive entre le rendement et un *pattern* d'enseignement fait de nombreuses sollicitations. Pour éviter ce biais, il est indi-

qué de calculer, pour chaque type de *feedbacks*, une fréquence pondérée en fonction du nombre de réponses émises par les élèves pendant l'activité. C'est ce qui sera fait dans des études ultérieures, notamment celles de Brophy et Evertson et celles de Good et Grouws (cf. ci-dessous).

2. D'une manière générale, les *encouragements académiques* et plus encore les critiques académiques sont de faible fréquence : en moyenne, on observe une critique pour 100 mauvaises réponses d'élèves et 0,5 critique pour 100 réponses de tous types (bonnes ou mauvaises). Il est, partant, logique d'observer de faibles corrélations entre ces paramètres et les gains d'apprentissage. En outre, ce constat empêche d'opposer des enseignants qui critiqueraient souvent et d'autres qui critiqueraient fort peu. Plus justement, il faut situer leurs différences entre un pôle où l'on n'observe aucune critique et un autre où l'on observe des critiques rares. Autrement dit, la corrélation positive entre les critiques académiques et le rendement ne doit pas être interprétée comme le signe qu'une attitude fondamentalement critique est profitable aux élèves. Il semble plus juste de considérer que des critiques occasionnelles peuvent avoir un effet positif sur l'apprentissage des élèves. Celles-ci, consécutives à des signes manifestes d'inattention ou de comportements désinvoltes, sont peut-être interprétées par

les élèves comme l'expression d'une attente positive teintée d'exigence de la part de l'enseignant.

The California ECE Study

En 1977, Stallings, Cory, Fairweather et Needle publient les résultats d'une étude portant sur l'évaluation d'un programme – *The California Early Childhood Education Program*, dont l'objectif est d'améliorer l'enseignement de base destiné à des élèves en difficulté (*low achievers*). Le programme ECE, financé par l'État de Californie, fournit aux écoles-pilotes un personnel complémentaire et fait participer certains parents aux activités scolaires. Il s'adresse à des classes sélectionnées au sein d'écoles dont le score moyen à un test d'entrée est inférieur au 20^e percentile : c'est dire le caractère défavorisé de ces classes. Un élément substantiel distingue cette étude de l'évaluation des programmes *Follow Through* : les classes ECE n'ont pas à respecter les lignes directrices d'un programme conçu par des spécialistes ; l'enseignement est organisé en fonction d'un projet pédagogique local.

Des observations ont été menées dans 45 classes de 3P. Stallings a retenu 49 variables en fonction des résultats de son étude précédente. Parmi celles-ci, 33 sont corrélées significativement avec le rendement en lecture. Pour l'essentiel, les résultats confirment ceux obtenus dans l'étude précédente. Les classes où l'on enregistre les gains les plus élevés se caractérisent par un temps impor-

tant dévolu à l'enseignement de la lecture et à celui des autres matières académiques. Inversement, on enregistre dans ces classes peu de temps alloué aux jeux ou à d'autres activités de socialisation. Autre variable significative : le temps durant lequel les enseignants sont en contact avec les élèves à propos des disciplines académiques. Plus précisément, il apparaît que plus l'enseignant consacre du temps à organiser ou à régler des problèmes de discipline posés par des élèves isolés, moins l'ensemble des élèves apprend. De façon complémentaire, les gains d'apprentissage sont moindres lorsque le temps d'enseignement consacré à des élèves isolés est élevé, c'est-à-dire lorsque l'enseignant procède par individualisation. Les gains sont importants lorsque l'enseignant privilégie l'organisation collective. Dans ces classes caractérisées par une haute implication dans les apprentissages scolaires, on voit encore que les enseignants donnent beaucoup d'explications, posent beaucoup de questions portant sur les apprentissages et réagissent par de nombreux *feedbacks*. Quant aux élèves, ils posent fréquemment des questions et initient des échanges avec le maître portant sur les disciplines académiques. Notons enfin que, dans la présente étude portant sur des classes de 3P, l'organisation de petits groupes est liée positivement au rendement, alors que ce n'était pas le cas pour ce niveau d'enseignement dans l'étude *Follow Through*.

La troisième étude de Stallings et le *Descriptive-Correlational-Experimental Loop*

Dans un article consacré au bilan des recherches sur l'enseignement, Rosenshine et Furst (1973) discutent en profondeur les limites de l'approche corrélationnelle. Remarquant que cette approche ne permet pas de mettre en évidence des liens de causalité, ils suggéraient d'adopter un schéma de recherche qu'ils nomment *descriptive-correlational-experimental loop*. Pour l'essentiel, il s'agit d'articuler plusieurs étapes de recherches : décrire et analyser le processus d'enseignement, d'abord, enquêter sur les liens processus-produit au moyen du calcul de corrélation, ensuite et, enfin, tirer parti des corrélations significatives observées pour formuler des recommandations qui sont transmises à un groupe d'enseignants, constituant de ce fait le groupe expérimental (GE). Parallèlement, est constitué un groupe de contrôle (GC) composé de classes aussi semblables que possible aux classes expérimentales. Un prétest et un post-test encadrent les deux groupes.

Dans un troisième programme de recherche, Stallings organise ses investigations selon ce schéma : une étude corrélationnelle (Stallings, Cory, Fairweather & Needle, 1978) est suivie par un *training experiment* (Stallings, Needle & Sayrook, 1979).

L'étude corrélationnelle porte, à nouveau, sur l'enseignement de la lecture, mais concerne cette fois-ci l'enseignement secondaire. Elle implique l'observation de 43 classes qui, toutes, accueillent des élèves en difficulté de

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

lecture. Une fois encore, la quantité d'enseignement ciblé sur les apprentissages visés constitue la clé du succès. Le tableau qui suit résume les principales tendances.

Tableau 2 : Principales corrélations observées par Stallings, Cory, Fairweather et Needle (1978)

Sont corrélés avec les gains	
Positivement	Négativement
<ul style="list-style-type: none"> - Le temps durant lequel les élèves lisent à voix haute - Le temps consacré à l'enseignement de la lecture en petit ou en grand groupe, ainsi que le temps consacré à revoir ou à discuter des travaux fournis aux élèves - La fréquence des <i>feedbacks</i> positifs (consécutifs à des bonnes réponses) - La fréquence des <i>feedbacks</i> correctifs (lorsque les élèves ne donnent pas la réponse correcte) 	<ul style="list-style-type: none"> - Le temps durant lequel l'enseignant n'interagit pas avec les élèves et s'occupe de tâches administratives - Le temps durant lequel il donne des consignes d'organisation plutôt que d'enseigner - La fréquence des activités au cours desquelles les élèves ont le choix des activités et/ou des objectifs - Le temps durant lequel les élèves travaillent individuellement à une lecture silencieuse ou sur du matériel écrit - La fréquence des interactions sociales qui n'ont pas pour objet les apprentissages scolaires - La fréquence des interactions négatives

En résumé, les gains sont de faible importance lorsque l'enseignement n'est pas ciblé sur la lecture. Il faut, de surcroît, faire état de spécificités liées au niveau initial des élèves en lecture. Avec les élèves dont le niveau est particulièrement bas, les enseignants les plus efficaces privilégient les activités de déchiffrage, tout en réussissant à combiner cet accent sur le décodage avec une attitude encourageant la compréhension du sens. De façon synthétique, la leçon *efficace* se présente comme suit : explication du vocabulaire et des concepts, lecture orale du texte ponctuée de questions pour développer et vérifier la compréhension, encouragements et louanges pour les bonnes réponses, aide et *feedbacks* correctif en cas de mauvaise réponse ou de difficulté. Avec les élèves dont le niveau correspond à une ou deux années de retard par rapport à ce qui est attendu au secondaire, l'enseignant efficace organise moins de lecture à haute voix et davantage d'activités de lecture silencieuse et de travaux sur du matériel écrit. Le temps d'enseignement frontal reste important. Toutefois, de longues minutes sont consacrées à des travaux individuels supervisés scrupuleusement par l'enseignant qui parcourt les rangées de bancs, prêt à intervenir lorsque les élèves le sollicitent ou prêt à relancer l'activité des élèves qui, pour une raison ou une autre, se détournent de leur tâche.

Dans la *training experiment*, une série de recommandations inspirées des résultats évoqués ci-dessus ont été transmises et expliquées aux enseignants du GE. Une répartition précise du temps des leçons leur est suggé-

rée : celles-ci doivent toutes comporter des moments d'activités collectives, des travaux de groupe et des phases de travaux individuels supervisés. Des conseils sont également donnés sur la façon de poser des questions et de réagir aux réponses des élèves. Les recommandations formulées par Stallings et son équipe tiennent également compte du niveau de départ des élèves en lecture.

L'analyse indique que les enseignants du GE se distinguent par des degrés variables d'implantation des consignes. Il apparaît notamment que, n'étant guère habitués à demander à leurs élèves de lire à voix haute, les enseignants du niveau secondaire ont peu appliqué cette recommandation. Quoiqu'il en soit, le traitement statistique indique que plus le comportement des enseignants s'approche des lignes directrices formulées par Stallings et al., plus les gains des élèves sont élevés. Cette tendance nette ne conduit toutefois pas à une différence significative entre le groupe expérimental et le groupe de contrôle. Bref, succès il y a, mais il reste mitigé.

CHAPITRE III

POLÉMIQUE AUTOUR DU LEADERSHIP DE L'ENSEIGNANT

Les recherches de Flanders : influence directe *versus* influence indirecte

S'inspirant des constats réalisés par Lewin, Lippitt et White (1939) quant aux effets des différents styles de *leadership* (autoritaire, laisser-faire et démocratique), Flanders (1960) formule l'hypothèse selon laquelle un style d'enseignement de type démocratique est plus efficace qu'une approche pédagogique directive. En termes plus précis, cela signifie que l'apprentissage des élèves et leur sentiment de bien-être seraient supérieurs là où les enseignants parlent peu et exercent une influence indirecte. À l'opposé, l'efficacité pédagogique serait moindre là où les enseignants parlent beaucoup et se complaisent dans de fréquents exposés magistraux.

Conséquent avec cette hypothèse générale, Flanders (1960) a construit une grille d'observation (la FIAC) dont le vecteur est l'opposition « influence indirecte »

versus « influence directe » de l'enseignant. Dans la version originale de la FIAC, l'influence indirecte regroupe quatre catégories : (1) Accepte les sentiments des élèves (2) Fait des éloges ou encourage (3) Accepte ou utilise les idées des élèves (4) Pose des questions. L'influence directe se décline en trois catégories : (5) Fait un exposé *ex cathedra* (6) Donne des directives (7) Critique ou en appelle à son autorité. Complémentairement, la FIAC oppose également (8) Réponses de l'élève aux questions de l'enseignant et, (9) Prises de parole spontanées.

Eu égard à cette grille, les hypothèses spécifiques de Flanders (1970) sont les suivantes :

H 1 L'influence indirecte est liée positivement aux attitudes des élèves (vis-à-vis de l'école) et à leurs apprentissages (rendement aux tests de connaissances). Ce qui (techniquement) consiste à supposer une corrélation positive entre la fréquence des comportements relevant de l'influence indirecte d'une part et, d'autre part, les mesures d'attitude et de rendement.

H 2 À l'opposé, l'influence directe et, plus particulièrement, toutes les manifestations d'autorité excessive de l'enseignant sont liées négativement avec, à la fois, les attitudes des élèves et leurs apprentissages. Autrement dit, il y aurait une corrélation négative entre la fréquence des comportements d'autorité d'une part et, d'autre part,

les mesures d'attitude et de rendement.

H 3 Les prises de parole de l'enseignant sont liées négativement avec les attitudes des élèves ainsi qu'avec leurs apprentissages. Autrement dit, il y aurait une corrélation négative entre la fréquence des prises de parole du maître et les mesures d'attitude et de rendement aux tests.

H 4 Les comportements négatifs de l'enseignant (c'est-à-dire les critiques adressées aux élèves, mais aussi les *feedbacks* négatifs) sont liés négativement avec les attitudes des élèves et avec leurs apprentissages. Autrement dit, il y aurait une corrélation négative entre la fréquence des comportements négatifs du maître et les mesures d'attitude et de rendement.

H 5 Les comportements positifs de l'enseignant (c'est-à-dire les approbations et autres marques de sollicitudes) sont liés positivement avec les attitudes des élèves ainsi qu'avec leurs apprentissages.

H 6 Enfin, la flexibilité de l'enseignant – c'est-à-dire sa capacité à exercer à l'occasion une influence directe – est liée positivement, tant avec les attitudes des élèves qu'avec leurs apprentissages.

Entre 1959 et 1967, Flanders (1970) a mené cinq études visant à mettre ces hypothèses à l'épreuve des faits. Le tableau 3 donne un premier aperçu de ces cinq recherches.

Tableau 3 : Descriptions synthétiques des études menées par Flanders

	Année	Lieu	Total (classes)	Classes observées	Niveau
1	1959-60	Minnesota	38	16	Gr 8 (3S)
2	1959-60	Minnesota	37	15	Gr 7 (2S)
3	1964-65	Michigan	101	30	Gr 5 (1S)
4	1965-66	Michigan	72	16	Gr 4 (4P)
5	1966-67	Michigan	100	16	Gr 2 (2P)

Flanders adopte le même plan, en trois dimensions, dans les cinq études, à quelques détails près.

- 1) Un échantillon représentatif des classes d'une région géographique donnée est sélectionné. Par exemple, dans l'étude 1, 38 classes du *Grade 8* (soit l'équivalent de la 2^e année du secondaire) ont été sélectionnées dans le Minnesota. Dans ces classes, un inventaire d'attitudes est présenté aux élèves en début d'année. La moyenne des scores à l'inventaire est calculée pour chaque classe. Seules les classes situées aux extrémités de la distribution des résultats sont retenues pour l'observation. Ainsi, dans l'étude 1, seulement 16 classes (8 à chaque extrémité de la distribution) ont été retenues. Notons cependant que pour les études 3 et 4, des classes du milieu de la distribution ont également été conservées. Enfin, dans la dernière étude, on se trouve face à un échantillon de juge-

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

ment : les classes ont été sélectionnées de sorte à maximiser la variabilité des interactions enseignant-élèves.

- 2) Dans les classes sélectionnées, les interactions maître-élèves sont observées et codées par des observateurs entraînés à l'aide de la FIAC. Les observateurs notent les interactions dans la classe pendant trois secondes, puis les codent dans une des dix catégories de la grille. Ensuite, ils observent à nouveau pendant trois secondes. Les enseignants travaillent dans leur classe habituelle avec leurs élèves, mais quand ils sont observés ils enseignent des unités de cours dites expérimentales. Celles-ci sont proches des unités d'enseignement habituelles, mais portent sur des matières différentes. Elles leur sont proposées toutes préparées par les chercheurs. Cette procédure a pour fonction de neutraliser l'intervention de facteurs dont on considère qu'ils affectent les apprentissages scolaires : le contenu enseigné et le degré de maîtrise par les enseignants de ce contenu.
- 3) Les apprentissages scolaires des élèves sont mesurés, par un prétest et un post-test, avant et après ces leçons. Le prétest permet, en outre, de prendre en considération les habiletés initiales des élèves dans les analyses.

À partir des codages obtenus lors des observations en classes, les chercheurs calculent les fréquences pour

chacune des catégories de la FIAC. Ensuite ils regroupent certaines catégories pour obtenir les dix indicateurs définis au tableau 3. Notons d'emblée le statut ambigu de la catégorie 4 (*pose des questions*), parfois considérée comme composante de l'influence indirecte et parfois pas ; nous y reviendrons.

Pour les 5 études, la corrélation entre chaque indicateur est calculée avec la moyenne (par classe) des résultats des élèves aux tests de connaissances, d'une part, et des scores d'attitude, d'autre part. Cela aboutit au tableau 4 (ci-contre). De ces résultats, on relèvera principalement les constats suivants.

Les enseignants ne parlent pas trop puisque, dans les cinq études, la variable *prise de parole du professeur* est corrélée positivement tant avec les apprentissages qu'avec les attitudes des élèves.

Les corrélations impliquant l'influence indirecte (variables 1 et 2) et les approbations tendent à être positives, alors que celles relatives aux comportements négatifs et aux manifestations fortes d'autorité tendent à être négatives.

Les scores de flexibilité sont généralement corrélés positivement avec les résultats et les attitudes des étudiants⁷.

⁷ Le ratio I/D est calculé séparément pour chaque observation de classe. Pour ce faire, on calcule la somme de 1 + 2 + 3 divisée par la somme de 1 + 2 + 3 + 6 + 7. Ensuite le plus bas de ces ratios est soustrait du plus élevé pour obtenir la dispersion.

Tableau 4 : Correlations observées dans les études de Flanders (1970)

Variable	Corrélation avec les scores de connaissances							Corrélation avec les scores d'attitude						
	2	4e	6 ^e	7 ^e	8	2	4 ^e	6	7	8				
Niveau d'étude/de grade														
Proportion de comportements indirects soit la somme des codes de (1) + (2) + (3) divisée par somme de (1) + (2) + (3) + (6) + (7)	-0,07	0,31	0,22	0,48	0,43	0,13	0,64	0,49	0,34	0,58				
Fréquence des comportements d'acceptation des idées des élèves, soit la somme des codes (3)	-0,45	0,19	0,30	0,40	0,19	0,13	0,52	0,40	0,33	0,31				
Fréquence des comportements indirects, soit la somme des codes (1) + (2) + (3) + (4)	0,05	-0,08	0,26	0,25	0,45	0,45	0,34	0,40	0,16	0,51				
Questions – soit la somme des codes (4)	0,07	-0,19	0,11	-0,06	0,44	0,49	-0,06	0,27	0,00	0,47				
Prises de parole du professeur soit la somme des codes (1) à (7)	0,30	0,08	0,11	0,02	0,45	0,38	0,10	0,24	0,15	0,61				
Prises de parole approuvées - soit la somme des codes (8) ou (9) suivis par un code (2)	-0,45	0,19	0,30	0,40	0,19	0,13	0,52	0,40	0,33	0,31				
Comportements négatifs (donne des ordres, critique) - Soit la somme des codes (6) et (7)	-0,10	-0,24	-0,04	-0,61	-0,34	-0,09	-0,17	-0,37	-0,43	-0,66				
Réactions <i>restrictives</i> - soit la somme des codes (8) et (9) qui étaient suivis des codes (6) et (7)	0,18	-0,34	-0,32	-0,50	-0,43	0,02	-0,32	-0,29	-0,47	-0,62				
Imposition, autorité négative - soit la somme des codes (6) suivis par codes (7) + somme des codes (7) suivis par codes (6)	0,05	-0,23	-0,15	-0,62	-0,25	-0,22	-0,22	-0,32	-0,43	-0,59				
Louanges - soit la somme des codes (2)	0,25	-0,13	0,36	-0,23	0,30	0,08	0,40	0,35	-0,34	0,38				
Flexibilité – Ecart entre le ratio/D le plus bas et le plus élevé	-0,07	0,46	0,19	0,37	0,43	0,12	0,08	0,41	0,13	0,43				
Nombre de classes	15	16	30	15	16	15	16	30	15	16				

Ceci reflète probablement le besoin et la pertinence d'adapter les stratégies d'enseignement à la situation, au lieu d'essayer de maximiser l'indirectivité à tout prix.

Les corrélations négatives tendent à être plus fortes et plus consistantes que les corrélations positives surtout en ce qui concerne les tests de connaissance ; ce phénomène se marque moins en ce qui concerne l'attitude des élèves.

Les corrélations obtenues dans l'étude concernant le *Grade 2* (2P) divergent fortement par rapport au schéma général décrit ci-dessus. Pour Brophy et Good (1986), ce constat pourrait s'expliquer par la nature des apprentissages assignés à ce niveau de la scolarité : techniques de lecture, algorithmes de calcul ; bref, des compétences supposant une part importante d'automatisation et, inversement, une faible part de raisonnement abstrait. Cette hypothèse explicative repose, en fait, sur les résultats des études menées par Soar et Soar (cf. ci-dessous) qui tendent à montrer que l'efficacité d'un *leadership* indirect dépend du type d'apprentissage visé.

Faut-il conclure des données produites par Flanders que, mis à part le début de la scolarité primaire, les effets de l'enseignement sur l'apprentissage des élèves et leur épanouissement socio-affectif sont d'autant plus importants que les maîtres exercent une influence indirecte ? Pour Brophy et Good (1986), ce serait aller vite en besogne et, ce pour plusieurs raisons. Ces auteurs rappellent, d'abord, qu'une corrélation ne reflète pas nécessairement une relation de cause à effet. Des résultats pro-

duits par Flanders, on pourrait tout aussi bien conclure que les résultats des élèves aux tests reflètent avant tout leur scores d'aptitude et que la relation entre scores élevés aux tests et taux élevé d'influence indirecte doit être interprétée en sens inverse : c'est le niveau d'aptitude des élèves qui détermine le degré d'influence indirecte de l'enseignant. Autrement dit, le niveau des élèves pourrait être la cause du style de *leadership* adopté par l'enseignant. Turner (1971) fut le premier à soulever cette question. Il suggérait que les élèves brillants produisent certains comportements qui augmentent la probabilité qu'un enseignant adopte un style indirect. Notamment, il remarquait que les étudiants qui ont des performances scolaires inférieures à la moyenne ont probablement moins d'idées et prennent moins souvent d'initiatives ; ceci rendrait plus difficile la possibilité de produire un *pattern* d'enseignement indirect. L'interaction en classe est, en effet, interactive, ce qui rend la suggestion de Turner tout à fait plausible. Flanders accepte pleinement cette éventualité et admet qu'il est probablement plus difficile pour l'enseignant de maintenir un *pattern* indirect avec des élèves de faibles compétences.

Par ailleurs, Brophy et Good (1986) contestent l'inclusion de la catégorie *pose des questions* dans l'influence indirecte, ainsi que celle *donne cours* dans l'influence directe. En outre, ils suspectent que les données recueillies par Flanders et, en particulier la corrélation entre rendement et influence indirecte, soient affectées d'un effet de colinéarité : les variables *scores des*

étudiants aux tests et degré d'influence indirecte des enseignants seraient toutes deux liées à une troisième variable dont le pouvoir explicatif serait déterminant. A l'appui de cette interprétation, Brophy et Good (1986) avancent plusieurs arguments fondés sur un examen scrupuleux des données recueillies par Flanders lors de sa deuxième étude. Celles-ci sont reproduites dans le tableau 4. On y trouve les fréquences moyennes de comportements observés durant les leçons de mathématiques et de sciences sociales de deux types de classes : sur la base d'observations du comportement des enseignants, certaines sont considérées comme appartenant à la catégorie *fonctionne de façon indirecte* tandis que d'autres classes appartiennent à la catégorie *fonctionne de façon directe*.

Sur l'ensemble des 31 classes, seulement 6,91 % des comportements relèvent de *l'influence indirecte* (du moins, lorsqu'on considère que les questions ne relèvent pas de cette catégorie générique) et seulement 9,69 % sont repris sous la rubrique *influence directe*. L'essentiel de l'action des enseignants consiste à *poser des questions* (10,90 %) et à *faire cours* (37,67 %). Autrement dit, les comportements d'influence indirecte à proprement parler ne représentent qu'une faible proportion de l'ensemble des événements qui surviennent en classe. Par ailleurs, lorsqu'on contraste les classes *indirectes* avec les classes *directes*, on constate bien que le pourcentage de la catégorie *influence indirecte* est supérieur dans les premières mais il en va de même pour la catégorie *fait un cours ex*

cathedra. On ne peut donc exclure qu'il y ait colinéarité et, partant, que la supériorité des classes dites *indirectes* s'explique par la fréquence des comportements enregistrés sous la catégorie *fait un cours* plutôt que par la catégorie *influence indirecte* elle-même. Cette interprétation paraît d'autant plus plausible que l'analyse fine des données indique que les comportements d'influence indirecte apparaissent quand l'enseignant présente des informations, gère une activité de récitation, propose des exercices ou encore, dirige une discussion en grand groupe. Ceci renforce donc l'hypothèse selon laquelle les enseignants dits *indirects* obtiendraient un meilleur rendement, non pas parce qu'ils utilisent des méthodes indirectes mais parce qu'ils font plus d'enseignement collectif. En définitive, la démonstration à laquelle croyait être arrivé Flanders est peu convaincante (voir tableau 5 ci-après ; source : N. Flanders, 1965).

Par ailleurs, malgré sa popularité, la FIAC paraît inappropriée pour rendre compte de la complexité des processus d'enseignement.

A cet égard, on retiendra deux des critiques constructives qui lui sont adressées par Brophy et Good (1986) :

1. Les interactions de gestion et les interactions académiques devraient être codées séparément : ceci pour éviter diverses confusions dans l'interprétation des conduites produites par les enseignants.

Tableau 5 : Répartition moyenne (en %) des comportements d'enseignants et d'élèves codés selon la FIAC. Observations faites dans 4 types de classes de "junior supérieur"

	Types de classes (pourcentages des interactions observées)				Total
	Math (indirect)	Math (direct)	Sciences humaines (indirect)	Sciences humaines (direct)	
Nombre de classes	7	9	7	8	31
Nombre d'interactions observées	26,083	32,726	28,194	23,641	110,644
Influence indirecte					
	0,23	0,11	0,11	0,03	0,12
Accepte les sentiments de l'élève					
Fait des éloges et encourage	1,69	1,06	1,25	1,14	1,2
Accepte ou utilise les idées de l'élève	8,11	2,63	8,28	3,03	5,51
Total influence indirecte	10,03	3,80	9,64	4,20	6,91
Pose des questions	12,52	9,53	10,75	10,80	10,90
Fait un cours <i>ex cathedra</i>	46,72	40,83	37,45	25,67	37,67
Influence Directe					
Donne des directives	3,38	8,64	4,29	9,86	6,54
Critique, en appelle à son autorité	0,94	4,66	1,69	5,32	3,15
Total influence directe	4,32	13,30	5,98	15,18	9,69
Elève					
Répond aux questions	10,73	13,02	17,54	21,49	15,70
Prend spontanément la parole	6,12	6,74	9,48	8,70	7,76
Silence, confusion	9,56	12,79	9,16	13,94	11,36

La plupart des *directives* (code 6) émanant des enseignants sont relatives à la procédure et la plupart de leurs *critiques* (code 7) sont dues au manque de discipline. Ceci a pour conséquence que la catégorie *influence directe* reflète autant une façon d'enseigner qu'une façon de faire front aux problèmes de discipline.⁸

2. Par ailleurs, la fréquence des éloges et des critiques dépend de la fréquence des réponses correctes qui, elles-mêmes, dépendent pour partie des aptitudes des élèves et pour une autre partie des compétences de l'enseignant à structurer la matière et à interroger les élèves. Ces indicateurs reflètent donc mal le savoir-faire propre aux enseignants.

Ces problèmes de catégorisation des pratiques d'enseignement en liaison avec des *construits* pédagogiques sont au cœur de l'apport de Soar et Soar.

Soar et Soar : des clarifications conceptuelles aux vérifications empiriques

L'apport de Robert Soar est déterminant tant sur le plan conceptuel qu'empirique. Cet auteur a conduit cinq

⁸ En outre, la grille ne distingue ni les *feedbacks* positifs des éloges, ni les *feedbacks* négatifs des critiques. Ceci a pour effet d'accroître la confusion lorsqu'il s'agit d'interpréter un taux élevé de critiques ou de louanges.

études processus-produit dans les années 60 et 70, plusieurs en collaboration avec Ruth Soar. La première (Soar, 1966) porte sur 55 classes situées en région urbaine ; elle s'étale sur les niveaux troisième à cinquième primaire. Les élèves fréquentant ces classes sont de race blanche et issus en majorité de familles moyennes ou favorisées. La seconde étude (Soar et Soar, 1972) porte sur 20 classes de première année intégrées dans le projet d'éducation compensatoire (*Follow Through*). Les élèves de ces classes, d'origines ethniques diverses, sont issus de familles modestes. La troisième étude (Soar & Soar, 1973 et 1978) concerne 59 classes de cinquième année. À nouveau, les élèves sont d'origine modeste et de races mélangées. Enfin, la quatrième et dernière étude a été réalisée dans 22 classes de 1^{re} année provenant d'un milieu urbain (Soar & Soar, 1973 et 1978). Les élèves sont d'origine mixte, tant au plan social qu'ethnique. La cinquième étude est de grande envergure, puisqu'elle inclut 289 classes primaires, essentiellement fréquentées par des enfants issus de milieu modeste (Soar, 1973). Les observations menées s'intègrent dans l'évaluation du programme *Follow Through* et concernent à la fois des classes imbriquées dans un projet expérimental et des classes de contrôle.

Dans la première étude, Soar utilise de façon combinée la grille de Flanders (FIAC) et une grille portant sur le comportement non-verbal et les expressions affectives du maître. Les quatre autres études comportent quatre grilles, deux codées directement sur le terrain et deux

codées à partir d'enregistrements. Le premier instrument tente d'appréhender l'organisation de la classe, ainsi que les rôles assignés et assumés par les élèves dans cette organisation ; il est également conçu pour enregistrer l'expression des affects par le professeur et les élèves. Le second système véhicule des concepts inspirés de la pensée de Dewey ; il porte sur la façon dont le maître développe le contenu d'enseignement en tenant compte, ou non, des apports des élèves. Le troisième instrument correspond à la taxonomie de Bloom et renvoie donc au niveau cognitif des échanges. Enfin, le quatrième instrument est la grille de Flanders, quelque peu aménagée.

Soar et Soar proposent plusieurs clarifications conceptuelles. Soucieux d'échapper à la confusion de concepts engendrée par la FIAC et par le sens commun, ces chercheurs s'attachent à distinguer le climat émotionnel (les affects positifs et négatifs, manifestés par les élèves et par l'enseignant) et le contrôle social exercé par le maître. À leurs yeux, on ne peut poser d'emblée qu'un enseignant qui exerce un contrôle important sur l'activité des élèves génère nécessairement un climat émotionnel négatif ; les deux paramètres doivent être appréhendés séparément de telle sorte qu'une estimation de l'un ne détermine pas *ipso facto* l'estimation de l'autre. En ce qui concerne la gestion de la classe, Soar (1966) note d'abord que l'enseignant peut l'assumer de façon tacite en établissant, dès le début de l'année, des règles et des routines (*classroom management, established structure*). De façon complémentaire, il régulera le fonctionnement

quotidien de la classe par l'intermédiaire de comportements récurrents (*current interaction*) ; dans le flux des interactions, il formulera des consignes, posera des questions et interviendra de façon fréquente et explicite pour rappeler les règles, les préciser, les nuancer ou en imposer le respect. Ces règles se déclinent en trois groupes, selon que le contrôle du maître porte sur :

- le comportement des élèves (déplacements et modes de participation aux échanges) ;
- les tâches d'apprentissage (sélection des tâches d'apprentissage et consignes portant sur la façon de les exécuter) ;
- les démarches de pensée (dans quelle mesure les élèves ont-ils l'autorisation ou sont-ils encouragés à aborder la matière à des niveaux cognitifs différents ? dans quelle mesure sont-ils encouragés à la pensée divergente ?).

Les résultats confortent le bien-fondé de ces distinctions. On les résumera en fonction de ces catégories.

Gestion des comportements des élèves

Dans les quatre études, on observe une relation positive entre le résultat au test et la fréquence des prises de parole des enseignants ainsi que leur contrôle du comportement des élèves. Plus précisément, il apparaît de façon très claire que les gains d'apprentissage dans les matières scolaires (langue maternelle et mathématiques) sont plus importants dans les classes où les maîtres éta-

blissent une structure qui limite les choix des élèves, leurs déplacements, ainsi que les occasions de perturbations.

Gestion des tâches d'apprentissage

Les enseignants qui ont une orientation clairement ciblée sur les apprentissages, qui exercent un contrôle strict sur les tâches d'apprentissage et surtout sur l'implication des élèves dans les tâches académiques, obtiennent de ces derniers un niveau de connaissances élevé. Pendant les travaux individuels sur feuille (*individual seatwork*), les maîtres *efficaces* circulent entre les bancs pour surveiller la concentration des élèves. Notons que l'analyse fine des résultats montre que l'on est souvent en présence de relations curvilinéaires, c'est-à-dire qu'au-delà d'un certain seuil, la corrélation s'estompe : autrement dit, au-delà d'un certain niveau de directivité, les gains d'apprentissage s'amenuisent ; la tendance peut même s'inverser. On reviendra sur ce point important ci-dessous.

Gestion des démarches de pensée

A ce niveau, les résultats se diversifient en fonction de l'origine socioculturelle des élèves et de leur niveau d'études. Les sollicitations de haut niveau cognitif sont contre-productives avec les jeunes élèves et spécialement ceux d'origine modeste ou de niveau intellectuel faible. En revanche, elles sont bénéfiques avec les élèves plus âgés et, spécialement, ceux caractérisés par un niveau intellectuel élevé.

Climat émotionnel

Une relation dite non-linéaire apparaît. Plus exactement un climat négatif est corrélé négativement avec les mesures de produit. En revanche, un climat affectif n'est corrélé ni positivement ni négativement avec ces mêmes mesures. Autrement dit, si un climat émotionnel négatif est contre-productif, il n'est pas établi que le climat doit être positif ; il peut être neutre.

Outre leur apport conceptuel, les travaux de Soar et Soar sont intéressants dans la mesure où ils illustrent deux aspects fondamentaux des processus d'enseignement :

- les effets d'interactions sont sinon règle générale, au moins phénomène régulier ;
- la relation entre une conduite et son effet peut être de nature curvilinéaire.

Illustrons ces deux affirmations en prenant appui sur la recherche menée par Soar et Soar (1973-1978). Celle-ci porte sur 59 classes primaires composées d'élèves d'origines ethniques diverses et issus de milieux sociaux défavorisés. Outre l'observation des interactions en classes, le plan de recherche prévoit la passation d'un test d'intelligence en début d'année et celle, en fin d'année, de trois tests : un test de créativité (A), le second de vocabulaire (B) et le troisième de lecture (C). Il importe de préciser que le test de vocabulaire fait abondamment appel au raisonnement abstrait, tandis que le test de

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

lecture requiert peu d'inférences et mobilise donc peu la capacité d'abstraction.

Les résultats des élèves aux tests de fin d'année ont été ajustés en fonction de leurs aptitudes initiales. Ces scores ajustés peuvent être considérés comme une mesure de progrès ou de gain (*pupil growth*). Soar et Soar ont ensuite calculé les corrélations de ces scores avec le taux d'influence indirecte, d'une part, et le taux de critiques de l'enseignant, d'autre part.

Le schéma 1 présente le type de relations observées grâce au premier lot de corrélations.



Figure n° 1

La liaison entre l'influence indirecte et la créativité (A) est linéaire : le plus vaut le mieux. Cette conclusion ne peut pas être étendue lorsque le critère d'efficacité est le vocabulaire (B) ou la lecture (C). Dans ces deux cas, des points optimaux (X et Z) sont repérables. Si on

abaisse les coordonnées des points X et Z à l'axe des abscisses, on remarque une différence entre les taux optimaux d'influence indirecte propres à chaque type de compétences : le point optimal X (activité qui demande peu de raisonnement abstrait) est situé en aval du point optimal Z (activité demandant plus de raisonnement abstrait).

L'analyse de ce schéma suggère qu'une tâche qui nécessite peu de raisonnement demande un taux d'indirectivité inférieur à celle nécessitant beaucoup de raisonnement. Il y a donc à la fois curvilinearité des relations entre les conduites d'enseignement et leurs effets et interaction entre la nature de la compétence mobilisée et l'efficacité des comportements des maîtres.

En ce qui concerne le taux de critiques (schéma 2), on observe que la courbe A (créativité) n'est plus linéaire, mais qu'elle passe par un point optimal O. En revanche, pour la courbe B (raisonnement abstrait), aucun point optimal n'est repérable. Si on abaisse la coordonnée du point O à l'axe des abscisses, on remarque que la créativité s'accommode d'un faible taux de critiques. La courbe C (peu de raisonnement abstrait) présente également un point optimal R qui, abaissé à l'axe des abscisses, correspond à un taux de critiques plus important.

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

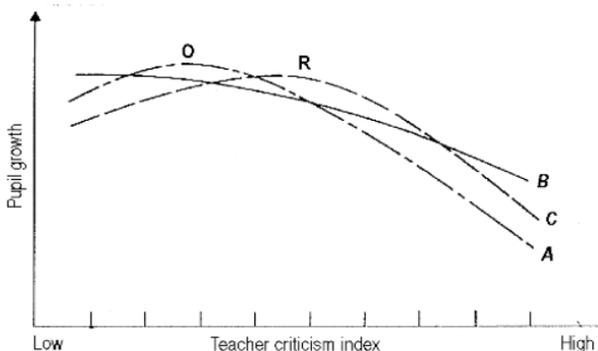


Figure n° 2

Il est également intéressant de comparer les courbes observées dans les deux schémas ou, plus simplement, de considérer simultanément les points optimaux identifiés dans l'un et l'autre cas :

- Seule la courbe C (test de lecture requérant peu de raisonnement abstrait) présente un point optimal dans les deux schémas : il semble, en définitive, que les élèves de l'échantillon sont plus performants pour ce type de tâche, quand le taux d'influence indirecte est bas et le niveau de critique est élevé.
- Les scores de créativité entretiennent une relation linéaire avec le taux d'influence indirecte et requièrent un faible taux de critique.
- Enfin, en ce qui concerne la tâche sollicitant le plus le raisonnement et les capacités d'abstraction des élèves,

MARCEL CRAHAY

un taux moyen d'influence indirecte et un taux minimal de critique constitueraient l'idéal.

Plus généralement, ces résultats indiquent que la nature de la relation entre un paramètre de processus et la mesure de produit varie selon la nature du critère d'efficacité envisagé. Nous reviendrons en conclusion sur cette importante observation.

CHAPITRE IV

À LA RECHERCHE DE L'ENSEIGNANT EFFICACE

De la stabilité de l'effet comme préalable à la recherche de l'enseignant efficace.

Les programmes de recherche de Brophy et Evertson, d'une part, et de Good et Grouws, d'autre part, marquent une autre avancée dans la réflexion méthodologique entourant les recherches sur les liaisons entre processus d'enseignement et produit d'apprentissage. Pour ces chercheurs, les études processus-produit postulent plus ou moins explicitement l'existence d'enseignants efficaces et d'autres qui le seraient moins. Le titre de leur ouvrage de 1975 *Teachers make a difference* en dit long à ce sujet. Or, logiquement, on ne peut parler d'enseignant efficace qu'à condition d'avoir pu montrer qu'un même praticien, opérant plusieurs d'années de suite à un même niveau de scolarité, engendre *grosso modo* des progrès d'importance équivalente et ceci, quel que soit le groupe d'élèves auquel il enseigne. C'est sur la base de ce raisonnement que Brophy (1973) d'une

part, Good et Grouws (1975) d'autre part, ont conduit des études de stabilité.

Utilisant des données recueillies grâce au *Metropolitan Achievement Test* (MAT), Brophy (1973) a pu étudier le profil d'efficacité de 88 enseignants de 2P et de 77 enseignants de 3P pendant trois années consécutives. Le MAT sert à la fois de mesure initiale (en début de chaque année) et finale (en fin de chaque année). Cette double passation d'une épreuve qui comporte cinq sous-tests (*word knowledge, word discrimination, reading, arithmetic computation and arithmetic concepts, problem solving*) permet d'appliquer l'analyse de régression aux données et, partant, de calculer un score résiduel (par élève) qui a valeur de mesure de gain. Pour chacune des classes de l'échantillon, on calcule la moyenne de ces scores ajustés ; on obtient ainsi un indicateur de l'efficacité des maîtres. Enfin, pour étudier la stabilité des enseignants, on procède au calcul des corrélations entre ces indicateurs d'efficacité au cours des 3 années successives. Ces corrélations sont généralement basses, mais toujours positives : la plupart sont de l'ordre de 0,30.

Tout aussi intéressant est l'examen des corrélations entre les sous-tests composant le MAT. Celles-ci sont nettement plus élevées que les coefficients de stabilité. Ainsi, lorsqu'on corrèle les scores au test de vocabulaire (*word knowledge*) de deux années successives, on obtient un coefficient de 0,30. En revanche, lorsqu'on corrèle pour une année particulière les scores à ce sous-test avec n'importe quel autre sous-test, les corrélations sont

supérieures à 0,70. D'une manière générale, il apparaît qu'un enseignant – à moins qu'il ne faille dire une classe – caractérisé(e) par un score moyen élevé dans un domaine, se caractérise également par des scores moyens élevés dans les autres domaines composant le test. Bref, quand il y a efficacité (ou non-efficacité), elle est générale. Il serait donc erroné de soutenir qu'au cours d'une même année, un enseignant peut être efficace dans un champ disciplinaire et inefficace dans un autre.

De même, il semble qu'il faille rejeter l'hypothèse selon laquelle il y aurait un effet du sexe de l'enseignant sur les résultats des filles et des garçons. En effet, Brophy (1973) constate d'abord qu'à âge égal, les filles réussissent mieux (en moyenne) que les garçons. Ensuite, il montre qu'il n'est pas possible d'identifier des enseignants chez qui les garçons auraient de meilleurs résultats et d'autres chez qui les filles seraient particulièrement avantagées. En définitive, quand il y a efficacité, celle-ci est globale : on n'observe pas d'effet différentiel des disciplines d'enseignement, ni du sexe des élèves.

Ces conclusions ont été confirmées plus tard par Ackland (1976) à un niveau supérieur de la scolarité (5P). Adoptant une méthode de recherche et de traitement statistique analogue à celle de Brophy, il observe des coefficients de stabilité quelque peu plus élevés. Pour les neuf sous-tests composant le test utilisé, il observe des valeurs supérieures à 0,40 lorsqu'on corrèle les scores obtenus dans 94 classes au cours de deux années successives.

Dans l'étude de stabilité de Good et Grouws (1975), les performances scolaires et les attitudes des élèves de 103 classes (3P et 4P) ont été mesurées en 1972-73 puis en 1973-74, en début et en fin d'année. Les classes proviennent de districts scolaires de type suburbain à la population de majorité blanche. Afin d'évaluer la stabilité des effets, Good et Grouws (1975) ont calculé des coefficients de corrélation à partir des mesures de gains ajustées aux sous-tests de l'*Iowa Test of Basic Skills*. Les résultats sont plutôt décevants puisque les coefficients de corrélation se situent, en moyenne, autour de 0,20 pour les performances scolaires et de 0,22 pour les mesures d'attitude. Par ailleurs, les mesures d'attitude et de rendement scolaire ne sont pas corrélées. D'une manière générale, les résultats relatifs aux attitudes sont difficiles à interpréter car les scores correspondant sont élevés dans quasiment toutes les classes ; ceci a pour conséquence de réduire leur variance, ce qui affecte en retour le calcul de corrélations.

En définitive, ces résultats sont troublants quant au postulat concernant les enseignants efficaces : les effets engendrés par les mêmes praticiens, opérant au même niveau scolaire pendant plusieurs années successives, ne sont pas corrélés de façon forte. Ceci oblige à admettre une variabilité relativement importante en fonction du groupe d'élèves. En définitive, on doit se demander s'il

ne conviendrait pas mieux de parler d'effet-classe⁹, plutôt que de parler d'effet-maître (Bressoux, 1994). Pour tester cette hypothèse concurrente, il serait intéressant de suivre des groupes-classe au cours de plusieurs années successives afin d'évaluer dans quelle mesure les performances moyennes des groupes d'élèves restent constantes à travers les niveaux de scolarité parcourus et ceci quel que soit le praticien responsable de leur enseignement.

Les recherches de Brophy et Everston

Brophy et Everston ont mené deux études processus-produit. La première – *The Texas teacher effectiveness study* – a trait à l'enseignement primaire, alors que la seconde – *The junior high study* – porte sur l'enseignement secondaire.

L'enseignant efficace au primaire

La *Texas Teacher Effectiveness Study* porte sur un sous-échantillon des maîtres impliqués dans l'étude de stabilité menée par Brophy (1973) : parmi les 165 maîtres observés pendant trois années consécutives, 31 enseignants de 2P et 28 de 3P ont été sélectionnés pour leur grande stabilité. Cet ensemble ne constitue ni un échantillon de maîtres efficaces, ni un échantillon contrasté : la distribution des scores moyens des classes au

⁹ L'observation par Brophy (1973) de fortes corrélations entre les sous-scores des tests supporte la plausibilité de cette hypothèse concurrente.

MAT épouse celle d'une courbe quasi normale. Les observations ont porté sur des leçons de mathématiques et de lecture. Enfin, les classes peuvent se ranger en deux catégories en fonction de l'origine sociale des élèves : il y a des classes composées majoritairement d'enfants de milieu modeste (*low SES*) et d'autres composées d'enfants de milieu aisé (*high SES*). Les 31 enseignants de 2P ont été observés pendant dix heures de classe et les 28 (de niveau 3P) ont été observés pendant 30 heures. Outre les mesures de processus, Brophy et Evertson (1976) ont pris en considération des variables de présage c'est-à-dire des mesures portant sur les antécédents et la personnalité des enseignants.

Les principaux résultats de cette étude renforcent, tout en les affinant, les conclusions des études précédentes.

Lorsqu'on croise l'analyse des variables de présage avec les indicateurs d'efficacité, on distingue trois groupes d'enseignants. Les enseignants des classes qui obtiennent les moins bons résultats d'apprentissage peuvent présenter deux profils différents. Le premier est typique d'enseignants désillusionnés. Ceux-ci se déclarent amers et expriment de nombreux griefs à l'égard de leurs élèves; ils soulignent la nécessité de se montrer autoritaires et parlent beaucoup de discipline. Le second profil correspond à des enseignants qui se montrent essentiellement préoccupés de la qualité de leurs relations avec leurs élèves ; ils donnent notamment priorité aux objectifs affectifs sur les objectifs cognitifs. Quant aux enseignants auxquels sont associés les effets d'ap-

prentissage les plus élevés, ils se définissent comme des personnes aimant travailler avec les élèves. Ils soulignent l'importance d'un enseignement ciblé sur les matières académiques et valorisent l'effort et la concentration sur le travail. Ils conçoivent leurs classes comme des environnements d'apprentissage et se décrivent comme des enseignants soucieux d'exploiter au mieux le temps qui leur est imparti. Au cours des interviews, les enseignants qualifiés d'*efficaces* manifestent un sens aigu de leurs responsabilités ; ils ont conscience du contrôle qu'ils exercent sur les activités d'enseignement et se perçoivent comme relativement efficaces. Lorsqu'on les questionne sur la planification de leur enseignement, ils manifestent une attitude pro-active : ils peuvent expliquer dans les grandes lignes comment ils vont organiser leur classe au cours des prochaines journées. Leurs réponses aux questions portant sur la contrôlabilité de la situation d'enseignement sont particulièrement éclairantes. Opposés à tout fatalisme, ils s'attribuent – au moins partiellement – les difficultés des élèves et, surtout, envisagent d'emblée des possibilités d'aménager leur enseignement pour aider ces élèves.

Quelles tendances observe-t-on au niveau des variables de processus ?

Les classes qui se caractérisent par les meilleurs scores aux mesures de produit se distinguent des autres par un taux particulièrement élevé d'implication des élèves dans les activités scolaires, ainsi que par un temps minimal consacré aux transitions ou aux problèmes de

procédure ou de discipline. Les enseignants de ces classes gèrent le groupe d'élèves et le temps scolaire en respectant les règles décrites par Kounin (1970) et conceptualisées par Evertson et Emmer (1982).

Kounin (1970) a inventé le terme de *withitness* pour définir ce style de *classroom management*. Le terme veut qualifier l'attitude d'implication constante et d'interventions appropriées des enseignants efficaces. Ceux-ci réussissent à impliquer l'ensemble de la classe dans les activités collectives et passent entre les bancs lors des travaux individuels, soucieux de la progression des tâches. Ils commettent rarement des erreurs de ciblage (*target errors*) ou des erreurs de temps de réaction (*timing errors*). On parle d'erreur de ciblage lorsqu'un enseignant blâme ou réprimande un élève qui n'a rien fait de mal et on parle d'erreur de timing lorsqu'un enseignant réagit tardivement à un problème de discipline ou à une erreur de type cognitif.

Soucieux de l'implication des élèves dans les activités scolaires, ces enseignants gèrent la classe de façon proactive. D'une part, ils explicitent très précisément leurs attentes quant aux règles à respecter en classe, et se montrent vigilants quant au respect de ces règles. D'autre part, ils réagissent (peut-être excessivement) au moindre incident de discipline et surtout préfèrent anticiper ces problèmes par l'un ou l'autre avertissement plutôt que de recourir *a posteriori* à des critiques ou à une punition. Ils sont également cohérents dans le temps

et justes dans leur façon de réagir aux infractions par rapport à ces règles.

Dans l'étude de Brophy et Evertson (1976), les enseignants efficaces se montrent moins soucieux d'asseoir leur autorité que d'engager les élèves dans les activités scolaires. Conscients de leurs responsabilités, ils préparent leur plan d'activités pour la journée et réussissent à passer d'une activité à l'autre de façon harmonieuse, c'est-à-dire avec un minimum de temps consacré aux transitions. Ceci a pour conséquence que les élèves attendent peu ou sont rarement inactifs. Les consignes relatives aux travaux individuels (*individual works assignments*) sont explicitées scrupuleusement. Elles sont calibrées sur les capacités des élèves, ce qui suppose un certain degré d'individualisation. Les élèves qui éprouvent des difficultés à accomplir une tâche reçoivent rapidement l'aide de l'enseignant ou d'un condisciple désigné (selon des règles établies au préalable ou en fonction d'une décision ponctuelle de l'enseignant). Aussi ces enseignants insistent sur la responsabilité des élèves quant à la qualité des travaux qu'ils remettent. Conséquents avec cette règle, ils supervisent et corrigent tous les travaux d'élèves, qu'il s'agisse de feuilles d'exercices réalisées en classe ou de devoirs à domicile.

Les tendances générales décrites ci-dessus prennent des accents différents selon l'origine sociale des élèves. Les élèves d'origine socio-économique élevée tendent à être plus enthousiastes que ceux issus de familles à statut modeste qui, eux, éprouvent davantage de difficultés et

se montrent plus souvent anxieux. Ces caractéristiques propres aux élèves entraînent des attitudes différentes chez les enseignants. Dans les classes peuplées en majorité d'élèves d'origine socio-économique favorisée, les enseignants les plus efficaces se montrent particulièrement stimulants sur le plan intellectuel et placent les élèves dans des situations-problèmes qui leur imposent de relever un défi. De surcroît, ils leur proposent des activités intellectuellement intéressantes dès qu'ils ont terminé leurs tâches scolaires. En revanche, dans les classes composées majoritairement d'enfants de milieux modestes, les enseignants les plus efficaces ont le souci de proposer des activités à la portée des élèves.

En matière de gestion des apprentissages, l'enseignant efficace se caractérise de manière générale par une capacité à organiser la progression des apprentissages de façon minutieuse et souple. Les leçons sont organisées selon un schéma classique : présentation du nouveau matériel et/ou rappel des connaissances acquises antérieurement ; vient ensuite un moment d'exercice ou de travail individuel durant lequel l'enseignant questionne les élèves, obtient des réponses le plus souvent correctes, qui sont suivies de *feedbacks*. Les questions sont abondantes, habituellement d'un bas niveau taxonomique, renvoyant à des faits plutôt qu'à des opinions. La progression se fait de façon rapide, mais à petits pas. Les élèves y expérimentent le succès, ce qui génère un effet « boule de neige ». Ce schéma général prend des accents quelque peu différents selon la composition des classes.

Dans les classes composées d'élèves favorisés, le rythme peut être plus rapide et la progression moins minutieuse. Dans celles peuplées d'enfants de milieux modestes, l'enseignant a intérêt à procéder par plus petites étapes en donnant davantage d'explications et en fournissant aux élèves l'opportunité de réaliser davantage d'exercices ; il donnera également plus de *feedbacks* et se caractérisera par un taux de redondance plus élevé que dans les autres classes.

La façon de réagir aux élèves diffère également selon la composition de la classe. Dans les classes fréquentées par des enfants de milieu favorisé, l'enseignant veille à contrôler les étudiants les plus participatifs et à impliquer ceux qui, plus hésitants, demandent qu'on les sollicite davantage. À l'occasion, il doit quelque peu freiner les plus actifs pour donner la parole aux autres. Toutefois, il ne semble pas fécond d'interroger sans cesse ces enfants, surtout lorsqu'ils ne connaissent pas la réponse. Dans les classes peuplées d'enfants d'origine modeste, beaucoup sont anxieux et manquent de confiance en leurs capacités intellectuelles, même quand ils connaissent les réponses aux questions posées. Il convient dès lors, pour l'enseignant, d'accepter toutes les réponses des élèves peu communicatifs et d'essayer de les améliorer par un jeu subtil de *feedbacks* correctifs. Bref, alors que dans les classes peuplées d'enfants favorisés, il convient de freiner l'enthousiasme de certains ; dans les classes à population défavorisée, il s'agit d'adopter l'attitude inverse : prendre garde à ne freiner aucun élève. À

l'occasion, ceci peut amener l'enseignant à accepter des réponses incorrectes ou incomplètes afin d'encourager l'élève dans sa participation, quitte à le corriger aimablement par la suite. Enfin, dans ces classes comme dans celles situées en milieu favorisé, les critiques sont rares et corrélées positivement avec le rendement. Rejoignant Stallings, Brophy et Evertson (1976) supposent que des critiques rares peuvent avoir un effet positif sur l'apprentissage de l'élève, surtout lorsqu'elles expriment indirectement une exigence positive à l'égard des élèves.

Brophy et Evertson (1976) notent encore l'absence de corrélation significative entre le rendement et une série de paramètres de processus dont la fonction consiste à motiver les élèves. Pour eux, ce constat est probablement tributaire du niveau d'études considéré. En effet, en début de scolarité, les élèves se caractérisent en grande majorité par une attitude extrêmement favorable vis-à-vis de l'école et l'on peut penser que les enseignants n'ont pas besoin de les motiver, ceux-ci s'engageant assez spontanément dans les apprentissages scolaires. En revanche, ce ne serait pas le cas des élèves plus âgés. La *Junior High Study*, menée au niveau secondaire permet d'envisager la pertinence de cette hypothèse.

L'enseignant efficace au secondaire

Evertson, Anderson, Anderson et Brophy (1980) présentent les résultats d'une étude – *The Junior High Study* – de type processus-produit menée auprès de 39 professeurs d'anglais (langue maternelle) et de 29 professeurs

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

de mathématiques. Chacun des enseignants retenus opère à deux niveaux scolaires : 7^e et 8^e *grade* (respectivement, 1S et 2S) ; le nombre de classes est donc double du nombre d'enseignants (respectivement, 78 classes d'anglais et 58 classes de mathématiques). Chaque enseignant a été observé une vingtaine de fois dans chacune de ses classes. Le nombre d'observations réalisées (2720 environ) est donc impressionnant.

Le niveau initial des élèves a été mesuré à l'aide du *Californian Achievement Test* (CAT) ; les sous-tests d'anglais et de mathématiques ont été présentés aux élèves au printemps précédant l'année de la recherche. Les post-tests ont été construits par les chercheurs après une analyse détaillée des opportunités éducatives offertes aux élèves, ceci afin d'ajuster la mesure des produits au curriculum enseigné. Tâche aisée pour les mathématiques ; en revanche, Evertson et al. (1980) ont éprouvé de grandes difficultés à construire un test solide pour apprécier les apprentissages en anglais. L'analyse des thèmes abordés dans les classes révèle une grande diversité d'objets d'enseignement : certains enseignants se concentrent sur la grammaire et les compétences de base, d'autres sur la compréhension en lecture et l'expression écrite, d'autres encore sur la poésie et le théâtre. En définitive, le post-test de mathématiques présente un taux de recouvrement (*content coverage*) élevé avec le curriculum enseigné dans les classes ; le taux de recouvrement est, en revanche, moins bon pour la langue maternelle.

À la fin de l'année d'expérience, les élèves ont été invités à évaluer leur enseignant sur une échelle de type Lickert. Plus précisément, il leur a été demandé d'une part d'indiquer dans quelle mesure ils trouvaient leur enseignant agréable, disponible et, d'autre part, s'ils avaient le sentiment d'avoir bénéficié de l'enseignement de ce professeur. Une analyse factorielle appliquée à leurs réponses a abouti à un premier facteur extrêmement solide, qui a été utilisé comme mesure de l'attitude des élèves à l'égard de cet enseignant.

Les premiers résultats sont fournis par les corrélations entre les différents tests. Les scores aux prétests rendent compte de 71 % de la variance des scores aux post-tests de mathématiques et de 85 % de cette variance aux post-tests d'anglais. En conséquence, il reste une part de variance explicable par les processus d'enseignement assez fluctuante : 29 % pour les mathématiques et 15 % pour l'anglais. Quant aux scores d'attitude, ils corrélaient positivement avec les mesures de gains (scores résiduels) en mathématiques (+0,32), mais négativement avec les gains en anglais (-0,24) ; nous y reviendrons.

L'étude porte sur des enseignants fonctionnant à deux niveaux de la scolarité, il est possible de calculer une corrélation entre les gains engendrés par chaque enseignant dans deux classes différentes. On retrouve ici le souci de Brophy d'estimer la stabilité des effets induits par l'action des enseignants. En mathématiques, cette corrélation est de +0,37 pour les mesures de gain ; elle est de +0,44 pour les mesures d'attitude. Par ailleurs, il

s'avère que cinq enseignants présentent des effets très différents selon les degrés ; deux d'entre eux se trouvent face à des classes dont le niveau initial est extrêmement différent (40 points de différence aux prétests). Lorsqu'on retire ces cinq enseignants, la corrélation avec le rendement monte à +0,57 et a la même valeur pour les attitudes. Bref, la stabilité des enseignants de mathématiques est relativement bonne. Ce n'est pas le cas pour les professeurs d'anglais. Pour les mesures de gain, le coefficient de stabilité est seulement de +0,05. En revanche, la corrélation entre les scores d'attitude des deux années est de +0,82¹⁰ (Emmer, Evertson et Brophy, 1979). Bref, les effets des professeurs d'anglais sur l'apprentissage des élèves ne sont pas stables. De surcroît, ils sont corrélés négativement (-0,24, cf. ci-dessus) avec les attitudes pour lesquelles on observe une stabilité certaine.

Des relations processus-produit émergent de l'analyse des résultats (Evertson *et al.*, 1980). Notamment, le rendement en anglais est corrélé positivement à l'absence de problèmes disciplinaires en classe et à la fréquence des encouragements et félicitations formulés par l'enseignant. Eu égard aux mesures d'attitude, il apparaît que les enseignants exigeants sont assez peu appréciés des élèves. Plus précisément, les élèves expriment des

¹⁰ Ces coefficients montent respectivement à + 0,29 et 0,83 lorsqu'on enlève treize enseignants dont les classes sont fortement contrastées quant au niveau initial.

attitudes positives à l'égard de l'enseignant et de l'école lorsque l'attitude de l'enseignant a été qualifiée par les observateurs de chaleureuse, enthousiaste et orientée vers les élèves. En revanche, les élèves expriment des avis moins positifs lorsque les observateurs ont noté que les enseignants consacrent un temps important à exposer la matière, font preuve d'exigence, posent des questions difficiles, se montrent à l'occasion critiques et s'efforcent d'améliorer les réponses erronées des élèves. Bref, l'évaluation que les élèves portent sur leur enseignant est nettement positive lorsque ce dernier est perçu comme « gentil » et qu'il ne se montre pas trop exigeant. Or, ces paramètres sont peu corrélés (voire corrélés négativement) avec le rendement. On se trouve donc face à une certaine incompatibilité entre les conduites favorables à l'apprentissage et celles propices à un bon climat de classe.

Cette opposition entre les variables d'enseignement associées au rendement et celles reliées au climat de classe ne se retrouve pas en mathématiques. Pour cette discipline, le maître efficace est également apprécié par les élèves. Avant tout, les bons enseignants de mathématiques sont des gestionnaires de classe efficaces. Même s'ils ont autant de problèmes de discipline à affronter, ils le font de façon plus avisée. Plus précisément, ils étouffent les problèmes dans l'oeuf en réagissant très rapidement dès qu'un incident s'annonce. Manifestant leur présence en classe par des déplacements fréquents parmi les élèves, ils font preuve de *withitness*.

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

Les mesures de qualité et de quantité d'enseignement sont encore plus directement reliées au rendement dans cette recherche que dans la précédente. Les enseignants efficaces interviennent directement dans le processus d'apprentissage : ils expliquent, démontrent, discutent et donnent des exercices. Le temps dévolu au travail individuel est réduit ; pendant celui-ci, les interventions de l'enseignant sont fréquentes et visent à assurer une guidance rapprochée de la progression des élèves dans les apprentissages.

Concernant les questions, la différence entre enseignants efficaces et enseignants moins efficaces est surtout quantitative. Les enseignants efficaces posent beaucoup de questions. La plupart de celles-ci sont orientées sur les produits plutôt que sur les processus. Toutefois, et comparativement avec la recherche concernant l'enseignement primaire, le pourcentage de questions orientées vers les processus est plus élevé ; il est, en outre, corrélé positivement (quoique faiblement) avec les résultats. En moyenne, 24 questions par période de 50 minutes sont posées dans les classes caractérisées par un gain d'apprentissage élevé ; 25 % de ces questions sont orientées sur les processus. Dans les classes caractérisées par un faible rendement, seulement 8,5 questions sont enregistrées par tranche de 50 minutes ; 15 % seulement de ces questions sont orientées sur les processus.

De manière générale, on trouve peu de travail en petits groupes. La plupart des leçons sont organisées en mode frontal. La gestion de la participation des élèves

dans ce contexte se fait selon un schéma classique : 45 % des réponses d'élèves sont obtenues par des sollicitations adressées à des élèves non volontaires, 25 % des réponses sont le fruit de sollicitations adressées à des volontaires qui lèvent la main et 25 % des réponses proviennent d'élèves qui appellent le maître (*calling out*). La fréquence des sollicitations adressées aux volontaires est corrélée positivement avec le rendement, ce qui ne veut pas dire que la fréquence des sollicitations adressées aux non volontaires soit corrélée négativement avec le gain. Dans les classes composées d'élèves de milieu modeste, les enseignants ont avantage à encourager les prises de parole spontanées, même intempestives ; cette stratégie favorise à la fois le rythme de travail – soutenu – et la participation des étudiants. Ainsi, le fait d'accepter des questions spontanées dans ce type de classe est corrélé positivement avec les gains. Les remarques positives formulées en grand groupe sont rares, mais corrélées positivement (quoique faiblement) avec les gains. En revanche, les félicitations et les critiques exprimées au cours d'interactions dyadiques (notamment pendant les travaux individuels) ne sont pas liées au rendement. Ce dernier résultat contraste avec ce qui est observé dans l'enseignement primaire où il est important d'interagir avec les élèves en situation de petits groupes ou en situation individualisée. Au secondaire, il semble plus important de rester en contact avec la classe entière et d'avancer à un rythme soutenu.

Utilisant les réponses des observateurs à des échelles d'évaluation des enseignants à haute inférence, Evertson et al. (1980) mettent en évidence que ceux-ci voient le professeur efficace comme un bon gestionnaire : il maintient l'ordre, gère la classe et veille au respect des règles. Les transitions sont peu fréquentes, courtes ; il y a peu d'interruptions. Les enseignants ont l'air compétents, confiants, crédibles, enthousiastes et réceptifs face à leurs élèves. Ils sont clairs dans leur présentation. Notamment, ils se montrent concernés par la réussite de leurs élèves ainsi que par leurs notes scolaires. Ils mettent l'accent sur l'apprentissage scolaire et gèrent le temps de façon stricte et efficace. En définitive, on trouve une description de l'enseignant efficace proche de celle issue des études de Stallings. Trois variables semblent particulièrement importantes, tant au niveau des petits degrés pour l'apprentissage de la lecture (Stallings & Kaskowitz, 1974) qu'au niveau supérieur pour l'enseignement des mathématiques (Evertson *et al.*, 1980) : 1. les opportunités d'apprentissage, 2. la gestion efficace du temps scolaire et 3. une méthode d'enseignement direct appelée par Good (1979) *instruction active*.

Le programme de recherche de Good et Grouws

Good et Grouws ont conduit un programme de recherche cohérent, décliné en trois étapes : à l'étude de stabilité discutée ci-dessus fait suite une recherche processus-produit, suivie d'une recherche quasi-

expérimentale avec intervention ; ces deux étapes de recherche concernent des classes de 4P.

L'étude corrélacionnelle

Cette recherche se situe dans le prolongement de l'étude de stabilité. Elle devait porter sur des enseignants caractérisés par des effets stables eu égard aux attitudes et aux apprentissages des élèves dans différentes matières scolaires. Bref, la recherche corrélacionnelle avait été conçue sur le postulat suivant : il est possible de trouver des enseignants globalement efficaces (indépendamment de la composition des classes et de la discipline scolaire) et, de surcroît, appréciés par des élèves. L'étude de stabilité ne confirmant pas ce postulat, Good et Grouws (1977) décident de faire porter la recherche processus-produit exclusivement sur les mathématiques ; les coefficients de stabilité observés dans ce domaine sont, en effet, plus stables que ceux calculés pour les autres disciplines scolaires. Dix-huit enseignants (hors des 103 ayant participé à l'étude de stabilité) sont ainsi retenus : neuf de ceux-ci peuvent être qualifiés d'*efficaces* eu égard au rendement moyen de leur classe, tandis que les neuf autres peuvent être désignés par l'étiquette *peu efficaces*.

Les classes de ces 18 enseignants ont été observées sept fois pendant des leçons de mathématiques. Le rendement des élèves a été mesuré à l'aide de l'*Iowa Test of Basic Skills* selon un schéma prétest/post-test (en automne, puis au printemps de la même année). Afin de

préservé l'anonymat de ces 18 enseignants, les mêmes données processus-produit ont été recueillies dans un lot complémentaire de 23 classes de 4^e année. En définitive, Good et Grouws (1977) disposent de données relatives à 41 classes ; celles-ci se prêtent au calcul de corrélations du type processus-produit. Parallèlement, les chercheurs ont procédé à une comparaison systématique des neuf classes *efficaces* avec les neuf classes dites *peu efficaces*.

Les résultats vont dans le même sens que les études précédentes. En effet, les enseignants *efficaces* se distinguent des autres par leur mode de gestion de classe. Ils passent peu de temps à « faire la discipline » et la durée des transitions entre les activités éducatives est réduite au minimum. Leurs élèves posent fréquemment des questions et répondent très spontanément à celles qui leur sont adressées. Bien que les enseignants de ces classes se caractérisent par l'importance qu'ils accordent aux apprentissages scolaires, le climat de classe est excellent : ceci est attesté par les réponses des élèves aux échelles d'attitude. Le rythme de progression dans la matière est élevé : dans ces classes, on étudie en moyenne 1,13 pages du manuel de mathématiques par jour, tandis que, dans les autres classes, le même indicateur donne la valeur de 0,71 (Good, Grouws & Beckerman, 1978). Cet indicateur de progression (nombre de pages moyen étudié par jour) est corrélé avec le rendement en mathématiques (0,49).

Les maîtres *efficaces* se distinguent également des autres maîtres par le petit nombre de questions qu'ils

adressent à leurs élèves et par leur adéquation : peu de questions débouchent sur des réponses incorrectes. De plus, lorsque les élèves répondent de façon erronée aux questions, les enseignants *efficaces* réagissent deux fois plus souvent que les autres enseignants par des *feedbacks* de processus : ils expliquent la procédure conduisant à la bonne réponse ; leurs collègues *moins efficaces* ont plutôt tendance à donner la bonne réponse.

Par ailleurs, des observations menées de façon plus qualitative révèlent que les enseignants *efficaces* se caractérisent par la clarté de leurs explications. Leur rythme de progression dans la matière est plus rapide et le temps consacré à revenir sur des consignes formulées précédemment est faible. A l'opposé, les enseignants *peu efficaces* fournissent des explications imprécises et sont obligés de revenir régulièrement sur les consignes données en début de leçon.

L'ensemble des observations dresse des enseignants *efficaces* un portrait cohérent, résumé dans le tableau 6 ci-après. Le rythme de progression dans la matière est, chez eux, particulièrement rapide et ce, pour plusieurs raisons. D'abord, ils fournissent des explications et des consignes d'emblée claires. Ensuite, ils captent la balle au bond pour intégrer des explications aux réponses fournies aux questions des élèves plutôt que de consacrer de longs moments aux explications. Aussi, la plupart de leurs questions sont de type factuel, susceptibles d'engendrer immédiatement des réponses correctes. Enfin, lorsque leurs élèves se retrouvent dans l'embarras,

ils se montrent prompts à donner des explications complémentaires plutôt que de simplement fournir la (ou les) réponse(s) correcte(s). Pendant les travaux individuels, les enseignants *efficaces* circulent parmi les bancs et se montrent prêts à encourager les élèves et à les aider en cas de difficulté. Dans ce contexte, ils répondent essentiellement aux sollicitations de leurs élèves : la moyenne des interactions suscitées par les enseignants est de 3 par heure alors qu'elle est de 6 dans les classes des enseignants *peu efficaces* ; celle des interactions sollicitées par les élèves est de 23 dans les classes *efficaces* alors qu'elle est de 12 dans les autres classes. En outre, il apparaît à nouveau que, pendant les travaux individuels, les *feedbacks* des enseignants *efficaces* consistent prioritairement en explications alors que ceux des enseignants *peu efficaces* consistent à donner directement la bonne réponse.

Good et Grouws (1977) décrivent les *feedbacks* des enseignants *efficaces* comme étant immédiats, non évaluatifs et appropriés : ils réagissent promptement aux réponses des élèves et se montrent essentiellement préoccupés par la qualité de la réponse des élèves. En conséquence, les maîtres *efficaces* adressent moins de louanges et moins de critiques à leurs élèves que les enseignants *peu efficaces*. Par ailleurs, lorsqu'ils expriment un jugement concernant la prestation d'un élève, celui-ci porte sur un aspect précis de la réponse de l'élève alors que les enseignants *moins efficaces* ont

tendance à adresser des louanges ou des critiques générales.

Tableau 6 : Portrait-robot des enseignants efficaces et peu efficaces selon Good et Grouws (1977).

<i>Interventions</i>	<i>Enseignants " efficaces "</i>	<i>Enseignants " peu efficaces "</i>
Temps passé à " faire " la discipline	Peu important	Important
Durée des transitions entre activités	Faible	Important
Questions posées par les élèves de l'enseignant	Fréquentes	Peu fréquentes
Réponses spontanées aux questions du maître	Fréquentes	Peu fréquentes
Climat	Bon	Bon
Rythme de progression dans la matière	Rapide	Lent
Nombre de questions posées par l'enseignant qui débouchent sur des réponses erronées	Faible	Elevé
Feedbacks de processus consécutifs à une erreur d'un élève	Fréquents	Rares
Explications	Claires	Imprécise
Questions factuelles	Fréquentes	Peu fréquentes

En définitive, l'élément distinctif le plus net des réactions des enseignants *efficaces* par rapport aux enseignants *moins efficaces* concerne le caractère de contingence de ces conduites d'enseignement.

L'intervention sous forme d'étude quasi-expérimentale

Dans la dernière phase de leur programme de recherche, Good et Grouws (1979) ont conduit une recherche de type quasi-expérimental contrastant un groupe de contrôle avec des classes animées par des enseignants bénéficiant d'un programme de formation. Les classes sont issues d'écoles urbaines fréquentées en majorité par des élèves de niveau socio-économique bas. Le programme de formation est défini par une série de principes éducatifs organisés en un modèle inspiré des résultats de l'étude corrélationnelle décrite ci-dessus.

Le modèle conçu par Good et Grouws (1979) comporte cinq phases :

1. **Rappel quotidien** (les 8 premières minutes de la leçon, excepté le lundi)
 - a) Revoir les concepts et les savoir-faire associés au devoir à domicile qui a été assigné aux élèves le jour précédent.
 - b) Examiner les devoirs à domicile des élèves et éventuellement procéder à des corrections spécifiques.
 - c) Proposer quelques exercices de calcul mental.

2. **Développement** (environ 20 minutes)
 - a) Bref rappel des concepts et procédures prérequis.
 - b) Présentation du nouveau concept ou de la nouvelle procédure en mettant l'accent sur la compréhension des élèves ; ceci peut impliquer le recours à des expli-

cations, à des démonstrations concrètes, à des illustrations, etc.

c) Jauger la compréhension des étudiants :

1. Susciter un processus interactif avec les élèves ;
2. Proposer des exercices mettant en oeuvre le concept ou la procédure expliquée et contrôler la réalisation de ces exercices.

d) Si nécessaire réexpliquer les concepts ou les procédures mal compris.

3. **Travail individuel** (environ 15 minutes)

a) Fournir aux élèves un exercice à réaliser dans un contexte préservé au maximum de perturbations.

b) Maintenir l'attention des élèves en circulant à travers les bancs.

c) Avertir les étudiants que leur travail sera évalué à la fin de l'exercice.

d) Evaluer le travail des élèves.

4. **Explication des devoirs à domicile**

a) Les devoirs doivent être assignés selon une base régulière (à la fin de chaque leçon de mathématiques, hormis le vendredi).

b) Ils doivent correspondre à approximativement 15 minutes de travail.

c) Ils doivent inclure un ou deux problèmes suscitant la révision des concepts ou procédures vus au cours.

5. Révision spécifique

a. Révision hebdomadaire :

Organiser une révision tous les lundis pendant 25 minutes au début du premier cours de mathématiques.

Cette révision doit être très précisément orientée vers des concepts et des procédures.

b. Révision mensuelle :

1. A organiser chaque lundi du mois.

2. Se concentrer sur les concepts et procédures étudiés dans le courant du mois qui précède.

Un manuel expliquant le modèle est distribué aux 21 maîtres du groupe expérimental. Il est présenté aux enseignants, et discuté, au cours de deux séances de 90 minutes chacune. Les chercheurs ont également rencontré les 19 maîtres du groupe contrôle afin de leur expliquer l'importance de l'enseignement des mathématiques ; aucune recommandation spécifique ne leur est donnée. Cette procédure a pour objectif de réduire les effets Hawthorne associés habituellement au fait de participer à une expérience. L'objectif des chercheurs est, en fait, de donner aux enseignants le sentiment de participer activement à l'expérience.

Entre octobre 1977 et fin janvier 1978, chaque enseignant membre du groupe expérimental et du groupe de contrôle a été observé six fois pendant qu'il enseignait les mathématiques. Dix-neuf des 21 enseignants du groupe expérimental ont implanté très scrupuleusement le programme expérimental. Au total, on se trouve bien

en face de deux groupes de classes différant quant à leur fonctionnement effectif. Toutefois, il faut noter que, contrairement à ce qui était recherché, le temps dévolu à la phase de développement (cf. ci-dessus) n'est guère plus important dans le groupe expérimental que dans les classes du groupe de contrôle.

Les résultats attestent de l'efficacité générale de l'intervention. D'une part, les élèves du groupe de contrôle progressent de façon sensible entre le pré et le post-test. Les progrès observés dans ces classes excèdent largement ceux enregistrés par les mêmes enseignants au cours des années précédentes. Ceci suggère la présence effective d'un effet Hawthorne. Celui-ci n'occulte, toutefois, pas l'effet du modèle d'enseignement des mathématiques mis à l'épreuve. En effet, les gains du groupe expérimental sont significativement supérieurs à ceux du groupe de contrôle pour les deux tests retenus : le SRA (un test standardisé de mathématiques) et un test ciblé spécifiquement sur le contenu enseigné pendant la période d'observation. Les progrès sont substantiels : en quelques mois, les élèves du groupe expérimental progressent en moyenne du 27^e percentile au 58^e, si l'on se réfère aux normes nationales. Les scores d'attitude des élèves des classes expérimentales sont également supérieurs à ceux enregistrés dans le groupe de contrôle. Enfin, il apparaît que les enseignants qui ont le mieux respecté les recommandations du modèle se caractérisent par les résultats les meilleurs.

CHAPITRE V

DEUX PISTES D'APPROFONDISSEMENT DE LA MÉTHODOLOGIE PROCESSUS-PRODUIT

Au cœur du *Descriptive-Correlational-Experimental Loop*, le calcul de corrélations entre variables de processus et mesures de produit joue un rôle-clé : les corrélations positives déterminent les recommandations adressées aux enseignants des classes expérimentales. En amont de toute la démarche de recherche, se situe le moment où le chercheur choisit les variables de processus à observer et autour desquelles il va construire sa (ou ses) grille(s) d'observation. En définitive, l'investigation a pour objet de confirmer ou d'infirmer les hypothèses ou intuitions préalables des chercheurs. Impossible dans le cadre de ce schéma de repérer d'autres variables déterminantes des apprentissages des élèves que celles intégrées dans les grilles d'observation. Pour échapper à cet écueil, les chercheurs de la *Beginning Teacher Evaluation Study* (BTES) se sont donnés le projet d'articuler approches quantitative et qualitative afin de voir si la

seconde approche confirmait les tendances mises en évidence par la première. Par ailleurs, la BTES accorde une attention toute particulière aux activités des élèves en classe.

On peut également reprocher à la méthodologie processus-produit canonique de ne pas être en mesure d'échapper aux illusions de la colinéarité. Ainsi, que ce soit dans la *Training experiment* de Stallings, Needle et Sayrook (1979) ou dans l'étude d'intervention de Good et Grouws (1979), les effets observés sont liés à un ensemble de variables de processus, plus ou moins bien implantées, dont il n'est pas sûr que toutes soient déterminantes. C'est afin d'échapper à cet autre écueil que des chercheurs de l'Université de Canterbury se sont attachés à mener des études quasi-expérimentales ciblées sur des variables de processus spécifiques.

Dans ce chapitre, nous examinerons, l'un à la suite de l'autre, ces deux programmes de recherche.

La *Beginning Teacher Evaluation Study (BTES)* ou l'articulation des approches quantitatives et qualitatives

En 1970, l'Etat de Californie, soucieux de réformer fondamentalement son système de formation des enseignants ainsi que leurs modalités de certification, établit une commission. Celle-ci dépose ses conclusions en 1972 en recommandant avec force de planifier une étude dont l'objectif est l'identification des compétences d'enseignement qui pourraient constituer un référentiel

pour évaluer les enseignants novices. Au fur et à mesure qu'elle progresse dans la conceptualisation du plan de recherche, la commission débouche sur la conviction qu'une recherche portant sur les relations entre conduites d'enseignement et apprentissages des élèves constitue une nécessité. Ainsi démarre la BTES, subsidiée par le *National Institute of Education* et impliquant la participation des chercheurs de l'*Education Testing Service* et du *Far West Regional Laboratory for Educational Research and Development* (Powell, 1980). Il faut souligner que le sigle désignant la recherche s'est révélé en définitive inapproprié : l'étude porte moins sur des enseignants débutants que sur des enseignants expérimentés ; en outre elle s'est nettement écartée des problèmes d'évaluation des enseignants pour se tourner vers la mise en évidence de relations processus-produit.

Achevée en 1977, la BTES correspond à un vaste programme de recherche qui s'est étalé sur sept années. Elle débute par une première phase (1970-1972) dont nous venons de rendre compte ; c'est la phase de réflexion sur la certification des enseignants qui débouche sur la conception de l'étude elle-même.

La seconde phase est désignée par l'expression *First Field Study*. Elle porte sur 41 enseignants de 2P et 54 de 5P, qui ont tous au moins trois ans d'expérience. Chacune des 95 classes est observée quatre fois. À deux occasions, les observateurs se centrent sur huit étudiants cibles, sélectionnés en fonction de leur sexe et du niveau de leurs performances scolaires lors de l'année anté-

rieure. La grille d'observation, appelée *APPLE*, est ciblée sur la nature et la durée des activités dans lesquelles s'impliquent les élèves. Pendant les deux autres journées d'observation, l'objectif est d'analyser la nature de l'enseignement dispensé lors d'activités de lecture et de mathématiques. Pour ce faire, les observateurs utilisent un instrument appelé *RAMOS*. Les résultats de cette étude se présentent sous la forme de mises en relation entre des variables processus et le rendement des élèves (Mc Donald & Elias, 1976 ; Donald, 1976, 1977).

La troisième phase se divise en deux études. La première recourt à la méthode ethnographique. Quarante enseignants sont concernés : 20 de 2P et 20 de 5P. Dans chacun de ces échantillons, dix enseignants sont considérés comme efficaces et les dix autres comme moins efficaces. La sélection des enseignants efficaces/moins efficaces est opérée au cours de deux semaines d'observation portant sur l'unité d'enseignement spécifique de la lecture. On parlera, dans ce cas-ci, de *known sample* dans la mesure où les caractéristiques des enseignants sont connues a priori par les chercheurs. En revanche, les observateurs réalisent leur visite dans les classes en aveugle : ils ne savent pas s'ils sont face à un enseignant efficace ou non. Quant à ces derniers, ils ignorent l'existence même de la distinction. L'essentiel des résultats de cette étude ethnographique sont présentés par Berliner et Tikunoff (1976, 1978).

Enfin, entre 1976 et 1977, une seconde *Field Study* est menée dans 50 classes : 25 classes de 2P et 25 de

niveau 5P. Les classes sont sélectionnées sur la base des résultats des élèves au cours des années antérieures. Plus précisément, les classes retenues comportent toutes au moins six élèves dont les scores d'entrée en mathématiques et en lecture se situent entre le trentième et le soixantième percentile. En définitive, les chercheurs de la BTES aboutissent à un échantillon diversifié du point de vue racial et ethnique, quoique légèrement surreprésenté par le milieu inférieur de la distribution des niveaux socio-économiques. Des mesures de produit sont opérées en octobre, décembre et mai. Toutes les classes ont été observées chaque semaine pendant une journée entière et ce sur une période de vingt semaines. Pour chacun des six élèves cibles de chaque classe, on procède toutes les quatre minutes à trois notations : 1) le contenu disciplinaire avec lequel il est en contact, 2) son degré d'attention ou d'implication dans la tâche (haut, modéré ou bas) et 3) le degré d'exactitude des réponses qu'il fournit. Lorsque l'enseignant s'adresse à l'élève observé, ces comportements sont encodés selon un système à trois composantes : présentation d'informations, *monitoring* et *feedbacks*. La fonction de présentation par l'enseignant se divise, elle, en quatre sous-catégories : 1) explication planifiée, 2) explication non planifiée, 3) consigne, 4) conseil ou guidance relatif à la tâche. La notion de *monitoring* renvoie aux comportements par lesquels l'enseignant observe l'élève, mais aussi le questionne. Quant au *feedbacks*, il concerne aussi bien les réponses scolaires des élèves que le contrôle de leur

attention ou de leur implication dans la tâche. En plus de ces observations, les enseignants sont interviewés, de façon approfondie, en automne et au printemps. Leurs impressions sont également recueillies, brièvement, chaque semaine. Ils sont également tenus de remplir un journal de bord. Toutes ces données sont recueillies afin d'apprécier la façon dont les enseignants planifient leur enseignement, mais aussi pour appréhender dans quelle mesure ils diagnostiquent les difficultés des élèves et de quelle façon ils s'efforcent d'y réagir. Les résultats de cette dernière phase de la BTES sont présentés dans divers rapports techniques (Berliner, Fisher, Filby & Marliave, 1978 ; Fisher *et al.*, 1978) ainsi que dans un chapitre rédigé par Fisher *et al.*, 1980).

Les résultats de la *First Field Study* ne sont pas présentés de façon classique. Plutôt que de présenter leurs observations sous la forme de corrélations reliant chaque variable de processus avec le rendement, Mc Donald et Elias (1976) ont eu recours à des calculs de corrélation et de régression multiple. Ceci signifie que l'on n'a de résultats que pour les variables de processus qui ajoutent une part d'explication significative à la variance des gains. Les mesures de produit correspondent, en effet, à des scores ajustés en fonction de l'origine sociale des élèves et de leurs performances scolaires au cours de l'année antérieure. Ceci rend évidemment quelque peu difficile la comparaison avec les résultats obtenus par les études précédentes.

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

Les différentes phases de cette étude sont résumées dans le tableau 7 :

Tableau 7 : Description synthétique des différentes phases de la BTES

Phase I 1970-72	Réflexion sur la formation et la certification des enseignants avec l'objectif de définir un référentiel de compétences nécessaires à l'enseignant débutant.									
Phase II 1973-74	<p><i>First Field Study</i> ciblée sur des enseignants ayant au moins 3 ans d'expérience</p> <ul style="list-style-type: none"> - 41 classes de 2P - 54 classes de 5P <p>Données recueillies sur leurs connaissances disciplinaires, etc.</p> <p>Observations en classes</p> <ul style="list-style-type: none"> - APPLE → durée des activités dans lesquelles se sont impliqués 8 élèves cibles, sélectionnés en fonction du sexe et de leur niveau de performance antérieur - RAMOS → observation des conduites d'enseignement <p>Chaque classe est observée 2 fois avec chacun des systèmes</p>									
Phase III-A 1974-75	<p><i>Ethnographic Study (know sample)</i></p> <table style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th style="text-align: center;"><i>Enseignants efficaces</i></th> <th style="text-align: center;"><i>Moins efficaces</i></th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">2P</td> <td style="text-align: center;">10</td> <td style="text-align: center;">10</td> </tr> <tr> <td style="text-align: center;">5P</td> <td style="text-align: center;">10</td> <td style="text-align: center;">10</td> </tr> </tbody> </table> <p>Efficacité mesurée au terme de deux semaines d'observation</p>		<i>Enseignants efficaces</i>	<i>Moins efficaces</i>	2P	10	10	5P	10	10
	<i>Enseignants efficaces</i>	<i>Moins efficaces</i>								
2P	10	10								
5P	10	10								

MARCEL CRAHAY

Phase III-B 1976-77	<i>Second Field Study</i> 25 classes de 2 P 21 classes de 5 P - Elles comportent au moins 6 élèves dont les performances initiales en mathématique et en lecture se situent entre la 30 ^e et le 60 ^e percentile de l'étalonnage national. - Mesures de produit : octobre, décembre, mai - Observation centrée sur les élèves + interview régulier des enseignants
---------------------------	--

Enfin, les chercheurs de la BTES ont privilégié les scores combinant différentes variables de processus plutôt que d'analyser chaque variable d'enseignement en particulier.

Malgré ces particularités méthodologiques, on retrouve les mêmes tendances que chez Brophy et Evertson et chez Good et Grouws. Les progrès scolaires les plus importants ou, plus justement, les gains de rendement les plus élevés sont observés dans les classes où les enseignants sont bien organisés, maximisent le temps alloué à l'enseignement des matières principales et réduisent autant que faire se peut le temps dévolu à la préparation des activités, à la discipline ou à d'autres tâches connexes. Ces enseignants *efficaces* consacrent beaucoup de temps à l'enseignement en grands groupes et, lorsqu'ils organisent des travaux individuels, ils circulent parmi les bancs pour assurer un *monitoring* des élèves. Dans ces classes, les élèves sont concentrés sur la tâche la plupart du temps, qu'il s'agisse d'activités collectives ou de travaux individuels. En 2P, le temps passé à l'exercice de *skills* scolaires (p. ex. calcul en mathématiques ou segmentation de mots en lecture) est lié posi-

vement avec le rendement en seconde année. En revanche, en 5P, ce même paramètre est associé négativement au rendement, mais le temps consacré, à ce niveau scolaire, à l'application de ces *skills* (résolution de problèmes mathématiques et compréhension de lecture) est, lui, lié positivement au rendement. Les *feedbacks* positifs et les encouragements sont liés positivement au rendement et ce, en 2P comme en 5P. Quant à la diversité du matériel, cet aspect entretient des liens complexes avec le rendement : en 2P on observe un lien positif avec la lecture, mais négatif avec les mathématiques ; en 5P, le lien est négatif dans les deux branches.

Mc Donald et Elias (1976) soulignent qu'aucune conduite d'enseignement ne se révèle être un prédicteur significatif de rendement à la fois dans les deux matières et aux deux niveaux de scolarité étudiés. Autrement dit, il n'est pas possible d'identifier des conduites d'enseignement qui soient efficaces pour les différentes matières scolaires et pour les différents niveaux d'enseignement. Cette observation relance la discussion à propos de l'existence d'une méthodologie générale d'enseignement. Cette discussion a notamment été développée par Gage (1979).

L'étude ethnographique menée dans le cadre de la phase III-A de la BTES présente un intérêt méthodologique tout particulier. Contrairement à la tradition faisant loi dans les recherches processus-produit, les observateurs se présentent en classe sans grille d'observation ; leur attention n'est donc pas orientée vers des événe-

ments spécifiques. Au contraire, ici, les différents observateurs sont invités à s'adonner à une observation flottante et à rendre compte (de façon libre mais détaillée) des événements qui se déroulent sous leurs yeux. Par ailleurs, alors que la méthode ethnographique est généralement utilisée pour des études de cas impliquant un petit nombre de classes, dans la BTES, elle est appliquée à un échantillon de classes suffisamment large pour permettre des comparaisons entre classes et certains traitements statistiques.

La démarche suivie mérite d'être détaillée. On distinguera quatre étapes principales.

1. Les observateurs, pour la plupart des étudiants diplômés en sociologie ou en anthropologie, sont recrutés en fonction de leur capacité à rédiger un protocole d'observation. Ils sont familiarisés avec les classes de 2^e et 5^e années au cours de quelques journées de présence en classe, avant d'être entraînés à la rédaction de protocoles détaillés décrivant l'enseignement en lecture et en mathématiques.
2. Quatre enseignants sont assignés à chaque observateur : deux considérés comme efficaces et deux jugés moins efficaces. Les observateurs n'ont pas connaissance de la façon dont les enseignants sont qualifiés. Chacun d'entre eux reçoit la visite d'un observateur pendant une semaine entière.

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

3. Utilisant ses notes manuscrites, chaque observateur fait un compte rendu oral de ses observations sur un enregistreur audio. Celles-ci sont retranscrites par des secrétaires avant d'être soumises à des juges représentant différents types d'expertise pédagogique. Ces juges sont invités à examiner des paires de protocoles (l'une venant d'un enseignant efficace et l'autre d'un enseignant moins efficace). Travaillant en groupes, ces experts ont généré 61 variables (ou dimensions) permettant de décrire la plupart des protocoles d'observations.
4. Les données finales sont obtenues par examen des protocoles par de nouveaux juges. Ceux-ci sont à nouveau mis en présence de paires de protocoles et sont invités à établir dans quelle mesure chacune des 61 variables retenues à l'étape précédente est présente, ou non, dans le protocole. Pour cette dernière phase, on procède à la totalité des pairages, soit 10 enseignants efficaces confrontés à 10 enseignants moins efficaces = 100 paires de protocoles à examiner, et ceci pour chaque niveau scolaire et chaque matière (lecture, mathématiques). Finalement, 36 paires de protocoles sont sélectionnées au hasard pour chaque discipline et chaque niveau scolaire. Ce sont ces protocoles qui sont analysés par les experts.

De ces analyses, il ressort que 21 des 61 variables retenues présentent des différences significatives entre

maîtres efficaces et maîtres moins efficaces. La plupart des contrastes retenus s'avèrent pertinents quel que soit le contenu-matière et le niveau d'étude. Plus exactement, les contrastes apparus pour une discipline et à un niveau scolaire se retrouvent le plus souvent tels quels dans l'autre discipline et l'autre niveau scolaire ; dans les autres cas, la distinction n'est pas nette, mais ne va jamais en sens inverse.

Quel portrait du bon enseignant nous donne cette étude de type ethnographique ?

Si l'on suit ces résultats, le maître efficace est un enseignant qui prend plaisir à enseigner et qui se montre généralement convivial dans les interactions quotidiennes avec ses élèves. Il est capable d'appeler chacun de ses élèves par leur nom ou prénom. Il se montre attentif à ce que chacun dit et accepte l'expression des sentiments. Il manifeste sa sympathie à l'égard des élèves par des gestes amicaux (toucher l'épaule, par exemple), par des sourires fréquents, des clins d'yeux et des pointes d'humour. Il les encourage pour la plupart de leurs réussites et s'attache à les impliquer dans des décisions relatives aux tâches scolaires.

Ceci ne l'empêche pas de se montrer très exigeant sur le plan des activités scolaires. Il exhorte les élèves à s'investir pleinement dans les tâches et souligne avec force leurs responsabilités dans l'apprentissage. Bref, il les incite à recourir à des attributions causales internes contrôlables pour expliquer leurs succès et/ou échecs. En

général, il se montre cohérent dans ses exigences comme dans ses réprimandes. Il supervise minutieusement les progrès de chacun. Ses connaissances disciplinaires sont bonnes. Il est soucieux d'organiser au mieux la progression des apprentissages à travers le curriculum et recourt, selon les besoins, soit à des explications collectives, soit à des travaux individuels. Il a la volonté d'impliquer tous les élèves dans les apprentissages plutôt que de se concentrer sur certains sous-groupes. Le plus souvent, il pose des questions ouvertes et laisse le temps nécessaire pour qu'un nombre suffisant d'élèves répondent. Il se montre accueillant à l'égard des autres adultes qui viennent l'aider pendant les activités scolaires. De leur côté, les élèves de ce type d'enseignant sont (en règle générale) très impliqués dans leur travail et motivés par les tâches scolaires.

En ce qui concerne la discipline, les enseignants *efficaces* réagissent très vite aux petits problèmes, sans jamais se montrer humiliants à l'égard des élèves indisciplinés. A cet égard, on pourrait dire que les classes *efficaces* se caractérisent par le respect mutuel de l'enseignant et des élèves. Enfin, les enseignants *efficaces* gèrent le *momentum* des activités, c'est-à-dire qu'ils passent d'une activité à l'autre de façon harmonieuse. Lorsqu'ils font des remarques de discipline, ils font peu d'erreurs de *targeting* (Kounin, 1970), c'est-à-dire qu'ils se trompent rarement d'élèves. Circulant entre les bancs pendant les travaux individuels, ils se montrent vigilants à l'égard de tout indice de bonne ou mauvaise

compréhension des consignes (et/ou des concepts) et font preuve d'une bonne aptitude à la *withiness* (Kounin, 1970). Bref, le *bon enseignant* se caractérise par une attitude clairement volontariste en matière d'apprentissage scolaire. Par son attitude quotidienne, il manifeste à ses élèves combien il est convaincu que chacun peut apprendre. Ce haut niveau d'exigence en matière d'apprentissage s'accompagne d'une attitude clairement positive à l'égard des élèves, d'un contact convivial et même chaleureux.

Ce portrait, dressé par des observations ethnographiques, rejoint celui obtenu par les recherches processus-produit menées de façon classique. Toutefois, ce mode d'analyse de la vie en classe enrichit les données par la prise en considération des dimensions affectives et relationnelles absentes des études antérieures. On se trouve en définitive face à une validation et à un enrichissement des analyses quantitatives par l'approche qualitative.

Les résultats de la seconde *Field Study* confirment ceux des études précédentes. Son apport spécifique concerne la gestion du temps scolaire (*allocated time*) et, plus encore, la relation entre l'implication des élèves dans les tâches (*time on task*) et leur apprentissage¹¹. Cette étude apporte également une contribution impor-

¹¹ Cette partie de la BTES est discutée en détail au chapitre 5 de notre livre *L'école peut-elle être juste et efficace ?* (Crahay, 2000).

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

tante concernant le taux de succès des réponses des élèves pendant l'apprentissage.

En moyenne, 58 % du temps journalier est consacré aux disciplines scolaires de base (lecture, mathématiques, science et *social studies*) et 24 % sont dévolus aux activités non académiques comme la musique, les arts et la narration d'histoires. Les transitions ainsi que les temps d'attente et les activités de gestion de classes (gestion des collations, etc.) occupent en moyenne 18 % du temps. Les élèves sont sous la supervision directe de l'enseignant pendant 30 % du temps consacré aux activités scolaires et sont supposés « travailler » de façon autonome pendant les 70 % restants. Autrement dit, le rapport entre temps alloué aux activités collectives et temps consacré aux travaux individuels (*independent seatwork*) est de l'ordre de 30-70. En ce qui concerne l'implication des élèves (*time on task*), il apparaît qu'en moyenne, les élèves sont concentrés sur la tâche pendant 70 à 75 % du temps dévolu aux apprentissages. Cet indicateur est supérieur à 80 % pendant les activités collectives pour se situer entre 67 et 72% pendant les travaux individuels.

Les gains d'apprentissage sont associés à la quantité de temps alloué aux disciplines de base, mais aussi au pourcentage de temps pendant lequel l'élève est concentré sur la tâche ainsi qu'au taux de bonnes réponses enregistrées. À cet égard, les chercheurs de la BTES distinguent trois possibilités :

1. Réussite *élevée* : un travail complètement correct, hormis quelques erreurs dues au hasard ;
2. Réussite *moyenne* : cas où l'élève produit la moitié de bonnes réponses et la moitié de mauvaises, se montrant capable d'une compréhension globale des problèmes et, plus particulièrement, des règles générales ;
3. Succès *faible* : des bonnes réponses données par chance et par une incompréhension, assez générale, des concepts.

Lorsqu'on examine les phases d'exercices organisés en classe, en moyenne 50 % des élèves se trouvent dans la situation de *succès élevé* et les 50 % restants dans celle de *réussite moyenne*. La situation de réussite faible est exceptionnelle ; elle correspond aux cas où l'enseignant introduit un exercice alors que certains, voire la majorité des élèves, n'ont pas compris les concepts ou les connaissances préalables. D'une manière générale, les analyses corrélationnelles montrent une relation clairement négative entre la fréquence de ces situations de réussite faible et les gains d'apprentissage. À l'opposé, on observe une relation franchement positive entre la fréquence de situation de réussite élevée et les mesures de produit. Se basant sur ce dernier constat et l'articulant au précédent, les chercheurs de la BTES ont proposé le concept d'*academic learning time*

(ALT)¹². Celui-ci correspond au temps durant lequel les élèves s'investissent dans une tâche d'apprentissage portant sur une compétence ou une connaissance incluse dans la mesure de produit et pour laquelle le taux de réussite des réponses est élevé. La régularité de ce *pattern* de conditions serait garante d'efficacité.

La liaison causale entre un taux de succès élevé et l'apprentissage a été mise en cause par Shulman (1987). Selon cet auteur, cette corrélation atteste simplement du fait que les bons élèves commettent moins d'erreurs que les élèves moyens ou faibles. Burstein (1980) a réanalysé les données recueillies par la BTES : plutôt que de prendre les élèves comme unités d'analyse, il se concentre sur le maître et produit des scores agrégés. Sa conclusion le conduit à recommander une alternance de moments où les élèves sont en situation de réussite élevée et d'autres où ils sont en situation de réussite moyenne. Ce pattern correspondrait à un enseignement où le degré de difficulté des activités serait à la fois à la portée des élèves tout en constituant une sorte de défi. Pour Brophy et Good (1986), il est difficile de tirer une conclusion générale à ce sujet. A leurs yeux, la combinaison optimale, eu égard au taux de réussite, varie selon le contexte (en particulier, le niveau d'étude) et la phase d'apprentissage. Un taux de réussite moyen leur paraît prévisible dans les

¹² Ce concept est discuté au chapitre 5 de notre livre *L'école peut-elle être juste et efficace ?* (Crahay, 2000).

premières étapes d'un apprentissage alors qu'un taux de réussite élevé doit être attendu en fin de séquence.

Sur le plan méthodologique, il convient de se demander s'il est possible de répondre à une question aussi subtile par l'intermédiaire d'une recherche processus-produit menée sur un échantillon de classes aussi important que celui retenu dans le cadre de la *Second Field Study* de la BTES. Assurément, la problématique sous-jacente à ces analyses est cruciale. Elle renvoie à la difficile question de l'ajustement du niveau de difficulté des activités aux caractéristiques des élèves et, en particulier, à leur niveau de connaissances. Rappelons que cette question est notamment au coeur de la théorie de Vygotsky et de sa notion de *zone proximale de développement*. Par ailleurs, on peut douter que ces interrogations puissent être résolues par le calcul de corrélations entre des scores à un test et des mesures moyennes de taux de réussite. Rappelons, en effet, que les élèves observés dans le cadre de cette étude le sont à vingt reprises et ce, pendant une journée entière à raison d'un codage toutes les quatre minutes. Ces multiples prises d'informations sont ensuite globalisées dans un score moyen par élève dans le cas de la BTES, ou d'un score moyen par enseignant dans le cas de Burstein (1980). L'indicateur ainsi construit est fruste. Il reflète une tendance générale, insensible aux fluctuations inhérentes à tout processus d'apprentissage. Il est logique de supposer que tout individu qui apprend passe d'abord par une phase où les notions lui sont étrangères avant d'atteindre un niveau

intermédiaire où sa maîtrise est partielle et donc ses réponses tantôt correctes, tantôt erronées. À moins d'adhérer à la théorie skinnérienne de l'apprentissage sans erreur, il est difficile d'imaginer un processus d'enseignement-apprentissage où les élèves produiraient exclusivement des bonnes réponses.

Retour sur les études quasi-expérimentales du groupe de Canterbury

Dans les années 70, quelques chercheurs de l'Université de Canterbury ont mené des études quasi-expérimentales très rigoureuses afin de cerner la relation entre certains paramètres précis du processus d'enseignement et les apprentissages des élèves¹³. Nous nous concentrerons sur les trois études de Hughes (1973), puis sur celle de Nuthall et Church (1973).

Hughes (1973) a mené une série de trois études de type quasi-expérimental portant sur la participation des élèves et les réactions des enseignants aux réponses des élèves. Dans tous les cas, les leçons ont pour thème les animaux et leur mode de vie ; les élèves fréquentent un septième *grade* (1S). Dans la première étude, Hughes distingue trois façons de gérer la participation des élèves et, en conséquence, constitue trois groupes. Dans le premier, les questions de l'enseignant sont adressées aux

¹³ Exception à ce choix méthodologique : l'étude de Wright et Nuthall (1970) est de type corrélationnel. Nous n'en parlons pas car elle n'apporte rien de neuf par rapport aux études corrélationnelles discutées précédemment.

élèves de façon aléatoire. Dans le second groupe, les questions sont distribuées selon un schéma prévisible puisque déterminé par la place des élèves dans la classe : l'enseignant interroge le premier élève à sa gauche, puis le second, puis le troisième, et ainsi de suite. Enfin, dans la troisième modalité, l'enseignant interroge uniquement les volontaires. Utilisant un plan de recherche type pré-test/post-test, Hughes constate que les gains moyens observés dans les trois groupes ne diffèrent pas. Il conclut donc qu'il n'y a pas de relation entre les trois modalités de questionnement des élèves et l'apprentissage.

La seconde étude de Hughes (1973) confirme la conclusion de la première. Dans celle-ci, la moitié des élèves de la classe (sélectionnés au hasard) sont interrogés systématiquement, alors que l'autre moitié n'a aucune chance de pouvoir répondre aux questions de l'enseignant. À nouveau, les gains cognitifs ne permettent pas de différencier les deux groupes.

La troisième étude de Hughes (1973) porte sur les réactions des enseignants aux réponses des élèves. Deux groupes sont constitués : dans le premier groupe, les enseignants ont pour consigne de réagir avec enthousiasme à toutes les réponses correctes des élèves. Lorsque ceux-ci produisent une réponse erronée, ils ont pour mission de leur adresser de légers reproches ou de les sermonner modérément. Dans le second groupe, les enseignants réagissent aux réponses des élèves en signalant simplement si celles-ci sont correctes ou erronées. Dans ce cas, les gains sont nettement différents selon

que les élèves appartiennent au premier ou au second groupe. De façon précise, les élèves qui bénéficient d'une attitude réactive de l'enseignant apprennent davantage que les autres.

En conclusion de ces trois études, Hughes suggère que les élèves peuvent bénéficier d'un enseignement sans participer de façon explicite à la leçon. En revanche, leur apprentissage est affecté de façon importante par la qualité des réactions de l'enseignant aux réponses. L'enthousiasme que l'enseignant met dans ses réactions semble déterminant. En effet, dans la troisième expérience, ce qui distingue les deux groupes, c'est bien la chaleur des réactions de l'enseignant. Dans les deux cas, ils reçoivent une information sur ce qui est correct ou erroné, mais les premiers bénéficient d'un supplément d'enthousiasme qui fait la différence au niveau des mesures de produits.

L'étude de Nuthall et Church (1973) peut être considérée comme un modèle de recherche quasi-expérimentale. L'objectif est d'étudier l'effet propre et l'effet conjugué de trois paramètres de l'enseignement :

L'insistance sur le contenu à enseigner ou *insistance curriculaire* (simple, double ou triple) ;

La façon de stimuler et de réguler la participation des élèves par les questions ;

La fréquence des informations en retour donnée sur la qualité des réponses des élèves.

Notons d'emblée que les paramètres 2 et 3 sont emboîtés : la possibilité de délivrer de fréquents *feedbacks*

aux élèves dépend de l'importance de leur participation, laquelle est déterminée par les questions de l'enseignant.

Le dispositif de recherche mis en place s'apparente à un plan factoriel, propice à l'analyse de variance (tableau 8). Toutes les leçons enseignées portent sur le même contenu : les principes élémentaires d'électricité. Les enseignants doivent faire référence à deux problèmes illustrant chacun des trois principes d'électricité. Ceux-ci sont enseignés à des élèves de dix ans, selon des scénarios de leçons (*scripts*) conçus par les chercheurs de façon à faire varier systématiquement certains paramètres d'enseignement. Au total, neuf scénarios de leçons sont élaborés et implantés dans des classes. À ceux-ci s'ajoute un groupe de contrôle : les élèves appartenant à ces classes ne reçoivent aucun enseignement en électricité élémentaire.

Tableau 8 : Plan de la recherche de Nuthall et Church (1973)

Insistance curriculaire	Participation des élèves		
	Réglée par questions fermées	Réglée par questions ouvertes	Réduite
Simple	A		
Double	B	D, E, F	
Triple	C		G, H, I

Les trois premiers scénarios correspondent à la même façon d'animer la classe : les enseignants sont invités à

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

expliquer les concepts et principes fondamentaux et à recourir à des questions fermées hautement structurées. C'est l'insistance curriculaire qui les distingue. Dans le scénario A, les enseignants expliquent chaque concept et chaque principe, une et une seule fois alors que ceux-ci font l'objet de deux cycles d'enseignement dans le scénario B et de trois cycles dans le C. Ceci a des conséquences sur la durée de la séquence d'enseignement, sur le nombre d'épisodes interactifs ainsi que sur le nombre moyen de réponses par élève. Au total, ces trois scénarios se caractérisent comme suit (tableau 9) :

Tableau 9 : Caractéristiques des trois premiers scénarios conçus par Nuthall et Church (1973)

Scénario	Cycles (nbre)	Nombre et durée des leçons	Episodes interactifs (nbre)	Moyenne de réponses par élève
A	1	1 leçon de 50 min	54	5
B	2	2 leçons de 35 min.	100	9
C	3	3 leçons de 40 min.	179	15

Les scénarios D, E, F sont comparables au script B du point de vue de l'insistance curriculaire (2 cycles d'enseignement pour chaque concept et chaque principe). À la différence des trois premiers scénarios, ils sont conçus de sorte à favoriser la participation des élèves. Pour ce faire, les enseignants doivent recourir à

des questions ouvertes (peu structurées et relativement vagues). Ce qui les distingue, c'est la façon de réagir aux mauvaises réponses des élèves.

Dans le cas du scénario D, lorsqu'un élève donne une mauvaise réponse, l'enseignant redirige la question vers d'autres élèves. À la fin de chaque épisode interactif, il reformule les informations traitées de façon à en donner une version totalement correcte.

Dans le cas du script E, l'enseignant redirige la question vers le même élève, en formulant une évaluation ou *feedbacks*, mais aussi en impliquant de plus en plus d'élèves dans la discussion jusqu'à obtenir l'information correcte.

Enfin, dans le cas du scénario F, l'enseignant fournit aux élèves qui ont commis une erreur des informations complémentaires ou des indices pouvant les aider à réajuster leurs réponses ; le cas échéant, il fait appel à un (ou des) élève(s) de niveau de compétence supérieur et leur demande d'aider leurs condisciples en difficulté.

Par ailleurs, ces trois scénarios se caractérisent comme indiqué dans le tableau 10.

Tableau 10 : Caractéristiques des scénarios D, E et F conçus par Nuthall et Church (1973)

Scénario	Cycles (nbre)	Nombre et durée des leçons	Episodes interactifs (nbre)	Moyenne de réponses par élève
D	2	2 leçons de 35 min	80	5
E	2	3 leçons de 35 min.	90	9
F	2	3 leçons de 35 min.	97	15

Les trois derniers scénarios (G, H, I) sont à l'antipode des *scripts* D, E, F puisque la participation des élèves y est très peu sollicitée. Chaque concept et chaque principe correspond à trois cycles d'enseignement (comme dans le scénario C).

Dans le cas du scénario G intitulé *infrequent question*, le nombre de questions est réduit, de telle sorte que leur somme correspond à celle prévue dans le cadre du traitement A. Autrement dit, deux tiers des questions qui émaillent les scénarios dans la condition C ont été transformées en épisodes interactifs durant lesquels l'enseignant expose la matière.

Dans le cadre du scénario H intitulé *infrequent response*, les élèves cibles (ou sujets expérimentaux), c'est-à-dire ceux dont les gains d'apprentissage sont mesurés, ne sont quasiment jamais interrogés.

La condition expérimentale I, intitulée *infrequent feedbacks*, est totalement similaire au scénario C, si ce

n'est que les élèves reçoivent un minimum de *feedbacks*. Ceci signifie que les épisodes interactifs sont rarement clôturés par une information sur la qualité des réponses produites par les élèves.

Notons encore qu'un groupe de contrôle est constitué : la matière étudiée dans les groupes expérimentaux n'y est pas enseignée.

Chaque traitement a été implanté dans trois classes minimum, idéalement quatre. Afin d'obtenir des groupes équivalents, Nuthall et Church ont procédé par étapes. Leur but consiste à conserver des classes comportant à peu près 20 élèves dont les compétences initiales en électricité se situent dans des marges clairement définies. À cette fin, un nombre important de classes ont été pré-testées. Ont été retenues les classes qui comportaient 20 élèves répondant aux trois critères suivants :

- 1) Le quotient intellectuel se situe entre 90 et 129.
- 2) Un score supérieur à 75 à un prétest sur les prérequis cognitifs (en sciences et en lecture) est nécessaire pour profiter de l'enseignement.
- 3) Un score de l'ordre de 50 % au prétest sur le contenu des leçons (ce score correspond à celui qui est obtenu par un élève qui donne des réponses au hasard).

Les résultats obtenus sont présentés dans le tableau 11, ainsi que dans la figure 3 correspondante.

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

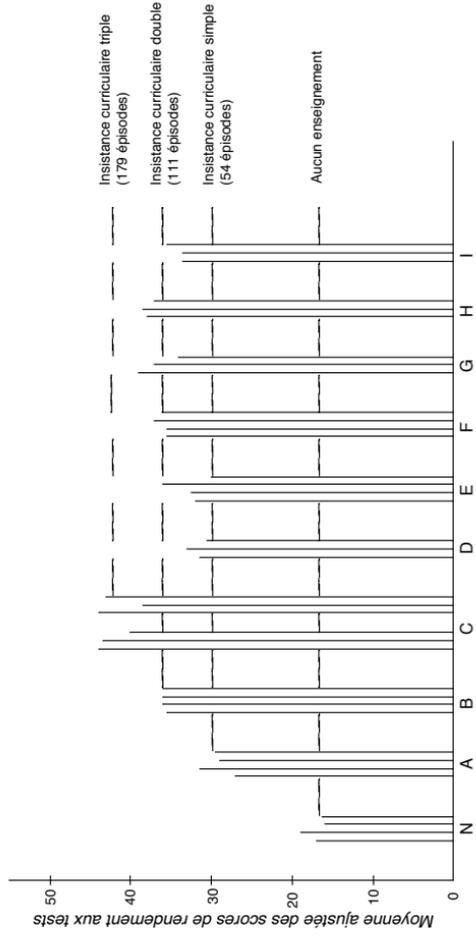
Tableau 11 : Résultats observés au terme de l'expérience de Nuthall et Church (1973) comparant neuf scénarios d'enseignement

Scénario	Intitulé	N° (classe)	Elèves (Nbre)	Moyenne ajustée
Contrôle	Pas d'enseignement	1	20	14.90
		2	20	16.34
		5	20	14.04
		6	20	14.25
A	Insistance curriculaire simple	33	20	23.92
		35	20	27.64
		36	17	25.36
		42	20	25.57
B	Insistance curriculaire double	29	20	30.57
		31	19	31.49
		34	20	31.39
		40	20	31.36
C	Insistance curriculaire triple	28	19	38.17
		30	18	36.13
		32	20	34.89
		58	19	38.29
		60	20	33.52
		62	18	37.09
D	lorsqu'un élève donne une mauvaise réponse, l'enseignant redirige la question vers d'autres élèves	41	19	27.80
		46	17	29.18
		51	19	26.89
E	l'enseignant redirige la question vers le même élève, en	43	16	27.98
		47	17	28.41
		48	20	32.23

MARCEL CRAHAY

	formulant une évaluation	55	16	26.55
F	l'enseignant fournit aux élèves qui ont commis une erreur des informations complémentaires	39	19	31.02
		45	14	31.02
		49	18	32.42
		52	15	31.13
G	<u>infrequent question</u> (nombre réduit de questions)	44	16	31.01
		57	15	32.67
		59	19	29.33
H	<u>infrequent response</u> (les élèves ne sont pas interrogés)	54	16	33.21
		64	19	33.89
		65	19	32.29
I	<u>infrequent feedbacks</u> (les élèves reçoivent un minimum de feedbacks)	53	18	29.72
		61	18	29.67
		63	18	31.18

Figure 3 : Effets des différents traitements expérimentaux conçus par Nuthall et Church.



Traitements expérimentaux

(chaque ligne du graphique représente la moyenne des scores obtenus pour une classe)

Outre la faiblesse des résultats obtenus dans le groupe contrôle, Nuthall et Church (1973) remarquent l'homogénéité des résultats à l'intérieur de chaque groupe ou traitement. Ils soulignent ensuite les différences de scores moyens entre les traitements A, B et C. Or, ces groupes se distinguent essentiellement par des insstances curriculaires différentes. De manière générale, ils évaluent à 37 % la part de variance expliquée par cette dimension. Discutant l'importance de ce constat, Nuthall et Church (1973) considèrent qu'il convient de retenir l'hypothèse selon laquelle une modification du temps consacré au contenu d'enseignement produit des effets d'apprentissage plus importants qu'une transformation des conduites d'enseignement visant à accroître la participation active des élèves.

Que dire de l'importance du nombre de questions adressées aux élèves ? Pour répondre à cette interrogation, il est pertinent de comparer les résultats du groupe G à ceux des groupes C et B. Il est alors possible de noter que réduire le nombre de questions de deux tiers engendre une réduction des effets d'apprentissage équivalente à celle que l'on obtient en diminuant le contenu enseigné d'un tiers. Dans le groupe H, les sujets expérimentaux bénéficient d'une insistance curriculaire équivalente à celle du groupe C, mais ils n'ont pas l'opportunité de répondre aux questions de l'enseignant (celles-ci sont essentiellement dévolues aux sujets non-expérimentaux de la classe). Il faut noter que la différence de rendement entre les groupes H et C est relative.

vement minime, ce qui conduit à penser que le fait de répondre aux questions de l'enseignant n'est pas aussi important qu'on pourrait l'imaginer.

Quel est l'effet des questions ouvertes ? Le fait de faire participer les élèves à l'aide de questions ouvertes a une première conséquence sur le temps d'enseignement : pour chaque segment de contenu, il faut plus de temps pour aboutir à l'information correcte. À cet égard, il est intéressant de constater que les gains d'apprentissage observés dans le groupe F sont équivalents à ceux du groupe B. Or, dans les deux cas, chaque concept et chaque principe est étudié à deux reprises et les élèves sont confrontés à quatre problèmes. Bref, dans ces deux groupes, le contenu enseigné est équivalent et les résultats également, peu importe la modalité de questionnement retenue. Par ailleurs, les scores observés dans les autres traitements utilisant les questions ouvertes sont inférieurs. Ceci conduit à la conclusion qu'en cas de gestion de la classe par questions ouvertes, il convient de réagir aux mauvaises réponses des élèves en leur fournissant des informations ou indices complémentaires plutôt qu'en proposant directement la bonne réponse (traitement D) ou en donnant un *feedbacks* évaluatif (traitement E). Enfin, il faut souligner que les scores observés dans les modalités D et E sont comparables à ceux observés dans le traitement I (réduction du nombre de *feedbacks*); on pourrait considérer que dans ces trois modalités d'enseignement, les élèves arrivent difficile-

MARCEL CRAHAY

ment à identifier quelle est, en définitive, l'information valide à retenir.

CHAPITRE VI

PENSER LE DÉPASSEMENT DU PARADIGME PROCESSUS-PRODUIT

L'enseignant efficace existe-t-il ?

Les études de stabilité d'Ackland (1976), Brophy (1973), Emmer, Evertson et Brophy (1979) et Good et Grouws (1975) obligent à poser cette question : leurs résultats sont loin d'avoir confirmé les attentes. De surcroît, les résultats obtenus au cours de la *First Field Study* de la BTES (Mc Donald & Elias, 1976) renforcent les doutes et l'interrogation. Pour rappel, dans cette étude, aucune conduite d'enseignement ne se révèle être un prédicteur significatif de rendement à la fois dans les deux matières et aux deux niveaux de scolarité étudiés. Alors, que penser du postulat de l'enseignant efficace s'il ne s'avère pas possible de désigner des conduites d'enseignement efficaces à la fois pour les différentes matières scolaires et pour les différents niveaux d'enseignement ? Nécessairement, il faut questionner la notion d'*effet-maître* et celle de compétences transver-

sales en matière d'enseignement ou, pour reprendre les termes de Gage (1979) de la généralité des dimensions de l'enseignement efficace.

Si un même enseignant, opérant plusieurs années successives dans le même établissement et au même niveau du cursus, produit des effets d'apprentissage qui varient de manière assez forte, nous pouvons en conclure qu'il n'est pas totalement responsable de ses effets ou du fait que son efficacité fluctue d'année en année. D'autres revues de la littérature (notamment Crahay, 2000) laissent penser que les caractéristiques des groupes d'élèves jouent un rôle. Ainsi, il est démontré que les effets d'*agrégation* prétéritent lourdement l'équité des systèmes éducatifs (Crahay & Monseur, 2006). Par ailleurs, l'examen contrasté des études expérimentales et « naturelles » relatives aux classes de niveau suggère que les *a priori* des enseignants quant aux potentialités des élèves affectent davantage l'amplitude des gains académiques que la composition du groupe-classe elle-même. Bref, il semble légitime de s'interroger sur la part de l'*effet-classe* dans ce qu'on appelle l'*effet-maître* (Bressoux, 1994).

Plus fondamentalement encore, il convient – pensons-nous – de mettre en discussion le glissement conceptuel induit par la méthodologie processus-produit. Car, entre la réflexion séminale de Gage (1963) et autres pionniers (notamment Dunkin & Biddle, 1974 ; Nuthall, 1968 ; Smith, 1960 ; Turner, 1971) et les programmes de recherches examinés ici, il faut reconnaître que l'on est

passé de l'analyse de l'enseignement à la recherche de l'enseignant efficace. Un glissement est contenu dans la méthode de recherche adoptée par la plupart des chercheurs : l'unité d'analyse retenue est la classe et les indicateurs de processus introduits dans le calcul de corrélation sont choisis de façon à rendre compte de la variance inter-enseignants ou inter-classes. Seul le groupe de Canterbury a fait un choix différent : l'unité d'analyse est la leçon ou la séquence de leçons.

Comme l'a dénoncé Doyle à diverses reprises (1977, 1978 et 1986 notamment), les recherches sur l'efficacité des enseignants tendent à neutraliser les variables de contexte plutôt que de chercher à en saisir les effets sur le processus d'enseignement. À nouveau, la méthode de recherche adoptée est en cause : une fois que les chercheurs procèdent à des observations dans un grand nombre de classes (lors d'activités sélectionnées aléatoirement) et que l'on calcule, pour chaque catégorie, la moyenne par classe, la variabilité intra-classe est effacée. Seule la variabilité inter-classes est retenue, mais comme moyen de mettre en évidence la variabilité des indicateurs de processus retenus et de les corréler avec les mesures de produit. Bref, les recherches processus-produit opèrent sur la base du postulat que les variables de processus reflètent la façon de faire des enseignants, leur personnalité, leur expérience professionnelle et, en définitive, leur expertise ou leur compétence (Medley, 1977). Quoiqu'elle ne soit pas niée sur le plan théorique, l'influence du contexte dans lequel se déroule

l'enseignement n'est pas conceptualisée. Or, celle-ci est bien réelle ; d'autres travaux en attestent (Bayer, 1979, 1992 ; Crahay, 1988 ; Doyle, 1979 a & b, 1983). C'est notamment ce qui a amené divers chercheurs à plaider en faveur d'une révolution paradigmatique, fondée sur un postulat inverse : « En classe, ce n'est pas le maître qui contrôle la situation, mais la situation qui contrôle le maître » (Doyle & Ponder, 1975, p. 184). En définitive, pour Doyle (1986), il conviendrait de conceptualiser l'enseignement dans une perspective inspirée de la psychologie écologique (Barker, 1968). Minimale, il convient de faire retour sur l'analyse de l'enseignement, d'en saisir les réalités et les dynamiques (Jackson, 1966). Ce retour conduit nécessairement au questionnement de Gage (1979) quant au caractère général des dimensions de l'enseignement.

Sans doute, en s'inspirant de Gage (1979), faut-il distinguer, à côté de dimensions transversales, d'autres trouvant leurs spécificités ailleurs : selon le contenu ou l'objet d'enseignement, les caractéristiques des élèves (principalement socioculturelles), le cadre institutionnel (le niveau d'enseignement notamment).

L'efficacité en matière d'enseignement

Les résultats des recherches processus-produit sont remarquablement cohérents en ce qui concerne l'importance d'un faisceau de paramètres : les opportunités éducatives, la couverture du contenu (*content coverage*), le temps alloué aux apprentissages scolaires et le

temps d'implication des élèves (ou *time on task*). La notion de *content overlap* (Barr, 1987) cerne sans doute au mieux ce qui est déterminant : l'enseignement doit être ciblé sur ce qui est mesuré par le test ; le temps dévolu aux apprentissages correspondants influe directement sur le temps durant lequel les élèves s'y engagent. Par voie de conséquence, les stratégies d'enseignement et/ou de décisions managériales (Doyle, 1979) qui ont pour effet de maintenir l'attention des élèves sur la tâche, sont également liées au rendement : transition douce entre activités et gestion proactive et avisée de la discipline (cf. à cet égard les concepts inventés par Kounin, 1970). Notons que ce ciblage sur les tâches scolaires n'est pas incompatible avec un climat émotionnel convivial et chaleureux. Au contraire, comme le montrent les travaux de Soar et Soar (1979) et l'étude ethnographique de la BTES, la complémentarité de ces caractéristiques est fréquente (donc possible) et souhaitable. En définitive, il semble que l'enseignement efficace consiste à créer les conditions (1) affectives et (2) organisationnelles propices à une implication forte des élèves dans les apprentissages scolaires.

- 1) Le climat émotionnel ou relationnel est crucial. Plusieurs recherches (principalement, Brophy & Evertson, 1976 ; Mc Donald & Elias, 1976 ; Soar & Soar, 1979) attestent de l'importance des attitudes, gestions, encouragements, etc. par lesquels l'enseignant manifeste son souci que tous les élèves apprennent et sa

confiance en leurs capacités. Les résultats de l'étude ethnographique de la BTES sont particulièrement clairs à cet égard : par un ensemble de signes verbaux et non-verbaux, l'enseignant efficace communique sa volonté que tous réussissent et sa conviction que ce but sera atteint.

- 2) La gestion des conditions de travail scolaire est également capitale. Toutes les recherches recensées ici confirment ce que Soar (1968) fût – semble-t-il – le premier à affirmer clairement : les gains d'apprentissage des élèves sont particulièrement importants lorsque l'enseignant est, à la fois clairement orienté vers les connaissances et les compétences à maîtriser et qu'il exerce un contrôle strict sur les tâches à réaliser. De manière générale, il est peu productif de laisser aux élèves le choix des activités. Bien plus, le rythme de progression dans les tâches doit être sous le contrôle de l'enseignant, tout particulièrement lorsque les élèves sont issus de milieux modestes. À cet égard, la conceptualisation et les travaux de Gump (1967) sont éclairants : ils montrent que l'engagement des élèves dans les tâches est plus labile lorsque ceux-ci sont supposés, étant donné le format de l'activité, gérer par eux-mêmes leur rythme de travail (*self-pacing activities*).

Ces résultats relatifs à la gestion des tâches d'apprentissage vont à l'encontre de convictions propres aux adeptes de l'Education Nouvelle. Gauthier (2005)

s'en est d'ailleurs servi pour argumenter à l'encontre du socioconstructivisme. Sans doute, cette opposition résulte-t-elle en partie des bases scientifiques auxquelles font recours les uns et les autres : les partisans actuels du socioconstructivisme enracinent leurs convictions dans la psychologie piagétienne tandis que Gauthier se situe dans la foulée de Gage (1963) pour réclamer des connaissances scientifiques sur la réalité de l'enseignement. Nous reviendrons ci-dessous sur la nécessité, selon notre point de vue, d'articuler *psychologie de l'apprentissage* et *savoirs sur l'enseignement*. À ce stade-ci de la réflexion, il nous faut enregistrer trois constats importants et complémentaires, établis par les recherches portant sur l'efficacité de l'enseignement :

- Celui-ci doit être direct c'est-à-dire que les opportunités d'apprentissage doivent être en lien direct avec les objectifs poursuivis et, parallèlement, avec la nature des épreuves qui seront utilisées pour évaluer les acquis des élèves.
- L'insistance curriculaire, c'est-à-dire le volume de ces opportunités d'apprentissage, ainsi que le temps d'implication des élèves dans ces tâches constitue des composantes- clé d'une pédagogie de la réussite.
- L'engagement des élèves dans les tâches scolaires est particulièrement délicat en situation de travail individuel : les occasions de distraction sont plus nombreuses qu'en situation de leçon magistrale et il im-

porte que l'enseignant assure un *monitoring* rapproché du travail des élèves.

Les caractéristiques des élèves sont à prendre en considération. Comme l'ont souligné Brophy et Evertson (1976), les élèves d'origine socio-économique élevée manifestent un rapport à l'école et aux apprentissages scolaires différents de ceux issus des minorités ethniques et/ou culturelles. Les premiers se montrent d'emblée confiants en leurs moyens et témoignent d'une *agentivité* (Bandura, 2003) de bon aloi. Les autres sont souvent incertains, anxieux voire timorés. Aux premiers, des situations problèmes, des défis peuvent être proposés avec de bonnes chances de succès. Face aux seconds, il convient – semble-t-il – de se montrer plus précautionneux : les activités proposées doivent être à leur portée car il serait préjudiciable de les placer en situation d'échec. De même, Turner (1967), discutant des effets possibles de l'influence indirecte (Flanders, 1970), indiquait que ce type de pédagogie était probablement mieux adapté aux élèves brillants qu'à ceux qui éprouvent des difficultés. De ces derniers, il faut attendre moins d'initiatives et moins d'attitudes assertives que des autres. Avec eux, le traitement des erreurs, le rythme de progression, le taux de redondances et la clarté des explications doivent être conçus de sorte qu'ils ne perdent pas pied (Brophy & Evertson, 1976). En matière de prise de parole, il convient de veiller à ne freiner aucun d'entre eux et d'éviter que les plus loquaces n'écrasent les plus taiseux.

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

Ces conclusions résument des résultats obtenus au terme de plusieurs programmes de recherche portant sur d'importants échantillons d'élèves de l'enseignement primaire aux Etats-Unis. Il n'est pas évident qu'elles puissent être généralisées à l'enseignement secondaire ; seule l'étude de Evertson, Anderson, Anderson et Brophy (1980) porte sur ce niveau du cursus scolaire et si les résultats relatifs aux mathématiques sont clairs, ceux relatifs à la langue maternelle sont plus difficiles à interpréter. Ces conclusions valent donc essentiellement pour les apprentissages de base en lecture et en mathématiques. Notons aussi le caractère déterminant des compétences prises pour critères d'efficacité. À cet égard, la démonstration de Soar & Soar (1979) est cruciale : même si les liaisons entre les paramètres de processus et les mesures de produit sont positives quel que soit le critère considéré, elles sont de type linéaire dans certains cas et curvilinéaires avec des points d'optimum variables dans d'autres cas. En termes courants, ces chercheurs montrent que la fréquence optimale des encouragements et des critiques dépend du type de compétence visée par l'enseignement : le développement de la créativité et de démarches de raisonnement demandent plus d'encouragement et moins de critiques que le développement de skills de lecture. Ceci est évidemment capital et atteste de l'importance de la mesure de produit adoptée dans les recherches sur l'efficacité de l'enseignement. Or, hormis Soar et Soar (1979), les autres chercheurs impliqués dans ce courant de re-

cherche n'ont guère problématisé la question de la mesure des produits, procédant comme si la notion d'efficacité allait de soi. Certes, il faut reconnaître que les tests standardisés qui inondent le marché scolaire américain sont ciblés sur les *key competencies* visées par l'école. Il n'en reste pas moins que la démonstration de Soar et Soar (1979) indique que des analyses fines des liaisons processus-produit en fonction du critère d'efficacité adopté eurent été opportunes. Cette critique n'a pas pour fonction d'annuler la puissance des résultats engrangés, mais d'en limiter la généralité. Portée sur le plan épistémologique, elle ouvre à une réflexion de fond sur les recherches sur l'efficacité de l'enseignement.

Comment l'enseignement contribue-t-il à l'apprentissage des élèves ?

Pour dépasser le paradigme processus-produit, il importe – pensons-nous – de revenir sur les fondements même de l'analyse de l'enseignement tels qu'ils ont été conceptualisés par les pionniers, Gage (1963) en tête. Pour ce chercheur, comme indiqué au chapitre I, l'enseignement doit être et a été considéré comme conceptuellement distinct du processus d'apprentissage. En quelque sorte, Gage (1963) mais aussi Dunkin et Biddle (1974), Nuthall (1968), Smith (1960), Turner (1971) prennent le contre-pied de la position initiée par Claparède (1946) dans *Psychologie de l'enfant et pédagogie expérimentale*. Pour le psychologue genevois, il faut enraciner la pédagogie dans la psychologie de l'enfant.

Ce positionnement épistémologique se prolonge et s'approfondit avec Piaget (1969) dans *Psychologie et pédagogie*. Il se perpétue aujourd'hui avec les adeptes du socioconstructivisme (par exemple Jonnaert, 2002). À l'opposé, Gage (1978) nourrit le projet scientifique de construire *The scientific basic of the art of teaching*, titre de son ouvrage-clé. Évidemment, les recherches processus-produit participent à cet ambitieux dessein.

Les limites d'une approche ancrée exclusivement dans la psychologie de l'enfant et/ou du développement cognitif ont été soulignées par Gage (1963, 1978), mais aussi par les didacticiens (Schneuwly, 1997). Ces critiques nous paraissent irréfutables et, en conséquence, on ne peut que souscrire à leur point de vue : la pédagogie et plus largement les sciences de l'éducation ne peuvent se contenter d'une référence théorique à la psychologie. Les bases scientifiques de l'art d'enseigner doivent être construites par le biais de recherches qui en font leur objet, mais celles-ci ne peuvent plus ignorer les apports de la psychologie. Bref, il convient de réconcilier le point de vue de Gage (1963 ; 1978) avec celui de Claparède (1946) et de Piaget (1969) ; il faut tenter d'articuler analyse de l'enseignement et psychologie de l'apprentissage et du développement.

Doyle (1978, 1986) va clairement dans ce sens lorsqu'il propose d'adopter, en matière de recherche sur l'efficacité de l'enseignement, le paradigme des processus médiateurs. Il suggère notamment de s'intéresser aux réponses *mathémagéniques* des élèves – littéralement,

comportements qui donnent naissance à l'apprentissage – et de distinguer les *stimuli* nominaux et les *stimuli* effectifs¹⁴. Les stimuli nominaux sont les aspects d'une manifestation d'enseignement objectivement définis ; les stimuli effectifs sont ceux qu'un élève traite réellement à un moment donné. La transformation sélective de stimuli nominaux en stimuli effectifs est opérée par l'apprenant. C'est donc lui et non l'enseignant qui construit des connaissances et/ou des compétences. Par ailleurs, en classe, c'est bien au départ de l'enseignement que l'élève apprend car les opportunités d'apprentissage sont nécessaires, même si elles ne sont pas suffisantes.

Le paradigme processus-produit procède d'une logique *plus-plus* : plus il y a d'opportunités d'apprentissage, de temps d'enseignement, de *feedbacks*, etc. plus il y aura de connaissances et de compétences maîtrisées. Le paradigme des processus médiateurs a l'ambition d'intégrer et de dépasser l'apport des recherches processus-produit. Pas question donc d'ignorer les acquis de ce courant de recherche. Plus justement, il s'agit de comprendre en profondeur comment l'enseignement contribue à l'apprentissage, en rendant à l'élève sa juste place dans le processus, celle d'un sujet qui sélectionne, interprète et transforme les stimuli nominaux rendus disponibles par l'enseignement. À cet égard, les travaux de Rothkopf (1970 ; 1976) font

¹⁴ Pour une recherche bâtie sur cette distinction, voir Crahay & Delhaxhe (1991).

figure de précurseurs. Ce chercheur a notamment montré que la sélection des stimuli à transformer est déterminée par la perception qu'a l'élève, d'une part, des exigences de la tâche d'apprentissage et, d'autre part, du degré d'adéquation des stimuli disponibles pour l'accomplissement de cette tâche. Plus largement, dans cette entreprise scientifique visant à coordonner l'analyse de l'enseignement et des démarches d'apprentissage, il semble nécessaire de prendre en considération deux types de représentations que se construit l'élève :

1. Les représentations de l'enseignant, de ses buts et objectifs en matière d'apprentissage (comment l'élève se représente-t-il ce que l'enseignant cherche à lui faire apprendre ? Quelle représentation se fait-il de ce que l'enseignant pense de lui, capable ou non d'apprendre ?)
2. Les représentations de la tâche et du type de savoir qu'il s'agit de construire. Ces représentations sont à voir à la fois comme résultantes de l'enseignement et comme variables médiatrices. L'élève se construit une représentation de l'école, de ses buts, de son fonctionnement et des façons de s'y comporter et d'y réussir, ainsi qu'une conception de l'apprentissage mais aussi un rapport au savoir à travers son expérience de l'enseignement. Quant à celui-ci, il n'affecte pas l'élève par voie directe : l'enseignement est perçu et interprété par l'élève en fonction de ses représentations et conceptions. Ici, agissent les stra-

tégies de l'enseignant régulièrement mobilisées en situation de classe (Bressous & Dessus, 2003), mais aussi des événements singuliers ou incidents critiques.

Comme l'argumente Doyle (1986), on peut raisonnablement soutenir que des modalités autres que la fréquence revêtent une importance égale sinon supérieure pour l'étude de l'efficacité. On est ainsi en droit de supposer que des comportements rares, c'est-à-dire exceptionnels par rapport au modèle général, peuvent être spécialement porteurs d'informations pour les élèves et produire d'importants effets sur leurs apprentissages. Enfin, des modalités qualitatives du comportement de l'enseignant et du matériel didactique (par exemple, des livres scolaires bien adaptés à la capacité en lecture de l'élève) peuvent aussi peser significativement (Berliner, 1976). Par conséquent, il semble nécessaire de penser les méthodes de recherche sur les effets de l'enseignement, en prenant pour acquis que la seule fréquence de certains comportements de l'enseignant ne peut rendre compte de la complexité des relations processus-produit qui se tissent au sein de la classe. C'est ce qu'avaient compris les chercheurs qui ont conçu l'étude ethnographique correspondant à la phase IIIa de la BTES et ceux que nous avons nommé le groupe de Canterbury. Bref, c'est sans doute en jonglant avec les approches corrélationnelles, ethnographiques et quasi-expérimentales qu'il sera possible de bâtir une base de connaissances relative au processus enseignement-apprentissage.

Pour dépasser le paradigme processus-produit, il convient également d'interroger la perspective naturaliste qui est en son cœur¹⁵. Celui-ci repose, en effet, sur le postulat implicite que toutes les lois du processus enseignement-apprentissage sont inscrites dans l'univers des situations de classe actuellement repérables. Il suffit donc d'observer les pratiques courantes des maîtres et d'identifier celles qui sont associées aux rendements scolaires les plus élevés – ceux-ci étant mesurés par des tests standardisés – pour déceler les bonnes pratiques. Ceci revient à assimiler la classe à un milieu naturel qui aurait ses lois propres auxquelles nul ne peut se soustraire. Postulat inacceptable à nos yeux, car il impose d'ignorer combien l'*habitus* des enseignants est marqué du sceau de l'histoire, de la tradition des Frères des écoles chrétiennes, de la pensée empiriste de Herbart et d'autres influences qui, au cours du temps, se sont sédi-

¹⁵ Soulignons que le point de vue *naturaliste* de Gage (1978) et, plus largement, des études processus-produit n'a rien d'exceptionnel. En fait, il est dominant dans le monde de l'éducation. On le retrouve au cœur même de l'Éducation Nouvelle et de la pédagogie fonctionnelle. Une fois encore, il suffit de se référer à Claparède (1912) pour en avoir la preuve. Selon le pédagogue genevois, « la vérité de la conception fonctionnelle de l'éducation tire (...) un haut degré de probabilité du fait qu'elle se déduit naturellement de lois biologiques générales qui se sont montrées elles-mêmes fécondes (la loi de l'évolution, la loi de l'intérêt et du besoin, la psychologie des fonctions opposées à la psychologie des facultés-entités), ... » (p. 129).

mentées en pratiques courantes. Sur le plan épistémologique¹⁶, le raisonnement sous-jacent à toute l'approche processus-produit illustre parfaitement ce que Popper (1982) a appelé *the common sense problem of induction* ; celui-ci consiste à considérer légitime de faire des prévisions à partir des régularités constatées dans le passé et à croire fermement en leur validité, voire en leur inéluclabilité, dans la mesure où celles-ci reposent sur un nombre important d'observations. Certes, ce type d'induction peut aider à rationaliser nos rapports au réel, mais risque aussi de limiter notre créativité. On ne peut s'empêcher ici de songer à la critique qu'adresse Papert (1981) aux recherches pédagogiques classiques qui ont pour but l'amélioration de l'enseignement en classe. Pour lui, ces recherches évoquent le travail d'un savant qui, au

¹⁶ Pour une discussion approfondie sur le plan épistémologique et philosophique du paradigme processus-produit, le lecteur se reportera à Fenstermacher (1980). On trouve également chez Charbonnel (1988), une réflexion critique portant sur les fondements de *la raison éducative*. Cette auteure souligne que celle-ci fonctionne le plus souvent au départ d'une confusion « entre lois idéales et lois existantes et même plus exactement entre principes normatifs idéaux d'une part, principes normatifs existants d'autre part, et lois de fonctionnement (ces dernières "inventées" par le seul chercheur) (...). Est mis sur le même plan ce qui devrait idéalement se passer (...) , ce qui est considéré dans une société comme devant se passer (...) et ce qui se passe (...) Tout cela est amalgamé dans une même notion de loi, ... » (p. 96).

XIX^e siècle, se serait obstiné à étudier les conditions optimales de fonctionnement des moyens de transport connus à l'époque. Or, il remarque :

L'invention de l'automobile, pas plus que celle de l'avion, n'ont découlé d'une étude détaillée destinée à savoir ce qui marchait ou ne marchait pas chez leurs ancêtres les voitures à chevaux (...). La vraie question est de savoir s'il est possible d'inventer l'équivalent de l'automobile (1981, p. 19).

Bref, si l'on suit le raisonnement de Papert, il est envisageable d'inventer de toutes pièces de nouvelles pratiques d'enseignement qui, à l'épreuve de la réalité, s'avèrent plus efficaces que celles qu'on observe dans les classes d'aujourd'hui.

BIBLIOGRAPHIE¹⁷

Liste des revues portant sur les recherches processus-produit.

Dunkin, M.J. (Ed.) (1987). *The International Encyclopedia of Teaching and Teacher Education*. Oxford : Pergamon Press.

Barnes, J. Teaching Experience, pp. 608-612.

Barr, R. Content Coverage, pp. 364-368.

Biddle, B.J. Teacher Roles, pp. 625-634.

Biddle, B.J. and Dunkin, M.J. Effects of Teaching, pp. 119-124.

Braun, C. Teachers' Expectations, pp. 598-605.

Calfee, R.C. and Piontkowski, D.C. Grouping for Teaching, pp. 225-232.

Doenau, S.J. Soliciting, pp. 407-413.

Doenau, S.J. Structuring, pp. 398-407.

Doyle, W. Paradigms for Research, pp. 113-119.

Dunkin, M.J. Lesson Formats, pp. 263-266.

¹⁷ Cette bibliographie a pour ambition de fournir un échantillon significatif de références clefs eu égard aux thématiques discutées dans cet opuscule.

MARCEL CRAHAY

- Gump, P.V. Activities: Structures and Functions, pp. 452-457.
- Levis, D.S. Teachers' Personality, pp. 585-589.
- Medley, D.M. Criteria for Evaluating Teaching, pp. 169-181.
- Pösner, G.J. Pacing and Sequencing, pp. 266-272.
- Rosenshine, B. Direct Instruction, pp. 257-262.
- Shavelson, R.J. Interactive decision-making, pp. 491-493.
- Shavelson, R.J. Planning, pp. 483-486.
- Shavelson, R.J. Teachers' Judgments, pp. 486-490.
- Slavin, R.E. Small Group Methods, pp. 237-243.
- Soar, R.S. and Soar, R.M. Classroom Climate, pp. 336-342.
- Withall, J. Teacher-centred and Learner-centred Teaching, pp. 327-336.
- Zahorik, J.A. Reacting, pp. 416-423.
- Witrock, M.C. (Ed.) (1986). *Handbook of Research and Teaching, Third Edition*. New York: Macmillan.
- Biddle, B.S. and Anderson, D.S. Theory, Methods, Knowledge, and Research on Teaching, pp. 230-252.
- Brophy, J.E. and Good, T.L. Teacher Behavior and Student Achievement, pp. 328-375.
- Cazden, C.B. Classroom Discourse, pp. 432-463.
- Clark, C.M. and Peterson, P.L. Teachers' Thought Processes, pp. 255-296.
- Doyle, W. Classroom Organization and Management, pp. 392-431.

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

- Rosenshine, B.V. and Stevens, R. Teaching Functions, pp. 376-391.
- Mitzel, H.E. (Ed.) (1982). *Encyclopedia of educational Research, Fifth Edition*. New York: The Free Press.
- Kleine, P.F. Teaching Styles, pp. 1927-1932.
- Medley, D.M. Teacher Effectiveness, pp. 1894-1903.
- Ryan, K. and Phillips, D.H. Teacher Characteristics, pp. 1869-1876.
- Weil, M.L. and Murphy, J. Instruction Processes, pp. 890-917.
- Travers, R.M.W. (Ed.) (1973). *Second handbook of Research on Teaching*. Chicago: Rand McNally.
- Bidwell, C.E. The Social Psychology of Teaching, pp. 413-449.
- Dreeben, R. The School as a Workplace, pp. 450-473.
- Lortie, D.C. Observations of teaching as Work, pp. 474-497.
- Nuthall, G. and Snook, I. Contemporary Models of Teaching, pp. 47-76.
- Rosenshine, B. and Furst, N. The Use of Direct Observation to Study Teaching, pp. 122-183.
- Ebel, R.L. (Ed.) (1969). *Encyclopedia of Educational Research, Fourth Edition*. New York: Macmillan.
- Anderson, R.H. and Ritscher, C. Pupil Progress, pp. 1050-1062.
- Biddle, B.J. Teacher Roles, pp. 1437-1446.
- Flanders, N.A. and Simon, A. Teacher Effectiveness. pp. 1423-1437.

MARCEL CRAHAY

- Gage, N.L. Teaching Methods, pp. 1446-1458.
Heathers, G. Grouping, pp. 559-570.
Gage, N.L. (Ed.) (1963). *Handbook of Research on Teaching*. Chicago: Rand McNally.
Getzels, J.W. and Jackson, P.W. The Teacher's Personality and Characteristics, pp. 506-582.
Wallen, N.E. and Travers, R.M.W. Analysis and Investigation of Teaching Methods, pp. 448-505.
Withall, J. and Lewis, W.H. Social Interaction in the Classroom, pp. 683-714.
Harris, C.W. (Ed.) (1960). *Encyclopedia of Educational Research, Third Edition*. New York: Macmillan.
Mitzel, H.E. Teacher Effectiveness, pp. 1481-1491.
Ryans, D.G. Prediction of Teacher Effectiveness, pp. 1486-1491.
Stiles, L.J. Instruction, pp. 710-715.
Wingo, G.M. Methods of Teaching, pp. 848-861.

Références bibliographiques

- Acland, H. (1976). Stability of teacher effectiveness : A replication. *Journal of educational Research*, 69, 289-292.
Adams, R.S. (1969). Location as a feature of instructional interaction. *Merril Parker Quarterly*, 15 (4), 309-321.

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

- Adams, R.S. & Biddle, B.J. (1970). *Realities of teaching : exploration with videotape*. New York : Holt, Rinehart and Winston.
- Alexander, L., Frankiewicz, R. & Williams, R. (1979). Facilitation of learning and retention of oral instruction using advance post organizers. *Journal of Educational Psychology*, 71, 701-707.
- Amidon, E. & Hunter, E. (1967). *Improving teaching: the analysis of classroom verbal interaction*. New York : Holt, Rinehart and Winston.
- Anderson, L., Evertson, C. & Brophy, J. (1979). An experimental study of effective teaching in first-grade reading groups. *Elementary School Journal*, 79, 193-223.
- Anderson, L.W. & Burnes, R.B. (1989). *Research in classrooms. The study of teachers, teaching and instruction*. Oxford : Pergamon Press.
- Bandura, A. (1997/2003). *Auto-efficacité : le sentiment d'efficacité personnelle*. Bruxelles : De Boeck.
- Barker, R.G. (1968). *Ecological psychology*. Stanford : Stanford University Press.
- Barr, R. & Dreeben, R. (1977). Instruction in classrooms. *Review of Research in Education*, 5, 89-162.
- Bayer, E. (1972). *Analyse multidimensionnelle de la communication d'enseignement*. Thèse de doctorat en sciences de l'éducation, Université de Liège.
- Bayer, E. (1973). L'analyse des processus d'enseignement. *Revue Française de Pédagogie*, 24, 30-40.

- Bayer, E. (1979). Sources de variance de l'indice d'influence de Flanders. *Scientia Paedagogica Experimentalis*, 3, 5-21.
- Bayer, E. (1986). Une science de l'enseignement est-elle possible ? In M. Crahay & D. Lafontaine (Ed.), *L'art et la science de l'enseignement* (pp. 483-507). Bruxelles : Labor.
- Bayer, E. & Ducrey, F. (1998). Une éventuelle science de l'enseignement aurait-elle sa place en sciences de l'éducation ? *Raisons éducatives*, 243-276.
- Bellack, A.A., Human, R.T., Smith, Fl. & Kiedbard, H.M. (1966). *The language of the classroom*. New York : Teachers College Press.
- Berliner, D. (1976). Impediments to the study of teacher effectiveness. *Journal of Teacher Education*, 27, 5-13.
- Berliner, D. & Tikunoff, W. (1976). The California Beginning Teacher Evaluation Study : Overview of the ethnographic study. *Journal of Teacher Education*, 27(1), 214-30.
- Berliner, D. & Tikunoff, W. (1977). Ethnography in the classroom. In G. Borich and K. Fenton (Ed.) *The appraisal of teaching : Concepts and process*. Reading, MA : Addison-Wesley.
- Bettencourt, E., Gillett, M., Gall, M. & Hull, R. (1983). Effects of teacher enthusiasm training on student on-task behavior and achievement. *American Research Journal*, 20, 435-450.

- Bissonnette, S. Richard, M. & Gauthier, C. (2005). Interventions pédagogiques efficaces et réussite scolaire des élèves provenant de milieux défavorisés. *Revue française de pédagogie*, 150, 87-141.
- Bloom, B.S. (1980). *The new direction in educational research : alterable variables*. Delat Kappa, 61, (6), 382-385.
- Borg, W. (1979). Teacher coverage of academic content and pupil achievement. *Journal of Educational Psychology*, 71, 635-645.
- Borg, W. (1980). Time and school learning. In C. Denham and A. Lieberman (Ed.), *Time to learn*. Washington, DC : National Institute of Education.
- Bressoux, P. (1994). Les recherches sur les effets-écoles et les effets-maîtres. *Revue française de pédagogie*, 108, 91-137.
- Bressoux, P. & Dessus, Ph. (2003). Les stratégies de l'enseignant en situation d'interaction. In M. Kail & M. Fayol (Ed.). *Les sciences cognitives et l'école. La question des apprentissages* (pp. 221-257). Paris : PUF.
- Brophy, J. (1973). Stability of teacher effectiveness. *American Educational Research Journal*, 10, 245-252.
- Brophy, J. (1979). Teacher behavior and its effects. *Journal of Educational Psychology*, 71, 733-750.
- Brophy, J. (1981). Teacher praise : A functional analysis. *Review of Educational Research*, 51, 5-32.

- Brophy, J. & Evertson, C. (1976). *Learning from teaching: A Developmental perspective*. Boston : Allyn and Bacon.
- Brophy, J. & Evertson, C. (1978). Content variables in teaching. *Educational Psychologist*, 12, 310-316.
- Brophy, J., Good, T. (1986). *Teacher behavior and student achievement*. In *Handbook of research on teaching*. A project of the American Educational Research Association (pp. 328-375). Wittrock : Merlin.
- Charbonnel, N. (1986), *Pour une critique de la raison éducative*. Bern : Peter Lang.
- Claparède, E. (1912), *Pourquoi les sciences de l'éducation ?* Lausanne et Genève : Payot.
- Claparède, E. (1946). *Psychologie de l'enfant et pédagogie expérimentale*. Neuchâtel : Delachaux et Niestlé.
- Clark, C., Gage, N., Marx, R., Peterson, P., Stayrock, N. & Winne, P. (1979). A factorial experiment on teacher structuring, soliciting, and reacting. *Journal of Educational Psychology*, 71, 534-552.
- Coleman, J., Campbell, E., Hobson, C., Mc Partland, J., Mood, A., Weinfield, F. & York, R. (1966). *Equality of educational opportunity*. Washington, DC : US. Office of Health, Education, and Welfare.
- Comber, L., & Keeves, J. (1973). *Science education in nineteen countries*. New York : Halsted Press.
- Cooley, W. & Leinhardt, G. (1980). The Instructional Dimensions Study. *Educational Evaluation and Policy Analysis*, 2, 7-25.

- Crahay, M. (1988). Contraintes de situation et interactions maître-élève : changer sa façon d'enseigner, est-ce possible ? *Revue Française de Pédagogie*, 88, 67-94.
- Crahay, M. & Delhaxhe, A. (1991). How do preschool teachers' requests influence children's behaviors? *Teaching and Teacher Education*, 7 (3), 221-229.
- Crahay, M. (1996). Recherche en éducation et réflexion de l'action éducative. In Ch. Hadji & J. Baillé (Ed.), *Recherche en éducation. Vers une nouvelle « alliance »* (pp. 125-160). Bruxelles : De Boeck.
- Crahay, M. (2000). *L'école peut-elle être juste et efficace ?* Bruxelles : De Boeck.
- Crahay, M., & Monseur, C. (2006). Différences individuelles et effets d'agrégation en ce qui concerne les performances en lecture. Analyse secondaire des données PISA 2000. In C. Houssemand, R. Martin & P. Dickes. *Perspectives de psychologie différentielle* (pp. 23-34). Rennes : Presses Universitaires de Rennes.
- de Landsheere, G. (1969). *Comment les maîtres enseignent*. Bruxelles : Direction Générale de l'Organisation des Etudes, Pédagogie et recherche, n°1.
- de Landsheere, G. (1974). Analysis of Verbal Interaction in the Classroom. In J. Chanan (Ed.), *Towards a Science of Teaching*. London : NFER.
- de Landsheere (1978). L'évaluation des enseignants. In M. Debesse et G. Mialaret (Ed.), *Traité des sciences pédagogiques*, Vol. 7, Paris : PUF.

- de Landsheere, G. & Delchambre, A. (1979). *Les comportements non verbaux de l'enseignant. Comment les maîtres enseignent, II*. Paris : Nathan.
- Doyle, W. & Ponder, G.A. (1975). Classroom Ecology : Some Concerns About a Neglected Dimension of Research on Teaching. *Contemporary Education*, 46, 183-188.
- Doyle, W. (1977). Paradigms for research on teacher effectiveness, *Review of Research in Education*, 5, 163-198.
- Doyle, W. (1978). Paradigms for Research on Teacher Effectiveness. In L.S. Shulman (Ed.) *Review of Research in Education* (vol. 5). Itasca : Peacock.
- Doyle, W. (1979a). Making managerial decisions in classrooms. In D.L. Duke (Ed.), *Classroom management* (Yearbook of the Society for the Study of Education). Chicago : University of Chicago Press.
- Doyle, W. (1979b). Classroom task and students' abilities. In P. Peterson and H. Walberg (Ed.), *Research on teaching: Concepts, findings, and implications*. Berkeley, CA : Mc Cutchan.
- Doyle, W. (1983). Academic work. *Review of Educational Research*, 53, 159-199.
- Doyle, W. (1986). Paradigmes de recherches sur l'efficacité des enseignants. In M. Crahay et D. Lafontaine (Ed.), *L'art et la science de l'enseignement* (pp. 435-481). Bruxelles : Labor.
- Dunkin, M. & Biddle, B. (1974). *The study of teaching*. New York : Holt, Rinehart and Winston.

- Dunkin, M.J. (1986). Concepts et modèle dans l'analyse des processus d'enseignement. In M. Crahay & D. Lafontaine (Ed.), *L'art et la Science de l'enseignement* (pp. 39-80). Bruxelles : Labor.
- Ebmeier, H. & Good, T. (1979). The effects of instructing teachers about good teaching on the mathematics achievement of fourth grade students. *American Educational Research Journal*, 16, 1-16.
- Emmer, E., Evertson, C. & Anderson, L. (1980). Effective classroom management at the beginning of the school year. *Elementary School Journal*, 80, 219-231.
- Evertson, C., Anderson, C., Anderson, L. & Brophy, J. (1980). Relationships between classroom behaviors and student outcomes in junior high Mathematics and English classes. *American Educational Research Journal*, 17, 43-60.
- Evertson, C. & Emmer, E. (1982). Effective management at the beginning of the school year in junior high classes. *Journal of Educational Psychology*, 74, 485-498.
- Evertson, C., Emmer E. & Brophy, J. (1980). Predictors of effective teaching in junior high mathematics classrooms. *Journal for Research in mathematics Education*, 11, 167-178.
- Fenstermacher, G.D. (1980). On Learning to Teach Effectively from Research on Teacher Effectiveness. In C. Denham & A. Lieberman (Ed.), *Time to learn*. Washington D.C. : National Institute of Education.

- Flanders, N. (1970). *Analysing teacher behavior*. Reading, MA : Addison-Wesley.
- Flanders, N. & Simon, A. (1969). Teacher effectiveness. In R. Ebel (Ed.), *Encyclopedia of Educational Research* (4th ed.). New York : Mac Millan.
- Fortune, J. (1967). *A study of the generality of presenting behaviors in teaching preschool children* (Final Report for U.S. Office of Education Project No. 6-4868). Memphis, TN: Memphis State University (ERIC Document Reproduction Service N° ED016 285).
- Gage, N.L (1963). Paradigms for research on teaching. In N.L. Gage (Ed.), *Handbook of Research on teaching*. American Educational Research Association. Chicago : Rand Mc Nally.
- Gage, N.L. (1965). Desirable behaviors of teachers, *Urban Education*, 1, 85-95.
- Gage, N.L. (Ed.). (1976), *The Psychology of Teaching Methods*, Seventy-fifth Yearbook of the National Society for the Study of Education (pp. 335-371, Part 1.), Chicago : University of Chicago Press.
- Gage, N. L. (1978). *The scientific basis of the art of teaching*. New York : Teachers College Press, Columbia University.
- Gage, N. L. (1979). The generality of dimensions of teaching. In P. Peterson and H. Walberg (Ed.), *Research on teaching : Concepts, findings, and implications*. Berkeley, Ca : Mc Cutchan.

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

- Gage, N. L. (1983). When does research on teaching yield implications for practice ? *Elementary School Journal*, 83, 492-496.
- Gage, N.G. (1986). Comment tirer un meilleur parti des recherches sur les processus d'enseignement ? In M. Crahay & D. Lafontaine (Ed.), *L'art et la Science de l'enseignement*. Bruxelles : Labor.
- Gauthier, C. (2005). Interventions pédagogiques efficaces et réussite scolaire des élèves provenant de milieux défavorisés. In S. Bissonnette, M. Richard & C. Gauthier, *Revue française de pédagogie*, 150, 87-141.
- Getzels, J.W. & Jackson, P.W. (1963). The teacher's personality and characteristics. In N.L. Gage (Ed.), *Handbook of research on teaching*. Chicago : Rand Mc Nally.
- Glaser, R. (1986). Enseigner comment penser. In M. Crahay & D. Lafontaine (Ed.), *L'art et la science de l'enseignement* (pp. 251-280). Bruxelles : Labor.
- Good, T. (1979). Teacher effectiveness in the elementary school: What we know about it now. *Journal of Teacher Education*, 30, 52-64.
- Good, T., Biddle , B. & Brophy, J. (1975). *Teachers make a difference*. New York: Holt, Rinehart and Winston.
- Good, T. & Brophy, J. (1984). *Looking in classrooms* (3rd ed.). New York : Harper and Row.
- Good, T. & Grouws, D. (1977). Teaching effects : a process-product study in fourth grade mathematics

- classrooms. *Journal of Teacher Education*, 28, 49-54.
- Good, T. & Grouws, D. (1979). The Missouri Mathematics Effectiveness Project : An experimental study in fourth grade classrooms. *Journal of Educational Psychology*, 71, 355-362.
- Good, T., Grouws, D. & Beckermn, T. (1978). Curriculum pacing : Some empirical data in mathematics. *Journal of Curriculum Studies*, 10, 75-81.
- Good, T., Grouws, D. & Ebmeier, M. (1983). *Active mathematics teaching*. New York : Longman.
- Gump, P. V. (1967). *The Classroom Behavior Setting: Its Nature and Relation to Student behavior*. (Rapport de recherche n° OE-4-10-107). US : Bureau of Research, H.E.W.
- Harris, A. & Server, B. (1966). The CRAFT Project: Instructional time in reading research. *Reading Research Quarterly*, 2, 27-57.
- Hedges, L.V., Giacomia, R.M. & Gage, N.L. (1981). *Meta-analysis of the effects of Open and Traditional Instruction*. Final Report. Stanford University Program on Teaching Effectiveness, Meta-analysis project, Vol. 2, Stanford, California.
- Hughes, D. (1973). An experimental investigation of the effects of pupil responding and teacher reacting on pupil achievement. *American Educational Research Journal*, 10, 21-37.
- Hume, D. (1988), *A treatise on human nature*. Oxford : Selby-Bigge.

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

- Jackson, P.W. (1966). *The Way Teaching Is*. In P.W. Jackson (Ed.). *The Way Teaching Is : A Report for the Learner on Teaching*. Washington, D.C. : Association for Supervision and Curriculum Development and the Center for the Study of Instruction of the National Education Association.
- Jackson, P.W. (1986). *The practice of teaching*. New York : Teachers College Press.
- Jonnaert, Ph. (2002). *Compétences et socioconstructivisme. Un cadre théorique*. Bruxelles : de Boeck.
- Karweit, N. & Slavin, R. (1982). Time-on-Task: Issues of timing, sampling, and definition. *Journal of Educational Psychology*, 74, 844-851.
- Kounin, J. (1970). *Discipline and group management in classrooms*. New York : Holt, Rinehart and Winston.
- Martin, J. (1979). Effects of teacher higher-order questions on student process and product variables in a single-classroom study. *Journal of Educational Research*, 72, 183-187.
- Mc Gaw,, B., Wardrop, J.L. & Bunda, M.A. (1972). Classroom observation schemes : where are the errors. *American educational Research Journal*, 9, 13-27.
- Mc Donald, F. (1976). Report on Phase II of the Beginning Teacher Evaluation Study. *Journal of Teacher Education*, 27 (1), 39-42.
- Mc Donald, F. (1977). Research on teaching: Report on Phase II of the Beginning Teacher Evaluation Study. In G. Borich and K. Fenton (Ed.), *The appraisal of*

- teaching: Concepts and process.* Reading, MA : Addison-Wesley.
- Mc Donald, F. & Elias P. (1976a). *The effects of teaching performance on pupil learning. Beginning Teacher Evaluation Study. Phase II, 1974-1976* (Final report, 5 vols.). Princeton, NJ : Educational Testing Service.
- Mc Donald, F. & Elias P. (1976b). *Executive summary report : Beginning Teacher Evaluation Study, Phase II.* Princeton, NJ : Educational Testing Service.
- Medley, D. (1977). *Teacher competency and teacher effectiveness: A review of process-product research.* Washington DC : American Association of Colleges for teacher Education.
- Medley, D. (1979). The effectiveness of teachers. In P. Peterson and H. Walberg (Ed.). *Research on teaching: Concepts, findings, and implications.* Berkeley, Ca : Mc Cutchan.
- Medley, D.M. (1988). Evolution of research on teaching. In M.J. Dunkin (Ed.), *The international encyclopedia of teaching and teacher education.* Oxford : Pergamon Press.
- Mitzel, J.W. & Jackson, P.W. (1963). The Teacher's Personality and characteristics. In N.L. Gage (Ed.) *Handbook of research on teaching* (pp. 506-582). Chicago : Rand Mac Nally.
- Morine-Dershimer, G. (1982). Pupil perceptions of teacher praise. *Elementary School Journal*, 82, 421-434.

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

- Nuthall, G.A. (1968). Studies of Teaching II : Types of Research on Teaching. *N.Z. J. of Educ. Studies*, 3 (2), 125-247.
- Nuthall, G. & Church, J. (1973). Experimental studies of teaching behaviour. In G. Chanan (Ed.), *Towards a science of teaching*. London : National Foundation for Educational Research.
- Papert, S. (1981), *Jaillissement de l'esprit : ordinateur et apprentissage*, Paris : Flammarion.
- Peterson, P. (1979). Direct instruction reconsidered. In P. Peterson and H. Walberg (Ed.), *Research on teaching : Concepts, findings, and implications*. Berkeley, Ca : Mc Cutchan.
- Peterson, P. (1979). *Research on teaching: Concepts, findings, and implications*. Berkeley, Ca : Mc Cutchan.
- Piaget, J. (1969). *Psychologie et pédagogie*. Paris : Denoël-Gonthier.
- Popham, W. (1971). Performance tests of teaching proficiency : Rationale, development, and validation. *American Educational Research Journal*, 8, 105-117.
- Popper, K.R. (1982), *La connaissance objective*, Paris : Editions Complexes.
- Powell, M. (1980). The Beginning Teacher Evaluation Study: A brief history of a major research project. In C. Denham and A. Lieberman (Ed.), *Time to learn*. Washington, DC : National institute of Education.
- Rosenshine, B. (1970a). Evaluation of instruction, *Review of Educational Research*, 40, 279-301.

- Rosenshine, B. (1970b). Experimental classroom studies of indirect teaching. *Classroom Interaction newsletter*, 5 (2), 7-11.
- Rosenshine, B. (1971). *Teaching behaviours and student achievement*. London : National Foundation for Educational Research.
- Rosenshine, B. (1979). Content, time, and direct instruction. In P. Peterson and H. Walberg (Ed.), *Research on teaching : Concepts, findings, and implications*. Berkeley, Ca : Mc Cutchan.
- Rosenshine, B. & Furst, N. (1973). The use of direct observation to study teaching. In R.M.W. Travers (Ed.), *Second Handbook of Research on Teaching*. Chicago : Rand Mc Nally.
- Rosenshine, B.V. (1986). Vers un enseignement efficace des matières structurées. In M. Crahay & D. Lafontaine (Ed.), *L'art et la science de l'enseignement* (pp. 81-97). Bruxelles : Labor.
- Rothkopf, E.Z. (1970). The concept of mathemagenic activities. *Review of Educational Research*, 40, 325-336.
- Rothkopf, E.Z. (1976). Writing to teach and reading to learn : A perspective on the psychology of written instruction. In N.L. Gage (Ed.), *The psychology of teaching methods*. 75th Yearbook of the National Society for the Study of Education, Pt. 1. Chicago : University of Chicago Press.

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

- Rowley, G.L. (1976). The reliability of observational measures. *American educational Research Journal*, 13, 1, 51-59.
- Schneuwly, B. (1997). Psychologie et pédagogie. Le paradigme psychopédagogique et son contraire. *Skholê*, 7, 27-45.
- Shulman, L.S. (1986). Paradigms and research programs in the study of teaching. A contemporary perspective. In M.C. Wittrock (Ed.), *The third handbook of the research on teaching*, 2-35. New York : Mac Millan.
- Smith, B.O. (1956). On the Anatomy of Teaching. *J. of Teacher Education*, 7, 339-346.
- Smith, B.O. (1960). A concept of teaching. *Teachers' College Record*, 5, 229-241.
- Smith, B.O. (1961). A Concept of Teaching. In B.O. Smith & R.H. Ennis (Ed.), *Language and Concepts in Education*. Chicago : Rand Mc Nally.
- Smith, B.O., Meux, M.O., Coombs, J., Nuthall, G.A. & Precians, R.P. (1967). *A Study of the Strategies of Teaching*. Urbana : Bureau of Educational Research, Université de l'Illinois.
- Soar, R.S. (1968). Optimum teacher-pupil interaction for pupil growth. *Educational Leadership*, 26, 275-280.
- Soar, R.S. (1973). *Follow Through classroom process measurement and pupil growth (1970-1971, final report)*. Gainesville : University of Florida, Institute for Development of Human Resource. (ERIC Document Reproduction Service N°ED 106 297).

MARCEL CRAHAY

- Soar, R.S. (1977). An integration of findings from four studies of teacher effectiveness. In G. Borich and K. Fenton (Ed.). *The appraisal of teaching: Concepts and process*. Reading, MA : Addison-Wesley.
- Soar, R.S. & Soar, R.S. (1972). An empirical analysis of selected follow through programs; An example of a process approach to evaluation. In I. Gordon (Ed.), *Early Childhood Education*. Chicago : National Society for the Study of Education.
- Soar, R.S. & Soar, R.S. (1979). Emotional climate and management. In P. Peterson and H. Walberg (Ed.), *Research on teaching : Concepts, findings, and implications*. Berkeley, Ca : Mc Cutchan.
- Stalling, J. (1975). Implementation and child effects of teaching practices in Follow Through classrooms. *Monographs of the Society for Research in Child Development*, 40 (7-8) n°163.
- Stalling, J. (1980). Allocated academic learning time revisited, or beyond time on task. *Educational Researcher*, 8 (11), 11-16.
- Stalling, J., Cory, R., Fairweather, J. & Needels, M. (1977). *Early Childhood Education classroom evaluation*. Menlo Park, Ca : SRI International.
- Stalling, J., Cory, R., Fairweather, J. & Needels, M. (1978). *A study of basic reading skills taught in secondary schools*. Menlo Park, Ca : SRI International.
- Stalling, J. & Kaskowitz, D. (1974). *Follow Through Classroom Observation Evaluation 1972-1973* (SRI

UN BILAN DES RECHERCHES PROCESSUS-PRODUIT

- Project URU-7370). Stanford, Ca : Stanford Research Institute.
- Stalling, J., Needels, M. & Stayrook, N. (1979). *The teaching of basic reading skills taught in secondary schools. Phase II and Phase III*. Menlo Park, Ca : SRI International.
- Tikunoff, W., Berliner, D. & Rist, R. (1975). *An ethnographic study of the forty classrooms of the Beginning Teacher Evaluation Study known sample* (Tech. Rep. No. 75-10-5). San Francisco : Far West Laboratory.
- Turner, R.L. (1971). Conceptual foundations of research in teacher education. In B.O. Smith (Ed.), *Research in teacher education : A symposium*. Englewood Cliffs, N.J. : Prentice-Hall.
- Vygotsky, L.S. (1985). Le problème de l'enseignement et du développement mental à l'âge scolaire. In B. Schneuwly & J.P. Bronckart (ed.), *Vygotsky aujourd'hui*. Neuchâtel : Delachaux et Niestlé.
- Wesbury, I. (1967). The reliability of measures of classroom behavior. *Ontario Journal of Educational Research*, 10 (2), 125-138.
- Wright, C. & Nuthall, G. (1970). Relationships between teacher behaviors and pupil achievement in three experimental elementary science lessons. *American Educational Research Journal*, 7, 477-491.

Carnets des sciences de l'éducation

La collection des Carnets des sciences de l'éducation est une édition de la Section des sciences de l'éducation, produite par le Groupe Publications, en lien avec le comité de rédaction de la collection Raisons éducatives.

La collection des Carnets des sciences de l'éducation se propose de publier des textes scientifiques propres aux divers champs disciplinaires et thématiques des sciences de l'éducation. Ces textes, toujours signés par un·e enseignant·e de la Section, sont articulés aux enseignements dispensés dans le cadre des différents cursus de formation, dont ils préparent ou prolongent l'étude.

A l'instar des Cahiers de la Section des sciences de l'éducation et de la collection Raisons éducatives, la collection des Carnets tient aussi à promouvoir et entretenir une culture de la lecture et de la recherche.