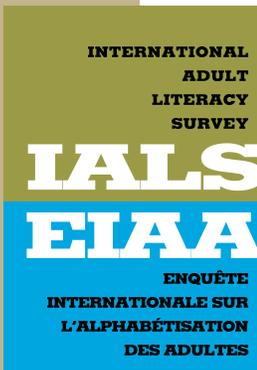


Littératie, affectation professionnelle et rendement de la surinstruction et de la sous-instruction

Daniel Boothby



Comment obtenir d'autres renseignements

Toute demande de renseignements au sujet du présent produit ou au sujet de statistiques ou de services connexes doit être adressée à : Yvan Clemont, Statistique Canada, Ottawa, Ontario, K1A 0T6 (téléphone : (613) 951-3326).

Pour obtenir des renseignements sur l'ensemble des données de Statistique Canada qui sont disponibles, veuillez composer l'un des numéros sans frais suivants. Vous pouvez également communiquer avec nous par courriel ou visiter notre site Web.

Service national de renseignements	1 800 263-1136
Service national d'appareils de télécommunications pour les malentendants	1 800 363-7629
Renseignements concernant le Programme des bibliothèques de dépôt	1 800 700-1033
Télécopieur pour le Programme des bibliothèques de dépôt	1 800 889-9734
Renseignements par courriel	infostats@statcan.ca
Site Web	www.statcan.ca

Renseignements sur les commandes et les abonnements

Le produit n° 89-552-MPF au catalogue est publié (irégulier) en version imprimée standard et est offert au prix de 10 \$CAN l'exemplaire. Les frais de livraison supplémentaires suivants s'appliquent aux envois à l'extérieur du Canada :

Exemplaire

États-Unis	6 \$CAN
Autres pays	10 \$CAN

Ce produit est aussi disponible sous forme électronique dans le site Internet de Statistique Canada, sous le n° 89-552-MIF au catalogue, gratuitement. Les utilisateurs peuvent obtenir des exemplaires ou s'abonner en visitant notre site Web à www.statcan.ca et en choisissant la rubrique Produits et services.

Les prix ne comprennent pas les taxes de ventes.

La version imprimée peut être commandée par

- Téléphone (Canada et États-Unis) **1 800 267-6677**
- Télécopieur (Canada et États-Unis) **1 877 287-4369**
- Courriel **order@statcan.ca**
- Poste
Statistique Canada
Division de la diffusion
Gestion de la circulation
120, avenue Parkdale
Ottawa (Ontario) K1A 0T6
- En personne au bureau régional de Statistique Canada le plus près de votre localité ou auprès des agents et librairies autorisés.

Lorsque vous signalez un changement d'adresse, veuillez nous fournir l'ancienne et la nouvelle adresse.

Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois, et ce, dans la langue officielle de leur choix. À cet égard, notre organisme s'est doté de normes de service à la clientèle qui doivent être observées par les employés lorsqu'ils offrent des services à la clientèle. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1 800 263-1136.

Le papier utilisé dans la présente publication répond aux exigences minimales de l'American National Standard for Information Sciences — Permanence of Paper for Printed Library Materials, ANSI Z39.48 — 1984.



Enquête internationale sur l'alphabétisation des adultes

Littératie, affectation professionnelle et rendement de la surinstruction et de la sous-instruction

Daniel Boothby

Développement des ressources humaines Canada

L'Enquête internationale sur l'alphabétisation des adultes (EIAA) a été menée dans 22 pays entre 1994 et 1998. Le volet canadien de l'EIAA, réalisé à l'automne de 1994, était financé principalement par la Direction générale de la recherche appliquée et le Secrétariat national à l'alphabétisation de Développement des ressources humaines Canada.

Publication autorisée par le ministre responsable de Statistique Canada

© Ministre de l'Industrie, 2002

Tous droits réservés. Il est interdit de reproduire ou de transmettre le contenu de la présente publication, sous quelque forme ou par quelque moyen que ce soit, enregistrement sur support magnétique, reproduction électronique, mécanique, photographique, ou autre, ou de l'emmagasiner dans un système de recouvrement, sans l'autorisation écrite préalable des Services de concession des droits de licence, Division du marketing, Statistique Canada, Ottawa, Ontario, Canada, K1A 0T6.

Janvier 2002

N° 89-552-MPF au catalogue, n° 9

ISSN 1482-7476

ISBN 0-660-96653-0

N° 89-552-MIF au catalogue, n° 9

ISSN 1480-9524

ISBN 0-662-86668-1

Périodicité : irrégulière

Ottawa

Statistique Canada

Développement des ressources humaines Canada

Nous désirons toutefois souligner que l'interprétation des données et les recommandations générales exposées dans le présent rapport sont propres à l'auteur et ne reflètent pas nécessairement celles des organismes subventionnaires ou des évaluateurs.

Données de catalogage avant publication de la Bibliothèque nationale du Canada

Boothby, Daniel W.

Littératie, affectation professionnelle et rendement de la surinstruction
et de la sous-instruction.

(Enquête internationale sur l'alphabétisation des adultes)

Publié aussi en anglais sous le titre: Literacy skills, occupational assignment and the
returns to over- and under education.

ISBN 0-660-96653-0 (papier)

ISBN 0-662-86668-1 (Internet)

CS89-552-MPF no 9

CS89-552-MIF no 9

1. Alphabétisation – Canada – Statistiques.
 2. Marché du travail, Effets de l'éducation sur les – Canada – Statistiques.
 3. Éducation – Canada – Statistiques.
 4. Salaires, Effets de l'éducation sur les – Canada – Statistiques.
- I. Statistique Canada. II. Canada. Développement des ressources
humaines Canada. III. Titre: Littératie, affectation professionnelle et
rendement de la surinstruction et de la sous-instruction. VI. Collection.

HD5728 B6614 2002

331.11/423/0971021

C2002-988001-7

Remerciements

La présente étude est réalisée en vertu d'un contrat avec Développement des ressources humaines Canada; les tableaux sont produits en vertu d'un contrat avec Statistique Canada. Les opinions exprimées dans le présent document sont celles de l'auteur et ne reflètent pas nécessairement le point de vue de Développement des ressources humaines Canada (DRHC) ni de Statistique Canada ni du gouvernement fédéral.

L'auteur tient à remercier Gilles Bérubé, de Développement des ressources humaines Canada (DRHC), pour son soutien à ce projet et pour ses observations utiles à l'égard des versions antérieures. Benoît Delage, de DRHC, a également formulé des observations utiles à l'égard des versions antérieures, tout comme Scott Murray et Marie Drolet, de Statistique Canada. L'auteur tient également à remercier Yves Bélanger, de Statistique Canada, pour ses conseils techniques concernant l'EIAA.

L'auteur reste seul responsable de toute erreur de fait ou d'interprétation.

Note de reconnaissance

Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population, les entreprises, les administrations canadiennes et les autres organismes. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques précises et actuelles.

PUBLICATIONS ÉLECTRONIQUES DISPONIBLES À
www.statcan.ca



Table des matières

Résumé	7
Chapitre 1	Introduction et survol de la documentation 9
	Rendement de la surinstruction et de la sous-instruction 10
	Surinstruction, sous-instruction et littératie 11
	Les différences liées aux compétences en littératie expliquent-elles le rendement de la surinstruction et de la sous-instruction? 12
Chapitre 2	Sources des données et mesures de la scolarité requise, de la surinstruction et de la sous-instruction 15
	Source des données 15
	Mesures de la scolarité requise, de la surinstruction et de la sous-instruction 16
	Années de scolarité et durée de la formation nécessaire (en années) 16
	Surinstruction ou sous-instruction? 17
Chapitre 3	Déterminants de la concordance avec la scolarité requise 19
	Littératie et concordance avec la scolarité requise 19
	Estimations des déterminants de la scolarité requise 20
Chapitre 4	Surinstruction, sous-instruction, littératie, utilisation des compétences en littératie et gains 25
	Résultats des régressions 25
	Gains, littératie, scolarité et jumelage emploi-travailleur 30
Conclusions	33
Bibliographie	37
Annexe 1	Restrictions applicables aux échantillons et effets sur la taille de l'échantillon 39
Annexe 2	Questions concernant la forme fonctionnelle de l'équation des gains 41
	Non-linéarité du rendement de la scolarité 41
	Biais dans la mesure de la surinstruction et de la sous-instruction 42
Annexe 3	La relation entre les études, la littératie et les notes en littératie 45
	Construction des « valeurs plausibles » dans l'EIAA 45
	Estimation des effets du conditionnement 47
	Quelques conclusions 50
Notes	53

Table des matières

Tableau 1	Niveaux moyens d'années de scolarité, de FG et de FG+PPS selon le niveau de scolarité le plus élevé et selon le sexe	17
Tableau 2	Régressions de référence relatives aux déterminants des exigences scolaires de la profession	21
Tableau 3	Effets du nombre d'années de scolarité sur les exigences scolaires de la profession	22
Tableau 4	Régressions de référence du logarithme des gains annuels moyens	26
Tableau 5	Effets de la scolarité, de la surinstruction et de la sous-instruction, des compétences en littératie et de l'utilisation de la littératie sur le logarithme des gains annuels	28
Tableau A.1	Corrélations entre note en littératie, proportion de bonnes réponses et nombre d'années de scolarité	48
Tableau A.2	Estimations de régression de la note en littératie et de la proportion de bonnes réponses	48
Tableau A.3	Estimations de régression relatives aux déterminants de la note en littératie	50

Résumé

La présente étude utilise les données du volet canadien de l'Enquête internationale sur l'alphabétisation des adultes pour examiner les relations entre la scolarité, la littératie et l'affectation professionnelle et pour déterminer dans quelle mesure le rendement de la surinstruction et de la sous-instruction constitue en réalité le rendement de la littératie.

Nous employons deux mesures de la durée de la formation nécessaire à l'emploi, toutes deux fondées sur le tableau détaillé des professions. L'une est le niveau de formation générale (FG) de la profession; l'autre est la somme des durées de la FG et de la préparation professionnelle spécifique (PPS).

L'analyse de régression de la durée de la formation nécessaire à l'emploi montre que la littératie constitue un déterminant important de l'affectation professionnelle par niveau de compétence, compte tenu de la scolarité. Les compétences acquises au moyen de la formation en cours d'emploi peuvent aussi jouer un rôle important dans l'affectation professionnelle.

Les travaux de recherche sur le rendement de la surinstruction et de la sous-instruction analysent la relation entre les compétences des travailleurs, mesurées d'après leur niveau de scolarité, et les exigences professionnelles de leur emploi. Les constatations habituelles de ces travaux de recherche sont les suivantes : 1) les travailleurs surinstruits (scolarité supérieure à celle qu'exige leur emploi) gagnent plus que ceux qui occupent un emploi aux exigences scolaires comparables, mais dont le niveau de scolarité (inférieur) correspond à ces exigences; 2) les travailleurs surinstruits gagnent moins que ceux dont la scolarité est comparable dans un emploi qui exige ce niveau de scolarité; et 3) les travailleurs sous-instruits (scolarité inférieure à celle qu'exige leur emploi) gagnent plus que ceux dont le niveau d'instruction est comparable et dont l'emploi correspond à leur scolarité, mais moins que les travailleurs qui occupent un emploi aux exigences scolaires comparables et dont la scolarité correspond à ces exigences.

Nous observons cette tendance du rendement de la surinstruction et de la sous-instruction chez les femmes et chez les hommes de notre échantillon en effectuant une analyse de régression (du logarithme) des gains des travailleurs à temps plein. Lorsqu'on ajoute à ces régressions des mesures des compétences en littératie, les coefficients estimatifs de la surinstruction et de la sous-instruction diminuent en valeur absolue chez les hommes et ceux de la sous-instruction augmentent chez les femmes. Lorsqu'on ajoute une mesure de l'utilisation des compétences en littératie au travail, cette variable a un coefficient positif et les valeurs absolues des coefficients de la surinstruction et de la sous-instruction diminuent encore, tant chez les femmes que chez les hommes.

Nous concluons que les compétences en littératie jouent un rôle important dans l'affectation professionnelle, indépendamment de celui de la scolarité; que le rendement de la sous-instruction, tant chez les femmes que chez les hommes, constitue dans une large mesure le rendement de compétences en littératie supérieures à la moyenne pour leur niveau de scolarité; et que chez les hommes, le rendement de la surinstruction constitue dans une large mesure le rendement de compétences en littératie supérieures à la moyenne pour leur emploi. Ces conclusions semblent indiquer que les employeurs sont en mesure de déterminer les compétences en littératie de leurs employés de manière plus exacte qu'en se fondant simplement sur leur niveau de scolarité.

Chapitre 1

Introduction et survol de la documentation

Les questions concernant les compétences—en quoi elles consistent, comment les mesurer, comment on les acquiert, comment elles sont rétribuées—revêtent une grande importance pour la recherche sur le marché du travail axée sur l’élaboration de politiques. Dans le présent document, nous nous penchons sur une question empirique, à savoir si le rendement observé de la « surinstruction » et de la « sous-instruction » constitue en réalité le rendement des compétences en littératie. Comme nous le mentionnons plus loin, la réponse à cette question empirique a des répercussions sur notre conception des compétences.

Les travaux de recherche dont nous faisons état portent sur divers aspects des compétences. Ainsi que le laisse entendre le titre, le premier de ces aspects concerne la « surinstruction », la « sous-instruction » et leur rendement respectif.

La notion de « surinstruction » découle de l’observation selon laquelle un nombre important de personnes travaillent dans un emploi dont les exigences scolaires sont inférieures au niveau d’études des titulaires. Selon une version populaire de cette observation, véhiculée par les journaux, de nombreux diplômés universitaires travaillent comme chauffeurs de taxi.

Parmi les auteurs universitaires, David Livingstone (1999) a publié un ouvrage intitulé *The Education-Jobs Gap*. Il y présente une série chronologique de données comparables montrant que, selon une enquête menée en Ontario à intervalles de deux ans, de 18 % à 23 % des travailleurs ontariens occupent un emploi où ils sont « sous-employés » et de 22 % à 26 % occupent un emploi où ils sont « sous-qualifiés » (tableau 2.6, p. 76). Des données comparables provenant de l’Enquête sociale générale semblent également indiquer qu’environ 20 % des travailleurs sont « sous-employés ». À partir de données d’auto-évaluation provenant de la *National Survey of Class Structure and Labour Process* menée au Canada, Vahey (2000) constate qu’en 1982, 30 % des travailleurs et 32 % des travailleuses étaient surinstruits, alors que 24 % des hommes et 17 % des femmes étaient sous-instruits. Il précise que dans presque tous les cas, le défaut de complémentarité correspondait à un niveau de scolarité; par exemple, des diplômés d’un collège communautaire travaillaient dans un emploi exigeant, à leur avis, un diplôme d’études secondaires.

Ce phénomène a de quoi préoccuper le public et les décideurs à cause de la perception selon laquelle la « surinstruction » constitue un gaspillage des ressources privées et sociales consacrées à l'instruction. Cette notion voulant qu'environ 20 % des travailleurs « surinstruits » occupent un emploi où leur instruction est gaspillée repose sur les hypothèses suivantes : à chaque emploi correspond un niveau d'exigences scolaires; ce niveau est mesuré avec exactitude; les personnes très instruites ne peuvent offrir un rendement supérieur à celui que supposent les exigences scolaires de l'emploi; le niveau d'instruction constitue une mesure exacte de la capacité d'une personne de répondre aux exigences professionnelles de l'emploi. Si l'on adhère à cette conception du marché du travail, on doit se préoccuper autant de la « sous-instruction » que de la « surinstruction », puisqu'il semble exister une proportion considérable de la population active dont les niveaux de compétence ne correspondent pas à ceux qu'exige l'emploi.

Rendement de la surinstruction et de la sous-instruction

Les résultats des travaux de recherche sur le « rendement de la surinstruction » remettent en question la notion selon laquelle la « surinstruction » constitue un gaspillage. Ces travaux aboutissent habituellement aux constatations suivantes :

1. les travailleurs « surinstruits » gagnent plus que ceux qui occupent un emploi aux exigences scolaires comparables, mais dont le niveau de scolarité (inférieur) correspond à ces exigences;
2. les travailleurs « surinstruits » gagnent moins que ceux dont la scolarité est comparable dans des emplois qui exigent ce niveau de scolarité;
3. les travailleurs « sous-instruits » gagnent plus que ceux dont le niveau d'instruction est comparable et dont l'emploi correspond à leur scolarité, mais moins que les travailleurs qui occupent un emploi aux exigences scolaires comparables et dont la scolarité correspond à ces exigences.

De manière plus formelle, à l'exemple de Sicherman (1991), supposons que S représente le nombre d'années de scolarité d'une personne et R , le nombre d'années de scolarité requises dans l'emploi de cette personne. Le nombre d'années de surinstruction O est défini comme suit : $O = \{S-R \text{ lorsqu'il est positif, } 0 \text{ dans le cas contraire}\}$. Le nombre d'années de sous-instruction U est défini comme suit : $U = \{R-S \text{ lorsqu'il est positif, } 0 \text{ dans le cas contraire}\}$. Dans une régression des gains qui comprend R , O et U comme variables explicatives, appelons respectivement r , o et u les coefficients estimatifs de ces variables. On peut alors formuler comme suit les trois énoncés du paragraphe précédent : 1) $r > o > 0$, 2) $u < 0$, et 3) $r + u > 0$ (de sorte que le rendement net d'une année de sous-instruction est positif).

Pour un échantillon de travailleurs provenant d'une étude américaine sur la dynamique du revenu, Sicherman (1991) établit par régression le logarithme du salaire horaire en fonction d'une série de variables explicatives types et de R , O et U . Les coefficients estimatifs sont les suivants : $r = 0,048$, $o = 0,039$, $u = -0,017$. Ils sont tous statistiquement significatifs. L'hypothèse selon laquelle $r = o$ est fortement rejetée par un test de Wald.

Dans un article intitulé « The Economics of Over- and Under-schooling », paru dans un récent numéro spécial de la *Economics of Education Review*, Hartog (2000) passe en revue une série d'études de ce type.

Il conclut (p. 135) que des constatations comme celles de Sicherman sont courantes et que cette spécification de l'équation des gains est encore plus valable lorsqu'on la confronte à la « spécification de Mincer (tous les coefficients sont égaux) »¹.

Dans un autre article paru dans le même numéro, Vahey (2000), à partir de données canadiennes de 1981 provenant de la *National Survey of Class Structure and Labour Process*,

observe le rendement positif de la surinstruction chez les hommes qui possèdent un baccalauréat et, à un degré moindre, une pénalité salariale chez les hommes sous-instruits qui occupent un emploi exigeant un niveau d'instruction peu élevé; il n'observe aucun rendement de la surinstruction ni de la sous-instruction chez les femmes². À notre connaissance, il s'agit de la seule étude, à part la nôtre, qui présente des estimations du rendement de la surinstruction et de la sous-instruction au Canada.

Surinstruction, sous-instruction et littératie

Pourquoi les travailleurs surinstruits tireraient-ils parti de leurs années de scolarité en sus des exigences de leur emploi si cette scolarité « excédentaire » ne contribuait pas à leur productivité? Pourquoi les travailleurs sous-instruits gagneraient-ils moins que les travailleurs suffisamment instruits dans des emplois dont les exigences scolaires sont semblables si les travailleurs sous-instruits n'étaient pas moins productifs?

Borghans et de Grip (2000) proposent un modèle dans lequel les travailleurs possédant différents niveaux d'instruction offrent différents niveaux de productivité dans le même emploi. Les travailleurs sont spécialisés dans un éventail de professions. Ce modèle tient compte de la possibilité de la surinstruction. Or, même en situation de surinstruction, l'instruction supplémentaire n'est pas entièrement gaspillée, car elle se traduit par une plus grande productivité dans l'emploi et par des gains plus élevés chez les travailleurs surinstruits que chez leurs collègues possédant le niveau d'instruction « approprié ».

Le modèle en question offre donc une explication de la tendance observée du rendement de la surinstruction. Les travailleurs surinstruits gagnent plus que leurs collègues parce qu'ils sont plus productifs. Ils gagnent moins que les travailleurs possédant un niveau d'instruction semblable et travaillant dans un emploi qui exige ce niveau d'instruction, car leur type de formation les oblige à travailler dans une profession où le niveau d'études est relativement excédentaire. Les auteurs opposent cet aspect au relèvement des exigences scolaires, dans lequel l'évolution de la technologie accroît la demande de travailleurs plus instruits au sein d'une profession.

Sicherman (1991) propose une conception de la surinstruction et de la sous-instruction qui suppose une substitution entre différentes méthodes d'acquisition du capital humain. Ses résultats empiriques appuient les hypothèses selon lesquelles, en moyenne, les travailleurs surinstruits ont moins d'expérience de travail, et les travailleurs sous-instruits en ont plus, que les travailleurs possédant le niveau de scolarité approprié pour leur emploi. Il soutient également l'hypothèse selon laquelle certains travailleurs surinstruits travaillent temporairement dans un emploi qui leur permet d'acquérir les compétences nécessaires à leur avancement professionnel. Enfin, il soutient la conception du jumelage emploi-travailleur selon laquelle, en raison du défaut de complémentarité emploi-compétences, les travailleurs surinstruits sont plutôt portés à quitter leur emploi. Dans la mesure où ces défauts de complémentarité sont temporaires, le « gaspillage » attribuable à la surinstruction est peu important.

Pryor et Scheffer (1997, 1999) préconisent une conception très différente de la surinstruction. Ils cherchent à expliquer pourquoi, aux États-Unis, des diplômés de l'enseignement universitaire occupent un emploi exigeant un diplôme d'études secondaires, alors qu'un écart croissant entre les salaires des travailleurs ayant fait des études secondaires et ceux qui ont fait des études universitaires semble justifier un accroissement de la demande relative de diplômés de l'enseignement universitaire. Les auteurs répondent que l'accroissement de la demande concerne les aptitudes cognitives, que les travailleurs ayant fait des études universitaires possèdent, en moyenne, des niveaux plus élevés d'aptitudes cognitives, mais que les diplômés de l'enseignement universitaire occupant un emploi qui exige un diplôme d'études secondaires possèdent des niveaux d'aptitudes cognitives

correspondant à ceux des diplômés de l'enseignement secondaire. Ils soutiennent leur point de vue à l'aide de données provenant de la *National Adult Literacy Survey*, qui montrent que les diplômés de l'enseignement universitaire occupant un emploi qui exige un diplôme d'études secondaires possèdent des niveaux moyens de compétences en littératie inférieurs à ceux des diplômés de l'enseignement secondaire qui occupent un emploi exigeant un diplôme d'études universitaires.

À l'aide des données de l'Enquête internationale sur l'alphabétisation des adultes (EIAA) de 1994, Green et Riddell (2001) montrent qu'au Canada, les compétences en littératie ont une incidence sur les gains qui compte pour environ le tiers de l'incidence totale des études sur les gains. Osberg (2000) explore la relation entre les gains, la scolarité et les notes en littératie en effectuant diverses transformations de la somme des trois notes en littératie provenant des données canadiennes de l'EIAA. Il conclut que, sans égard à la transformation effectuée, une proportion considérable du rendement des études constitue un rendement des compétences en littératie, surtout chez les hommes. Charette et Meng (1998) utilisent les données canadiennes de l'Enquête sur les capacités de lecture et d'écriture utilisées quotidiennement, menée en 1989. Ils constatent que lorsqu'on inclut des mesures des compétences en littératie et des capacités de calcul dans les équations du revenu, le rendement de la scolarité augmente chez les femmes et diminue chez les hommes³.

Boothby (1999) montre qu'au Canada, les faibles niveaux de compétences en littératie sont liés au défaut de complémentarité entre l'emploi et les études. À l'aide de données provenant de l'EIAA de 1994, il montre que chez les diplômés universitaires, le risque de « défaut de complémentarité » est inversement proportionnel aux compétences en littératie⁴. Il montre également que le « défaut de complémentarité professionnelle » a un effet négatif sur les gains des diplômés de l'enseignement postsecondaire.

Les différences liées aux compétences en littératie expliquent-elles le rendement de la surinstruction et de la sous-instruction?

Sans le mentionner explicitement, Pryor et Shaffer (1999) présentent des résultats qui pourraient expliquer le rendement de la « surinstruction » et de la « sous-instruction ». Premièrement, ils montrent (tableau 3.5, p. 66) que les niveaux moyens de compétences en littératie des diplômés de l'enseignement universitaire sont plus élevés dans les professions où les niveaux d'instruction sont élevés que dans celles où les niveaux d'instruction sont plus faibles. Toutefois, le même tableau montre qu'à l'intérieur de chaque niveau d'instruction des emplois, les notes moyennes en compétences en littératie augmentent avec le niveau de scolarité des titulaires de poste. Deuxièmement, les auteurs montrent (tableau 5.1, p. 104 et 105), pour tous les niveaux d'études (y compris chez les diplômés universitaires), une relation étroite entre le salaire et le niveau d'instruction correspondant à l'emploi. Le même tableau montre aussi qu'à chaque niveau d'instruction, les salaires augmentent en fonction du niveau d'études.

On peut en déduire qu'on est en présence du phénomène suivant : les travailleurs possédant un faible niveau de compétences en littératie pour leur niveau d'études peuvent convenir à des emplois dans lesquels le niveau d'instruction courant est inférieur au leur. Ils sont considérés comme surinstruits dans ces emplois. Toutefois, leur niveau de compétences en littératie peut être supérieur à celui des travailleurs possédant le niveau d'instruction « approprié » pour l'emploi. Pour le niveau d'instruction correspondant à l'emploi, les compétences en littératie supérieures à la moyenne se traduisent par des salaires moyens supérieurs à la moyenne. C'est ce que l'on considère comme le rendement de la surinstruction. Il est inférieur à celui de la scolarité requise, car les compétences en littératie de ces travailleurs sont inférieures à celles des personnes possédant le même niveau d'instruction et travaillant dans un emploi qui exige ce niveau d'instruction.

Inversement, les travailleurs sous-instruits sont ceux qui possèdent le niveau de compétences en littératie le plus élevé pour leur niveau d'instruction. Ils conviennent à des emplois dont les « exigences scolaires » sont supérieures à leur niveau de scolarité, mais possèdent des compétences en littératie inférieures à la moyenne pour ces emplois.

À l'instar de Pryor et Shaffer, le modèle que nous venons de décrire tient pour acquis que l'affectation professionnelle et les gains sont liés aux compétences en littératie. En principe, il ne se limite pas aux compétences en littératie. On pourrait construire le même type de modèle de la surinstruction et de la sous-instruction en fonction de l'aptitude à communiquer ou de toute autre compétence pouvant être corrélée avec la scolarité. Pour mettre le modèle à l'essai, nous devons nous limiter à tenir compte des compétences en littératie, car c'est pour ces compétences que l'on dispose de mesures directes.

Le modèle est semblable à celui de Borghans et de Grip (2000) (évoqué plus haut), du fait qu'il n'y a pas de niveau de compétence fixe dans une profession. Les travailleurs qui se situent à différents niveaux de compétence évoluent plutôt à différents niveaux de productivité au sein de la profession, et les gains varient en conséquence.

La partie empirique de notre étude consiste donc à utiliser les données de l'EIAA pour déterminer :

1. quelle est l'influence de la scolarité, des compétences en littératie et de l'expérience de travail des répondants sur les exigences professionnelles de leurs emplois,
2. si les données de l'EIAA montrent la tendance habituelle du rendement de la surinstruction et de la sous-instruction, et
3. dans quelle mesure le rendement des compétences en littératie explique celui de la surinstruction et de la sous-instruction.

Comme nous l'avons vu plus haut, notre étude s'attache beaucoup à la relation entre l'utilisation des compétences et le rendement de ces compétences. D'après l'ensemble de données de l'EIAA, Krahn et Lowe (1998) ont constaté qu'une proportion importante des travailleurs occupés possédaient un niveau élevé de compétences en littératie qu'ils n'utilisaient pas au travail. Green et coll. (2000) ont réparti les personnes occupées comprises dans l'échantillon britannique de l'EIAA par niveaux de compétences en littératie et par niveaux d'utilisation des compétences en littératie au travail. Ils ont constaté d'importantes pertes de gains chez les travailleurs possédant des compétences en littératie élevées dans des emplois faisant peu appel aux compétences en littératie. Nous examinerons donc également la relation entre les gains et l'utilisation des compétences en littératie au travail.

Voici le plan du document. Dans le chapitre 2, nous décrivons nos sources des données, notre mesure des exigences scolaires de l'emploi et les mesures de la surinstruction et de la sous-instruction qui en résultent. Dans le chapitre 3, nous présentons les résultats des estimations des déterminants des exigences scolaires de l'emploi des répondants. Dans le chapitre 4, nous présentons les résultats des estimations d'une série d'autres modèles du rendement des études, de la littératie, de la surinstruction, de la sous-instruction et de l'utilisation des compétences en littératie au travail. Nos conclusions suivent le chapitre 4.

PUBLICATIONS ÉLECTRONIQUES DISPONIBLES À
www.statcan.ca



Chapitre 2

Sources des données et mesures de la scolarité requise, de la surinstruction et de la sous-instruction

Source des données

Notre principale source de données est le fichier principal du volet canadien de l'EIAA de 1994. Ce fichier renferme des renseignements sur un ensemble standard de variables axées sur la scolarité, la démographie et le marché du travail à partir d'un échantillon national représentatif des Canadiens. Les répondants ont également écrit des tests de lecture. On a ainsi mesuré trois types de compétences en littératie : la compréhension de textes schématiques, de textes suivis et de textes au contenu quantitatif. Les données canadiennes de l'EIAA sont présentées dans Statistique Canada (sans date). On trouvera dans Murray et coll. (1998) la description la plus détaillée de l'EIAA.

Nous avons restreint notre échantillon aux non-étudiants, car la relation entre le niveau d'études et le niveau d'instruction exigé par l'emploi est habituellement très différente chez les étudiants et chez les non-étudiants. Nous avons limité l'échantillon aux Canadiens de naissance, car les résultats de Green et Riddell (2001) et de Kapsalis (2000) révèlent que les relations entre la scolarité, la littératie et les gains peuvent différer nettement entre les immigrants et les natifs. Nous n'avons retenu, pour notre échantillon, que les personnes occupées à temps plein, car les estimations préliminaires ont montré que les relations entre la scolarité, l'affectation professionnelle et les gains étaient très différentes chez les travailleurs à temps partiel. Nous attribuons ce phénomène à une forte concentration de travailleurs à temps partiel dans un petit nombre de professions où les exigences scolaires sont peu élevées, sans égard au niveau d'études de ces travailleurs. Nous n'avons retenu pour notre échantillon que les personnes ayant un code professionnel valide⁵.

Mesures de la scolarité requise, de la surinstruction et de la sous-instruction

Les données du fichier principal comprennent les codes professionnels détaillés correspondant aux emplois des personnes occupées. Le système de codification des professions utilisé est la Classification type des professions de 1980 (CTP 1980), ce qui nous permet d'intégrer à notre ensemble de données les estimations de la durée de la formation nécessaire pour les professions détaillées (groupe de base CTP 1980). Ces estimations de la durée de la formation sont fondées sur les exigences en matière de formation pour les professions très détaillées décrites dans la Classification canadienne descriptive des professions (CCDP) (Main-d'œuvre et Immigration, 1971 et 1973)⁶.

La CCDP fournit deux types d'exigences relatives à la durée de la formation nécessaire aux professions très détaillées (7 chiffres) : la formation générale (FG) et la préparation professionnelle spécifique (PPS). Au niveau des professions très détaillées, chacune de ces deux durées de formation a été traduite en années; puis, on a calculé la moyenne au sein du groupe de base pour obtenir des estimations de la durée de la formation nécessaire aux professions du groupe de base⁷. On a souvent utilisé des mesures de ce genre dans des études visant à cerner l'évolution des exigences professionnelles de l'économie. (Pour le Canada, par exemple, voir Hunter et Manley (1986), Myles (1988) et Boyd (1990).)

Années de scolarité et durée de la formation nécessaire (en années)

Nous avons utilisé deux mesures différentes de la durée de la formation nécessaire aux professions détaillées. La première est le niveau de FG de la profession, traduit en années. La seconde est la somme de la FG (en années) et de la PPS en années pour la profession. Le tableau 1 donne, par niveau de scolarité le plus élevé atteint par les femmes et par les hommes occupés, le nombre moyen d'années de scolarité et le nombre moyen d'années de scolarité requise pour l'emploi. Les niveaux moyens des deux mesures de la durée de la formation nécessaire (FG et FG+PPS) sont présentés dans le tableau 1.

Chacune de ces mesures présente des inconvénients. La mesure FG concerne uniquement les études générales, sans tenir compte de la formation professionnelle reçue dans des établissements d'enseignement. Elle a donc tendance à sous-estimer les exigences scolaires pour les professions, surtout celles qui exigent un niveau élevé de formation spécialisée dans des établissements d'enseignement. La PPS concerne le temps consacré à la formation nécessaire à certaines professions dans des établissements d'enseignement. Cette mesure présente une difficulté : elle tient compte également du temps consacré à d'autres formes de formation professionnelle, par exemple, la formation en cours d'emploi. Par conséquent, la somme de la FG et de la PPS risque de surestimer les exigences scolaires de nombreuses professions, notamment celles où bon nombre de compétences sont acquises au moyen de la formation en cours d'emploi, de l'apprentissage, etc.

Tableau 1 Niveaux moyens d'années de scolarité, de FG et de FG+PPS selon le niveau de scolarité le plus élevé et selon le sexe

Niveau de scolarité le plus élevé	Hommes			Femmes		
	Années de scolarité	FG	FG+PPS	Années de scolarité	FG	FG+PPS
Ensemble	13,1	11,6	14,2	12,8	11,7	13,7
Cours primaire partiel	4,5	8,5	9,2	6,1	9,2	9,5
Cours primaire	8,4	9,8	11,5	8,1	9,5	10,1
Études secondaires partielles	9,9	10,1	11,6	10,1	10,5	12,1
Études secondaires	12,4	11,1	13,5	12,1	11,5	13,4
Études postsecondaires (non universitaires)	14,4	11,9	14,8	14,2	12,0	14,0
Diplôme universitaire de premier cycle	17,2	13,6	17,1	17,5	13,8	16,8
Études supérieures	18,8	15,1	20,8	17,9	14,4	17,1
	n = 1 046	n = 1 049	n = 1 049	n = 914	n = 915	n = 915

Dans le tableau 1, le nombre moyen d'années de scolarité par niveau de scolarité le plus élevé atteint est conforme à ce que l'on attendrait; il est aussi très semblable chez les femmes et chez les hommes. Il convient de mentionner que les diplômés universitaires n'ayant pas fait d'études supérieures déclarent en moyenne plus de 17 années de scolarité accomplies. Selon la mesure FG, les diplômés universitaires sont donc, en moyenne, surinstruits, car le niveau de scolarité maximal exigé selon cette mesure est de 17 ans.

Selon les deux mesures, le nombre moyen d'années de formation requises augmente en fonction du plus haut niveau de scolarité atteint. La durée de formation nécessaire est très semblable chez les femmes et chez les hommes possédant le même plus haut niveau de scolarité atteint. En outre, l'écart entre les deux mesures (durée de la PPS) augmente.

Les femmes et les hommes ayant fait des études secondaires partielles travaillent dans des emplois qui exigent en moyenne environ dix années de FG et environ 11,5 années de FG et de PPS (donc, environ 1,5 année de PPS). La différence entre les deux mesures est de plus de deux ans chez les diplômés de l'enseignement secondaire et postsecondaire (non universitaire) et de plus de trois ans chez les diplômés universitaires. Ainsi, lorsque les niveaux d'instruction sont élevés, la durée de la formation nécessaire mesurée en fonction de la FG et de la PPS augmente beaucoup plus rapidement que la scolarité requise mesurée en fonction de la FG.

Surinstruction ou sous-instruction?

Ces différences entre les deux mesures de la scolarité requise entraînent des différences importantes dans les mesures de la surinstruction qui en résultent. Dans chaque case du tableau, on peut calculer le nombre moyen d'années de surinstruction ou de sous-instruction en soustrayant le nombre moyen d'années de scolarité accomplies du nombre moyen d'années de scolarité requise. En mesurant la scolarité requise en fonction de la FG, on constate qu'à chaque niveau de scolarité postsecondaire, les diplômés et les diplômées comptent en moyenne plus de deux années de surinstruction. En utilisant la mesure FG+PPS, on observe une correspondance assez étroite entre le nombre moyen d'années de scolarité des diplômés de l'enseignement post-secondaire et le nombre moyen d'années de formation exigé par leur emploi.

En moyenne, les diplômés de l'enseignement secondaire sont surinstruits selon la mesure FG et sous-instruits selon la mesure FG+PPS. Le nombre moyen d'années de scolarité des personnes ayant fait des études secondaires partielles correspond à peu près au nombre

moyen d'années de scolarité exigé par leur emploi, mesuré en fonction de la FG, mais représente environ deux ans de moins que la durée de la formation nécessaire mesurée en fonction de la FG+PPS. (Pour ces deux groupes, la différence entre le nombre d'années de scolarité et le nombre d'années de FG et de PPS s'explique, en tout ou en partie, par la préparation professionnelle spécifique reçue en dehors des établissements d'enseignement.)

Les deux mesures des exigences scolaires donnent donc des résultats contradictoires quant à l'ampleur de la surinstruction et de la sous-instruction au sein de la population active. La mesure FG suppose qu'en moyenne, les femmes comptent une année de surinstruction et les hommes, une année et demie. La mesure FG+PPS suppose que les hommes et les femmes comptent, en moyenne, une année de sous-instruction. Les deux mesures concordent sur un point : la surinstruction augmente (et la sous-instruction diminue) en fonction du plus haut niveau de scolarité atteint.

Ces deux mesures ont en commun certaines difficultés. Toutes deux décrivent un niveau moyen d'exigences relatives à la durée de la formation, mais sans tenir compte des différences concernant ces exigences entre des emplois au sein d'une profession détaillée. En outre, la FG et la PPS sont fondées sur des analyses des professions publiées pour la première fois au début des années 1970. Elles ne tiennent compte d'aucune hausse des exigences relatives à la durée de la formation professionnelle.

C'est d'ailleurs ce qui amène Hartog (2000, p. 132 à 133) à conclure que pour mesurer la surinstruction et la sous-instruction, il est préférable d'utiliser l'évaluation des exigences scolaires de l'emploi faite par les travailleurs. Or, ce type de mesure présente aussi des inconvénients, car il est possible que les employeurs embauchent en fonction de la disponibilité de personnes possédant un niveau d'instruction donné, plutôt qu'en fonction des véritables exigences de l'emploi. Il peut aussi créer une variance factice des exigences de l'emploi⁸.

Quoi qu'il en soit, l'EIAA ne demandait pas aux personnes occupées d'évaluer les exigences scolaires de leur emploi, de sorte qu'on ne dispose pas d'une mesure de la scolarité requise fondée sur l'évaluation des travailleurs. Comme le mentionne Hartog (2000, p. 133), « Peut-on préférer une mesure à l'autre? Le choix est habituellement dicté par la disponibilité des données... » [traduction]. Si nous utilisons les exigences relatives à la durée de formation fondées sur la profession, c'est parce qu'il s'agit des seules mesures disponibles des exigences scolaires de l'emploi.

Comme nous n'avons aucune raison impérieuse de préférer l'une ou l'autre mesure des exigences de l'emploi, nous les avons utilisées toutes les deux dans notre étude. Ainsi que nous l'avons mentionné, l'une de ces mesures suppose que les travailleurs canadiens sont, dans l'ensemble, nettement surinstruits pour les emplois qu'ils occupent; l'autre suppose qu'ils sont, dans l'ensemble, nettement sous-instruits. Cette divergence fait ressortir l'inexactitude des mesures de la scolarité requise que nous employons. Étant donné leur caractère approximatif, il nous semble utile de présenter des résultats fondés sur les deux mesures.

Chapitre 3

Déterminants de la concordance avec la scolarité requise

Littératie et concordance avec la scolarité requise

La présente section porte sur la relation entre les caractéristiques observées des personnes et les exigences scolaires de leur emploi. Une personne choisit un emploi parmi ceux qui lui sont accessibles. On suppose que les employeurs engagent uniquement les personnes qui possèdent les compétences exigées par l'emploi, que, pour le même salaire, ils préfèrent les personnes hautement qualifiées aux personnes moyennement qualifiées et que, pour un emploi donné, ils sont prêts à verser un salaire plus élevé aux personnes hautement qualifiées qu'aux personnes moyennement qualifiées. Cela suppose que pour les emplois dont les exigences professionnelles sont élevées, on embauche des travailleurs hautement qualifiés et que les travailleurs qui se situent à des niveaux de compétence différents peuvent occuper le même emploi à des taux de rémunération différents. Les suppositions qui sous-tendent cette description du comportement de l'employeur sont semblables au concept des profils études-productivité concernant les professions, dont il est question dans Borghans et de Grip (2000, p. 6 à 10)⁹.

On suppose que la scolarité produit des compétences, dont les compétences en littératie. Des personnes possédant le même nombre d'années de scolarité peuvent se situer à différents niveaux de compétence produits par la scolarité, notamment au chapitre des compétences en littératie. D'autres types d'investissement dans le capital humain, dont la formation en cours d'emploi, produisent également des compétences. La scolarité ne permet pas nécessairement d'acquérir certaines compétences produites par la formation en cours d'emploi. Les emplois dont les exigences scolaires sont très élevées peuvent aussi exiger des compétences très élevées produites par d'autres types d'investissement dans le capital humain. (La corrélation entre la FG et la PPS peut être attribuable en partie à une telle complémentarité entre les compétences.) Sicherman (1991) aborde bon nombre de ces aspects et ce qu'ils supposent pour la modélisation du jumelage emploi-travailleur, de la surinstruction et de la sous-instruction.

Parmi les emplois qui lui sont accessibles, soit ceux pour lesquels elle répond aux exigences professionnelles de l'employeur, une personne est réputée choisir l'emploi dont le

niveau de rémunération est le plus élevé. Pryor et Schaffer constatent que les travailleurs qui se situent à un niveau de scolarité donné mais dont les compétences en littératie sont inférieures à la moyenne pour leur scolarité travaillent souvent dans un emploi dont les exigences scolaires sont inférieures à leur niveau d'instruction. Boothby (1999) constate la même tendance chez les travailleurs canadiens ayant fait des études postsecondaires.

Deux raisons pourraient expliquer la concentration de travailleurs dont les compétences en littératie sont inférieures à la moyenne pour leur niveau de scolarité dans des emplois dont les exigences scolaires sont inférieures à leur niveau de scolarité. Premièrement, les travailleurs qui se situent au bas de l'échelle des compétences en littératie pour un niveau de scolarité donné ne répondent pas nécessairement aux exigences professionnelles minimales d'emplois dont les exigences scolaires correspondent à leur niveau de scolarité. Deuxièmement, il se peut qu'en raison du profil compétences-salaire de certaines professions, ces travailleurs gagnent plus dans des professions dont les exigences scolaires sont inférieures à leur niveau de scolarité. Ce serait le cas si le rendement de la surinstruction dans les professions pour lesquelles ces travailleurs sont surinstruits compensait amplement la pénalité salariale liée à des compétences en littératie inférieures à la moyenne chez ces travailleurs dans les professions dont les exigences scolaires correspondent à leur niveau de scolarité.

Dans un cas comme dans l'autre, on observe que les travailleurs « surinstruits » sont ceux dont les compétences en littératie sont inférieures à la moyenne pour leur scolarité. Si les travailleurs « surinstruits » possèdent des compétences en littératie supérieures à la moyenne pour la profession dans laquelle ils travaillent, le « rendement de la surinstruction » pourrait constituer plutôt, du moins en partie, le rendement de ces compétences en littératie.

De même, les travailleurs sous-instruits seraient ceux dont les compétences en littératie sont supérieures à la moyenne pour leur scolarité. Ces compétences supérieures leur auraient permis d'accéder à des professions dont les exigences scolaires sont plus élevées. La tendance observée du rendement de la sous-instruction serait attribuable aux compétences en littératie inférieures des travailleurs sous-instruits plutôt qu'à celles des travailleurs dont le niveau de scolarité correspond aux exigences de la profession.

Selon ce point de vue, la relation entre les caractéristiques des travailleurs et les exigences scolaires de leur emploi suppose les points suivants. Premièrement, les exigences scolaires de l'emploi constituent une mesure du niveau de compétence exigé par l'emploi, notamment au chapitre des compétences en littératie. Il peut s'agir de compétences complémentaires acquises au moyen d'autres formes d'investissement dans le capital humain que la scolarité. Deuxièmement, les travailleurs comptant un grand nombre d'années de scolarité et de formation en cours d'emploi sont plus susceptibles de travailler dans un emploi dont le niveau de compétence est élevé (mesuré d'après le niveau élevé des exigences scolaires).

Troisièmement, l'ajout d'une mesure des compétences en littératie à la modélisation du niveau de compétence correspondant à l'emploi devrait avoir deux effets. Les personnes dont les compétences en littératie sont élevées devraient occuper un emploi dont les exigences scolaires sont élevées. L'ajout d'une mesure directe des compétences en littératie devrait réduire l'effet de la scolarité sur le niveau de compétence correspondant à l'emploi, puisque l'effet de la scolarité est attribuable en partie à son effet sur les compétences en littératie¹⁰.

Estimations des déterminants de la scolarité requise

Le tableau 2 présente les résultats des estimations de régression des déterminants du niveau des exigences scolaires pour les emplois occupés par les travailleuses et les travailleurs à temps plein. Les résultats sont présentés séparément pour les deux mesures des exigences scolaires chez les femmes et chez les hommes.

Tableau 2 Régressions de référence relatives aux déterminants des exigences scolaires de la profession

	Hommes (n = 925)		Femmes (n = 836)	
	FG	FG+PPS	FG	FG+PPS
Constante	4,415 (<i>,451</i>)	2,001 (<i>,868</i>)	6,809 (<i>,435</i>)	6,372 (<i>,813</i>)
Années de scolarité	<i>,500</i> (<i>,024</i>)	<i>,840</i> (<i>,047</i>)	<i>,390</i> (<i>,026</i>)	<i>,590</i> (<i>,049</i>)
Expérience	<i>,052</i> (<i>,007</i>)	<i>,094</i> (<i>,014</i>)	<i>,026</i> (<i>,007</i>)	<i>,052</i> (<i>,139</i>)
Atlantique	-,084 (<i>,267</i>)	-,063 (<i>,514</i>)	-,468 (,241)	-,651 (<i>,452</i>)
Québec	,039 (<i>,185</i>)	-,053 (<i>,356</i>)	-,875 (,176)	-1,39 (,330)
Ontario	—	—	—	—
Prairies	,050 (<i>,213</i>)	,178 (<i>,410</i>)	-,527 (,202)	-,507 (<i>,379</i>)
Colombie-Britannique	-,654 (,266)	-1,07 (,512)	,047 (<i>,220</i>)	,591 (<i>,412</i>)
Grande région urbaine (>500 000 h.)	—	—	—	—
Petite région urbaine (<500 000 h.)	-,559 (,166)	-,895 (,319)	-,464 (,157)	-,954 (,294)
Région rurale	-,756 (,213)	-1,39 (,410)	-,676 (,185)	-1,19 (,347)
Jamais mariés	—	—	—	—
Non-célibataires	,047 (<i>,196</i>)	,109 (<i>,376</i>)	,183 (<i>,183</i>)	,014 (<i>,343</i>)
Veufs, divorcés, séparés	,466 (<i>,332</i>)	,963 (<i>,638</i>)	,284 (<i>,206</i>)	-,625 (<i>,386</i>)
r² (ajusté selon le degré de liberté)	,36	,30	,27	,20

Entre parenthèses : erreurs-types estimatives.

En gras et en italiques : chiffres significatifs à 1 %. En gras seulement : chiffres significatifs à 5 %.

Plusieurs aspects des effets estimatifs de la scolarité et de l'expérience de travail sont dignes de mention¹¹. Premièrement, tant chez les femmes que chez les hommes, les coefficients de ces variables sont beaucoup plus élevés selon la mesure FG+PPS que selon la mesure FG. Deuxièmement, selon les deux mesures, les coefficients de la scolarité et de l'expérience de travail sont beaucoup plus élevés chez les hommes que chez les femmes.

Si, dans les régressions de la scolarité requise, le coefficient de la scolarité équivalait à un, une augmentation d'une année de scolarité supposerait une augmentation d'une année de scolarité requise, ce qui laisserait inchangé le nombre d'années de surinstruction (ou de sous-instruction). Or, dans ces régressions, tous les coefficients de la scolarité sont inférieurs à un. Ainsi, en moyenne, la surinstruction augmente (ou la sous-instruction diminue) à mesure que le niveau de scolarité augmente. L'augmentation de la surinstruction en fonction de la scolarité est la plus faible dans la régression selon la mesure FG+PPS chez les hommes (coefficient de la scolarité = 0,840) et la plus forte dans la régression selon la mesure FG chez les femmes (coefficient de la scolarité = 0,390).

Un niveau élevé d'expérience de travail est associé à des exigences scolaires élevées de l'emploi, soit parce que l'expérience de travail remplace en partie les études, soit parce que certains emplois dont les exigences scolaires sont élevées comportent des exigences complémentaires en matière d'expérience de travail (voir Sicherman, 1991). Les coefficients de la scolarité et de l'expérience de travail révèlent qu'il faut compter au moins neuf années d'expérience de travail pour obtenir le même effet qu'une année de scolarité sur les exigences scolaires de l'emploi.

L'expérience de travail est associée à des exigences scolaires de l'emploi plus élevées chez les hommes que chez les femmes. C'est peut-être parce que l'approximation tenant lieu de l'expérience de travail—âge - années de scolarité - 6—surestime l'expérience de travail des femmes, d'où une sous-estimation possible du coefficient¹².

Outre la scolarité et l'expérience de travail, les régressions présentées dans le tableau 2 comprennent une série d'indicateurs liés à la géographie et à l'état civil. Les variables liées à l'état civil n'ont pas d'effets statistiquement significatifs. Par contre, le fait d'habiter une petite région urbaine (moins de 500 000 habitants) ou une région rurale a des effets négatifs importants sur les exigences scolaires de l'emploi. Ce phénomène est peut-être attribuable à une composition différente de l'emploi dans ces régions par rapport aux grandes régions urbaines.

Selon qu'il s'agit des femmes ou des hommes, les effets régionaux sont très différents¹³. Chez les hommes, le seul effet statistiquement significatif est un effet négatif du fait de résider en Colombie-Britannique (par rapport au fait de résider en Ontario). Chez les femmes, le fait de résider au Québec a un effet négatif statistiquement significatif sur les deux ensembles d'estimations et le fait de résider dans les provinces de l'Atlantique ou dans les provinces des Prairies a un effet négatif statistiquement significatif sur les estimations selon la mesure FG.

Le tableau 3 montre l'effet, sur les coefficients du nombre d'années de scolarité, de l'ajout de mesures des compétences en littératie aux régressions présentées dans le tableau 2. On utilise deux mesures des compétences en littératie : la moyenne des trois « premières valeurs plausibles » en compréhension de textes suivis, de textes schématisés et de textes au contenu quantitatif (appelée note en littératie) et une mesure fondée sur le pourcentage de bonnes réponses aux items de test de l'EIAA¹⁴. On utilise cette seconde mesure parce que les notes en littératie présentées dans l'EIAA sont conditionnées en partie par le niveau de scolarité du répondant. L'utilisation d'une mesure fondée directement sur les items de test constitue un moyen de repérer le biais pouvant résulter de l'utilisation de notes en littératie qui sont conditionnées par la scolarité. (Ces aspects sont abordés en détail dans l'annexe 3.)¹⁵

Tableau 3 Effets du nombre d'années de scolarité sur les exigences scolaires de la profession

	FG			FG+PPS		
Hommes (n = 925)						
Années de scolarité	,500	(,024)	,389	(,028)	,448	(,027)
Note en littératie	—		,014	(,002)	—	
Proportion de bonnes réponses	—		—		2,125	(,465)
r² (ajusté selon le degré de liberté)	,36		,40		,30	
Femmes (n = 836)						
Années de scolarité	,390	(,026)	,301	(,029)	,366	(,028)
Note en littératie	—		,011	(,002)	—	
Proportion de bonnes réponses	—		—		,972	(,402)
r² (corrige selon le degré de liberté)	,27		,30		,20	

Dans chaque régression présentée dans le tableau 3 et comprenant une mesure des compétences en littératie comme variable explicative, un niveau élevé de compétences en littératie est associé au fait de travailler dans une profession où les exigences scolaires sont élevées. Lorsqu'on inclut les compétences en littératie, le coefficient de la scolarité diminue. On peut en déduire qu'une partie de l'effet apparent de la scolarité sur l'affectation professionnelle constitue en réalité un effet des compétences en littératie. (Dans la mesure où les compétences en littératie sont produites par la scolarité, l'effet des compétences en littératie constitue un effet indirect de la scolarité.)

Il est également vrai que dans chaque cas, la baisse du coefficient de la scolarité est plus importante lorsqu'on utilise la note en littératie plutôt que le pourcentage de bonnes réponses comme mesure des compétences en littératie. Il se peut que le conditionnement des notes en littératie par rapport à la scolarité entraîne un biais du coefficient de la scolarité. Il est également possible que la régression qui utilise la note en littératie mesure mieux les effets de la scolarité (lorsqu'on tient compte des compétences en littératie) que la régression qui utilise le pourcentage de bonnes réponses.

Les résultats présentés dans le tableau 2 confirment l'hypothèse selon laquelle les compétences acquises au moyen de la scolarité et celles acquises au moyen de la formation en cours d'emploi ont tendance à accroître le niveau de compétence des emplois pour lesquels les travailleurs sont qualifiés, mesuré en fonction de la durée de la formation nécessaire. Les résultats présentés dans le tableau 3 confirment les hypothèses selon lesquelles 1) les compétences en littératie influencent les niveaux de compétence des travailleurs et 2) une partie des effets estimatifs de la scolarité dans les régressions du tableau 2 constitue en réalité un effet des compétences en littératie. Tous ces effets sont plus importants chez les hommes que chez les femmes.

PUBLICATIONS ÉLECTRONIQUES DISPONIBLES À
www.statcan.ca



Chapitre 4

Surinstruction, sous-instruction, littératie, utilisation des compétences en littératie et gains

Résultats des régressions

Nous abordons maintenant l'estimation du rendement de la surinstruction et de la sous-instruction ainsi que de la relation entre les compétences en littératie et ce rendement. Le tableau 4 présente des estimations de référence des régressions des gains des femmes et des hommes, avant l'inclusion des mesures de la surinstruction, de la sous-instruction, des compétences en littératie et de l'utilisation des compétences en littératie au travail. La variable dépendante est le logarithme des gains annuels.

Dans notre échantillon provenant de l'EIAA, on observe des effets statistiquement significatifs du nombre d'années de scolarité et du carré du nombre d'années de scolarité. Le coefficient du premier est positif; celui du second est négatif. Cette tendance à la baisse du rendement de la scolarité diffère des résultats résumés par Hartog (2000, p. 135). (Les questions de forme fonctionnelle sont abordées dans l'annexe 2.) D'après les tendances du rendement de la scolarité montrées dans le tableau 4, le rendement d'une année de scolarité supplémentaire serait positif jusqu'à concurrence de 21 années de scolarité tant chez les femmes que chez les hommes, mais plus important chez les femmes que chez les hommes jusqu'à concurrence d'environ 18 années de scolarité.

Le rendement marginal estimatif de l'expérience de travail est plus élevé chez les femmes que chez les hommes jusqu'à concurrence de treize années d'expérience de travail¹. L'approximation tenant lieu de l'expérience de travail ($\hat{\text{âge}} - \text{années de scolarité} - 6$) a tendance à surestimer l'expérience de travail des femmes, donc à entraîner une sous-estimation du rendement de cette expérience. Si l'on disposait d'une mesure plus exacte de l'expérience de travail pour l'échantillon, l'écart entre le rendement de l'expérience de travail des hommes et celui des femmes pourrait bien être supérieur à nos estimations.

Les tendances de la variation régionale des gains sont semblables chez les femmes et chez les hommes, sauf que les hommes de la Colombie-Britannique gagnent beaucoup moins que ceux de l'Ontario, alors que les femmes de cette province ne gagnent pas beaucoup moins que les Ontariennes. Les hommes et les femmes qui travaillent en région rurale gagnent beaucoup moins que leurs homologues des centres urbains. Les hommes non célibataires gagnent beaucoup plus que ceux qui n'ont jamais été mariés, mais il n'y a pas de lien significatif entre leur état civil et leurs gains. Chez les femmes, l'état civil n'a aucun lien significatif avec les gains (rappelons qu'il s'agit d'un échantillon de personnes occupées à temps plein).

Tableau 4 Régressions de référence du logarithme des gains annuels moyens

	Hommes (n=774)		Femmes (n=684)	
Constante	7,770	(,374)	6,773	(0,455)
Années de scolarité	,189	(,048)	,302	(,063)
Années de scolarité ²	-,004	(,002)	-,007	(,002)
Expérience	,043	(,009)	,062	(,008)
Expérience ²	-,0005	(,0002)	-,001	(,0002)
Atlantique	-,372	(,088)	-,416	(,086)
Québec	-,170	(,061)	-,148	(,064)
Ontario	—	—	—	—
Prairies	-,090	(,071)	-,157	(,074)
Colombie-Britannique	-,410	(,087)	,094	(,079)
Grande région urbaine (>500 000 h.)	—	—	—	—
Petite région urbaine (<500 000 h.)	-,056	(,054)	-,017	(,056)
Région rurale	-,158	(,073)	-,174	(,065)
Jamais mariés	—	—	—	—
Non-célibataires	,550	(,075)	-,034	(,071)
Veufs, divorcés, séparés	-,090	(,112)	-,043	(,072)
r² (ajusté selon le degré de liberté)	,28		,29	

Entre parenthèses : erreurs-types estimatives.

En gras et en italiques : chiffres significatifs à 1 %. En gras seulement : chiffres significatifs à 5 %.

Le tableau 5 montre les effets observés lorsqu'on ajoute aux régressions des gains du tableau 4 des mesures de la surinstruction, de la sous-instruction, des compétences en littératie et de l'utilisation des compétences en littératie au travail. À part les coefficients de ces mesures, seuls les coefficients estimatifs du nombre d'années de scolarité et de son carré figurent dans le tableau 5¹⁷.

La forme fonctionnelle des régressions présentées dans le tableau 5 et comprenant des mesures de la surinstruction et de la sous-instruction diffère de celle utilisée par Sicherman (1991), que nous avons évoquée plus haut. Le modèle estimé par Sicherman est formulé comme suit :

$$1. \quad \ln Y = c + rR + oO + uU + X'B + e$$

où Y représente les gains, R représente la scolarité requise dans la profession de la personne, O représente le nombre d'années de surinstruction et U représente le nombre d'années de sous-instruction; r, o et u sont leurs coefficients respectifs; c représente les constantes de régression, X est un vecteur d'autres variables explicatives, B, un vecteur de leurs coefficients et e, le terme d'erreur (nous avons omis les indices inférieurs pour ne pas alourdir l'équation). Supposons que S représente le nombre d'années de scolarité. Par

définition, $O = \{S-R \text{ lorsqu'il est positif, } 0 \text{ dans le cas contraire}\}$ et $U = \{R-S \text{ lorsqu'il est positif, } 0 \text{ dans le cas contraire}\}$. On peut reformuler l'équation comme suit :

$$2. \quad \ln Y = c + rS + o'O + u'U + X'B + e$$

où $o' := o-r$ et $u' := r + u$.

Rappelons que dans l'équation 1, la tendance prévue des signes est la suivante : $r > o > 0$, $u < 0$, $r + u > 0$.

D'où les prévisions suivantes pour l'équation 2 : $r > 0$, $o' < 0$, $r + o' > 0$, $r > u' > 0$. L'hypothèse nulle de l'équation 1 que Hartog (2000, p. 135) appelle spécification de Mincer ($H_0: r = o = -u$) est équivalente à l'hypothèse nulle de l'équation 2 : $H_0' : o' = 0$, $u' = 0$.

Dans l'équation 1, le coefficient o de O est habituellement appelé le rendement de la surinstruction. Appelons le coefficient o' de O dans l'équation 2 la pénalité liée à la surinstruction, car il s'agit de la valeur dont le logarithme des gains diminue pour une année de surinstruction dans l'équation 2. Le coefficient u de U dans l'équation 1 est la pénalité liée à la sous-instruction. Appelons le coefficient u' de U le rendement de la sous-instruction, car il s'agit de la valeur dont le logarithme des gains augmente pour chaque année de sous-instruction dans l'équation 2¹⁸.

À cause de la non-linéarité du rendement de la scolarité dans notre échantillon, nous avons préféré une forme fonctionnelle qui comprend explicitement le nombre d'années de scolarité. La forme fonctionnelle des régressions du rendement de la surinstruction et de la sous-instruction présentées dans le tableau 5 est la suivante :

$$3. \quad \ln Y = c + rS + sS^2 + o'O + u'U + X'B + e$$

Le vecteur X des autres caractéristiques comprend toutes les variables du tableau 4 (sauf les termes du nombre d'années de scolarité). Dans certaines régressions, il comprend également la note en littératie, ou cette note et une mesure de l'utilisation des compétences en littératie.

Tableau 5 Effets de la scolarité, de la surinstruction et de la sous-instruction, des compétences en littératie et de l'utilisation de la littératie sur le logarithme des gains annuels

	(1)	(2)	(3)	(4)	(5)
Hommes (n = 774)					
Années de scolarité	<i>,189</i> (<i>,048</i>)	<i>,213</i> (<i>,049</i>)	<i>,205</i> (<i>,047</i>)	<i>,134</i> (<i>,049</i>)	<i>,110</i> (<i>,048</i>)
Années de scolarité ²	<i>-,004</i> (<i>,002</i>)	<i>-,004</i> (<i>,002</i>)	<i>-,005</i> (<i>,002</i>)	<i>-,003</i> (<i>,002</i>)	<i>-,003</i> (<i>,002</i>)
Années de surinstruction (FG)	—	<i>-,070</i> (<i>,014</i>)	—	—	—
Années de sous-instruction (FG)	—	<i>,070</i> (<i>,023</i>)	—	—	—
Années de surinstruction (FG+PPS)	—	—	<i>-,055</i> (<i>,013</i>)	—	—
Années de sous-instruction (FG+PPS)	—	—	<i>,030</i> (<i>,008</i>)	—	—
Note en littératie	—	—	—	<i>,003</i> (<i>,001</i>)	—
Proportion de bonnes réponses	—	—	—	—	<i>,989</i> (<i>,157</i>)
Utilisation de la littératie au travail	—	—	—	—	—
r² (corrige selon le degré de liberté)	,28	,32	,32	,30	,31
Femmes (n = 684)					
Années de scolarité	<i>,302</i> (<i>,063</i>)	<i>,343</i> (<i>,062</i>)	<i>,317</i> (<i>,059</i>)	<i>,217</i> (<i>,064</i>)	<i>,251</i> (<i>,064</i>)
Années de scolarité ²	<i>-,007</i> (<i>,002</i>)	<i>-,006</i> (<i>,002</i>)	<i>-,006</i> (<i>,002</i>)	<i>-,005</i> (<i>,002</i>)	<i>-,006</i> (<i>,002</i>)
Années de surinstruction (FG)	—	<i>-,125</i> (<i>,017</i>)	—	—	—
Années de sous-instruction (FG)	—	<i>,100</i> (<i>,026</i>)	—	—	—
Années de surinstruction (FG+PPS)	—	—	<i>-,115</i> (<i>,017</i>)	—	—
Années de sous-instruction (FG+PPS)	—	—	<i>,035</i> (<i>,009</i>)	—	—
Note en littératie	—	—	—	<i>,004</i> (<i>,001</i>)	—
Proportion de bonnes réponses	—	—	—	—	<i>,508</i> (<i>,146</i>)
Utilisation de la littératie au travail	—	—	—	—	—
r² (corrige selon le degré de liberté)	,29	,38	,38	,33	,30
	(6)	(7)	(8)	(9)	
Hommes (n = 774)					
Années de scolarité	<i>,172</i> (<i>,050</i>)	<i>,166</i> (<i>,048</i>)	<i>,144</i> (<i>,051</i>)	<i>,147</i> (<i>,049</i>)	
Années de scolarité ²	<i>-,003</i> (<i>,002</i>)	<i>-,004</i> (<i>,002</i>)	<i>-,003</i> (<i>,002</i>)	<i>-,003</i> (<i>,002</i>)	
Années de surinstruction (FG)	<i>-,061</i> (<i>,014</i>)	—	<i>-,056</i> (<i>,015</i>)	—	
Années de sous-instruction (FG)	<i>,060</i> (<i>,025</i>)	—	<i>,047</i> (<i>,025</i>)	—	
Années de surinstruction (FG+PPS)	—	<i>-,049</i> (<i>,013</i>)	—	<i>-,048</i> (<i>,013</i>)	
Années de sous-instruction (FG+PPS)	—	<i>,026</i> (<i>,008</i>)	—	<i>,021</i> (<i>,008</i>)	
Note en littératie	<i>,002</i> (<i>,001</i>)				
Proportion de bonnes réponses	—	—	—	—	
Utilisation de la littératie au travail	—	—	<i>,054</i> (<i>,025</i>)	<i>,049</i> (<i>,025</i>)	
r² (corrige selon le degré de liberté)	,32	,33	,33	,33	
Femmes (n = 684)					
Années de scolarité	<i>,265</i> (<i>,064</i>)	<i>,255</i> (<i>,060</i>)	<i>,233</i> (<i>,063</i>)	<i>,238</i> (<i>,059</i>)	
Années de scolarité ²	<i>-,004</i> (<i>,002</i>)	<i>-,005</i> (<i>,002</i>)	<i>-,004</i> (<i>,002</i>)	<i>-,005</i> (<i>,002</i>)	
Années de surinstruction (FG)	<i>-,123</i> (<i>,017</i>)	—	<i>-,112</i> (<i>,017</i>)	—	
Années de sous-instruction (FG)	<i>,075</i> (<i>,026</i>)	—	<i>,047</i> (<i>,026</i>)	—	
Années de surinstruction (FG+PPS)	—	<i>-,114</i> (<i>,016</i>)	—	<i>-,110</i> (<i>,016</i>)	
Années de sous-instruction (FG+PPS)	—	<i>,026</i> (<i>,009</i>)	—	<i>,013</i> (<i>,009</i>)	
Note en littératie	<i>,003</i> (<i>,001</i>)				
Proportion de bonnes réponses	—	—	—	—	
Utilisation de la littératie au travail	—	—	<i>,134</i> (<i>,025</i>)	<i>,138</i> (<i>,025</i>)	
r² (corrige selon le degré de liberté)	,39	,39	,42	,42	

Entre parenthèses : erreurs-types estimatives.

En gras et en italiques : chiffres significatifs à 1 %. En gras seulement : chiffres significatifs à 5 %. En italiques seulement : chiffres significatifs à 10 %.

La première colonne du tableau 5 reproduit les coefficients du nombre d'années de scolarité et du carré du nombre d'années de scolarité présentés dans le tableau 4. Les deuxième et troisième colonnes montrent, respectivement, les effets de l'ajout des mesures de la surinstruction et de la sous-instruction à la mesure FG et à la mesure FG+PPS.

Tant chez les femmes que chez les hommes, l'ajout de mesures de la surinstruction et de la sous-instruction augmente significativement le rendement d'une année de scolarité. Le rendement d'une année de scolarité supplémentaire exigée par l'emploi est beaucoup plus élevé que celui d'une année de scolarité supplémentaire non exigée par l'emploi. (Ce dernier rendement correspond au rendement mesuré d'une année de scolarité moins la pénalité liée à la surinstruction.) Tant chez les hommes que chez les femmes, le rendement de la scolarité par rapport à son niveau dans la régression de référence (colonne 1) augmente moins lorsqu'on utilise la mesure FG+PPS (colonne 3) que lorsqu'on utilise la mesure FG (colonne 2).

À l'égard des gains, la pénalité liée à la surinstruction est plus importante chez les femmes que chez les hommes et est un peu plus élevée selon la mesure FG que selon la mesure FG+PPS. La pénalité liée à une année de surinstruction correspond au rendement d'une année de scolarité entre les quinzième et dix-huitième années de scolarité (selon l'estimation). Jusqu'à ce point, une année de scolarité supplémentaire qui résulte en une année de surinstruction augmente néanmoins la moyenne des gains.

Le rendement de la sous-instruction est positif et statistiquement significatif. En d'autres termes, il y a un avantage sur le plan des gains pour les personnes qui travaillent dans une profession où les exigences scolaires sont supérieures à leur niveau d'instruction. L'avantage sur le plan des gains associé à une année de surinstruction est plus important chez les femmes que chez les hommes. Le rendement de la surinstruction est plus élevé lorsqu'on utilise la mesure FG que lorsqu'on utilise la mesure FG+PPS.

La colonne 4 montre l'effet qu'on obtient en ajoutant la mesure de la note en littératie à la régression des gains de base de la colonne 1. Comme dans les estimations de Green et Riddell (2001), le rendement de la scolarité diminue nettement lorsqu'on inclut la mesure de la note en littératie.

La colonne 5 utilise la mesure des compétences en littératie fondée sur le pourcentage de bonnes réponses, plutôt que la mesure de la note en littératie. Comme dans le tableau 3, l'impact de la mesure fondée sur le pourcentage de bonnes réponses sur le rendement de la scolarité est inférieur à celui de la mesure de la note en littératie. Ici comme dans le tableau 3, nous ignorons si cette différence résulte du conditionnement de la note en littératie par rapport à la scolarité. Dans le reste du tableau 5, nous employons la mesure de la note en littératie. Dans les colonnes 6 à 9, les résultats de la régression varient très peu lorsqu'on utilise la mesure fondée sur le pourcentage de bonnes réponses.

Les colonnes 6 et 7 comprennent les mesures de la surinstruction et de la sous-instruction, ainsi que la mesure de la note en littératie. La prise en compte des effets de la surinstruction et de la sous-instruction dans les colonnes 6 et 7 se traduit par une augmentation considérable du rendement de la scolarité par rapport à celui que montre la colonne 4.

Chez les hommes, les coefficients de la surinstruction et de la sous-instruction diminuent (en valeur absolue) entre les colonnes 2 et 3 et les colonnes 6 et 7. Chez les femmes, le coefficient de la sous-instruction diminue par rapport aux chiffres des colonnes 2 et 3, mais celui de la surinstruction varie très peu. Tant chez les femmes que chez les hommes, le coefficient de la mesure de la note en littératie est moins élevé dans les colonnes 6 et 7 que dans la colonne 4. N'oublions pas que la variable « note en littératie » a un effet significatif sur les exigences scolaires de l'emploi, indépendamment de la scolarité. Dans la colonne 4, le coefficient de la note en littératie peut donc comprendre les effets indirects de la littératie sur les gains, par le biais de l'affectation professionnelle, ainsi que des effets directs.

Les colonnes 8 et 9 ajoutent aux modèles des colonnes 6 et 7 une mesure de l'utilisation des compétences en littératie au travail. Cette mesure est fondée sur les réponses à six items concernant la lecture au travail et quatre autres concernant l'écriture au travail¹⁹. L'utilisation des compétences en littératie a des effets positifs significatifs sur les gains des femmes et des hommes; ces effets sont beaucoup plus importants chez les femmes²⁰. Lorsqu'on ajoute la variable « utilisation des compétences en littératie », le rendement de la scolarité diminue quelque peu chez les hommes et légèrement chez les femmes. Entre les colonnes 6 et 7 et les colonnes 8 et 9, les coefficients de la mesure de la note en littératie sont, pour l'essentiel, inchangés.

Dans les colonnes 8 et 9, les pénalités liées à la surinstruction sont légèrement inférieures à celles que montrent respectivement les colonnes 6 et 7. Lorsqu'on ajoute au modèle la variable « utilisation des compétences en littératie », le rendement de la sous-instruction diminue. Dans la colonne 9, le coefficient de la sous-instruction chez les femmes n'est pas statistiquement significatif. C'est le seul cas, dans le tableau 5, où le coefficient d'une mesure de la surinstruction ou de la sous-instruction n'est pas statistiquement significatif.

Gains, littératie, scolarité et jumelage emploi-travailleur

Le tableau 5 explore la relation entre la scolarité, le jumelage emploi-travailleur, la littératie et les gains. Ici, on se penche sur deux grandes questions : premièrement, les mesures de la surinstruction et de la sous-instruction utilisées montrent-elles les tendances du rendement de la surinstruction et de la sous-instruction observées par d'autres chercheurs? Deuxièmement, une certaine partie du rendement de la surinstruction et de la sous-instruction (le cas échéant) constitue-t-elle en réalité un rendement des compétences en littératie?

Selon une recherche antérieure sur le rendement de la surinstruction et de la sous-instruction (résumée par Hartog, 2000), on observe un rendement positif de chaque année de surinstruction. Celui-ci est inférieur au rendement du nombre d'années de scolarité exigé par l'emploi. Les travailleurs sous-instruits sont avantagés du fait d'occuper un emploi dont les exigences sont supérieures à leur scolarité, mais ce rendement est inférieur à celui du nombre d'« années manquantes » par rapport à la scolarité requise. En d'autres termes, une année de surinstruction se traduit par un rendement positif, mais par une pénalité par rapport à une année de scolarité requise; une année de sous-instruction se traduit par un rendement positif, mais par une pénalité par rapport à l'« année manquante » de la scolarité requise.

Les résultats présentés dans le tableau 5 suivent cette tendance selon nos deux mesures de la surinstruction et de la sous-instruction²¹. La pénalité liée à une année de surinstruction est un peu plus élevée selon la mesure FG que selon la mesure FG+PPS; le rendement de la sous-instruction est plus élevé selon la mesure FG que selon la mesure FG+PPS (colonnes 2 et 3).

Les mesures de la surinstruction et de la sous-instruction expliquent les effets de la concordance entre la scolarité et les exigences de l'emploi. Lorsqu'on tient compte des effets du jumelage emploi-travailleur, le rendement estimatif d'une année de scolarité doit être interprété comme le rendement d'une année de scolarité qui est exigée par l'emploi de la personne. Comme le montre le tableau 5, ce rendement est plus élevé que celui de la scolarité lorsqu'on fait abstraction du jumelage emploi-travailleur (colonnes 2 et 3, par rapport à la colonne 1).

L'ajout des mesures des compétences en littératie aux modèles du tableau 5 abaisse le rendement de la scolarité (colonnes 4 et 5, par rapport à la colonne 1). Selon Green et Riddell (2001), cette baisse du rendement de la scolarité s'explique du fait qu'une composante de ce rendement constitue en réalité un rendement des compétences en littératie produites par la scolarité.

Les modèles qui comprennent la mesure de la note en littératie montrent également des pénalités liées à la surinstruction et un rendement de la sous-instruction chez les hommes et un rendement de la sous-instruction chez les femmes, lesquels sont quelque peu inférieurs par rapport aux modèles qui font abstraction des compétences en littératie (colonnes 6 et 7, par rapport aux colonnes 2 et 3). Il semble donc probable que le rendement de la sous-instruction constitue en partie un rendement des compétences en littératie supérieures à la moyenne pour le niveau de scolarité de la personne. Chez les hommes, le rendement à la surinstruction peut refléter en partie le rendement des compétences en littératie supérieures à la moyenne pour les hommes occupant des emplois au même niveau de scolarité exigé.

Ces modèles montrent également un rendement inférieur des compétences en littératie par rapport aux modèles qui comprennent la mesure de la note en littératie, mais non les mesures de la surinstruction et de la sous-instruction (colonnes 6 et 7, par rapport aux colonnes 4 et 5). Selon notre interprétation, cela signifie que l'effet des compétences en littératie sur les gains s'exerce en partie par le biais du jumelage emploi-travailleur. Les tableaux 2 et 3 montrent que le niveau de la durée de la scolarité requise dans l'emploi est influencé par les compétences en littératie. Si les compétences en littératie influent sur les gains par le biais du jumelage emploi-travailleur, ce serait donc parce que les personnes dont les compétences en littératie sont inférieures obtiennent des emplois où la scolarité requise est inférieure et que les gains sont inférieurs dans ces emplois.

Si l'on interprète le rendement de la littératie en termes de gains comme un rendement des compétences, il semble plausible que ce rendement dépende de l'utilisation des compétences. Lorsqu'on ajoute une mesure de l'utilisation des compétences en littératie au travail aux modèles qui comprennent des mesures des compétences en littératie, de la surinstruction et de la sous-instruction, le rendement de la scolarité et celui de la sous-instruction diminuent tant chez les hommes que chez les femmes et la pénalité liée à la surinstruction diminue chez les femmes. L'utilisation des compétences en littératie au travail se traduit par un rendement positif (colonnes 8 et 9, par rapport aux colonnes 6 et 7).

Selon notre interprétation de ces résultats, l'utilisation des compétences en littératie au travail mesure des variations des exigences professionnelles des emplois dont ne tiennent pas compte nos mesures du niveau professionnel des exigences scolaires de l'emploi. D'autres interprétations sont possibles. Selon DiNardo et Pischke (1997), le fait d'être assis au travail a un effet positif sur les gains. Ce résultat fait ressortir la nécessité de faire preuve de prudence avant de déduire qu'un lien entre une activité professionnelle et des gains plus élevés signifie que la capacité d'exercer l'activité est récompensée par des gains plus élevés.

PUBLICATIONS ÉLECTRONIQUES DISPONIBLES À
www.statcan.ca



Conclusions

Avant de passer à des conclusions spécifiques, nous formulons d'abord une observation générale sur les différences dans nos résultats concernant les femmes et les hommes. Les résultats qualitatifs sont semblables chez les femmes et chez les hommes. Toutefois, il y a souvent des écarts importants entre les coefficients de certaines de nos variables essentielles chez les unes et chez les autres. Prenons, par exemple, les écarts entre les femmes et les hommes pour ce qui concerne le coefficient de la scolarité dans les régressions pour le niveau de compétence correspondant à une profession. À notre avis, ces écarts sont attribuables pour une large part à des différences persistantes dans les répartitions professionnelles des femmes et des hommes. Ces différences nous ont incités, dans les sections qui précèdent, à présenter tous les résultats par sexe.

Notre premier ensemble de conclusions concerne les déterminants des niveaux de compétence liés aux professions des travailleurs, mesurés en fonction des exigences scolaires de la profession. Nous avons montré que les compétences en littératie constituaient un déterminant important de l'affectation professionnelle par niveau de compétence, lorsqu'on tient compte du rôle de la scolarité. Cette observation concorde avec les résultats de Pryor et Schaffer (1997, 1999) et de Boothby (1999). Nous avons également montré que la relation entre la scolarité et l'affectation professionnelle (abstraction faite des compétences en littératie) constituait pour une large part un effet des compétences en littératie (qui sont étroitement liées à la scolarité). Enfin, nous avons montré que les compétences acquises au moyen de la formation en cours d'emploi pouvaient aussi constituer un déterminant important de l'affectation professionnelle. Cette observation concorde avec la conclusion de Sicherman (1991) selon laquelle la surinstruction et la sous-instruction peuvent résulter de l'effet de compétences acquises en cours d'emploi sur l'affectation professionnelle.

Ceci dit, il semble opportun de réitérer que selon nos estimations, le nombre d'années de scolarité a, de loin, la plus grande influence sur l'affectation professionnelle par niveau de compétence. Des différences très importantes dans l'expérience de travail et dans la note en littératie influent moins sur l'affectation professionnelle qu'une seule année de scolarité. À notre avis, les effets de la scolarité sont attribuables en partie à des compétences autres que les compétences en littératie qui sont produites par la scolarité. Rien dans nos résultats ne nous permet de préciser de quelles compétences il s'agit ni dans quelle mesure l'effet de la scolarité sur l'affectation professionnelle est attribuable à ces compétences.

Notre deuxième ensemble de conclusions concerne l'existence du rendement de la surinstruction et de la sous-instruction. À ce que nous sachions, la seule autre étude visant à mesurer ce rendement au Canada est celle de Vahey (2000), qui utilise des données provenant d'une enquête menée en 1982. Nous avons utilisé deux mesures différentes de la scolarité requise, toutes deux fondées sur la profession détaillée des personnes comprises dans

l'échantillon de l'EIAA et occupées à temps plein. L'une de ces mesures montre que les travailleurs compris dans notre échantillon sont, en moyenne, surinstruits; l'autre montre qu'ils sont, en moyenne, sous-instruits.

Nos résultats fondés sur chacune de ces mesures de la scolarité requise débouchent sur les mêmes conclusions. Nous constatons qu'une année de surinstruction se traduit par un rendement positif net (jusqu'à 17 années de scolarité ou plus), mais que ce rendement est nettement inférieur à celui d'une année de scolarité lorsque la scolarité correspond aux exigences de l'emploi. Chaque année de scolarité requise au-delà de la scolarité accomplie d'une personne (année de sous-instruction) donne un rendement, mais celui-ci est inférieur à celui que suppose le fait d'occuper le même emploi et de compter une année de scolarité supplémentaire. Telles sont les constatations habituelles des travaux de recherche sur le rendement de la surinstruction et de la sous-instruction résumés par Hartog (2000). Elles font ressortir le fait que les gains dépendent essentiellement de la concordance entre la scolarité et la profession, et non de la scolarité seule.

Comme nous l'avons souligné plus haut, les compétences en littératie constituent un déterminant important de l'affectation professionnelle. Notre prochain ensemble de conclusions concerne notre démarche visant à évaluer dans quelle mesure le rendement de la surinstruction et de la sous-instruction est déterminé par les compétences en littératie. À cette fin, nous avons inclus, dans nos modèles de régression du rendement de la surinstruction et de la sous-instruction, des mesures des compétences en littératie (seules) ainsi que des compétences en littératie et de l'utilisation de ces capacités au travail.

Lorsqu'on inclut les mesures des compétences en littératie, les coefficients estimatifs de la surinstruction et de la sous-instruction diminuent en valeur absolue chez les hommes et les coefficients estimatifs de la sous-instruction diminuent chez les femmes. Le fait d'inclure à la fois la mesure des compétences en littératie et celle de l'utilisation de ces capacités se traduit par d'autres réductions de tous ces coefficients, y compris ceux de la surinstruction chez les femmes.

Ces résultats concordent avec le modèle suivant (évoqué dans la section 1.3 ci-dessus) : les travailleurs dont les compétences en littératie sont inférieures à la moyenne pour leur niveau d'instruction travaillent habituellement dans une profession dont les exigences scolaires sont inférieures à leur niveau d'instruction. Ils possèdent habituellement des compétences en littératie supérieures à la moyenne pour leur profession et peuvent avoir tendance à occuper, au sein de cette profession, un emploi présentant un niveau élevé d'activités faisant appel aux compétences en littératie. Inversement, les travailleurs dont les compétences en littératie sont supérieures à la moyenne pour leur niveau d'instruction travaillent habituellement dans une profession dont les exigences scolaires sont supérieures à leur niveau d'instruction. Ils possèdent habituellement des compétences en littératie inférieures à la moyenne pour cette profession.

Les gains supérieurs des travailleurs surinstruits (par rapport aux gains de ceux qui possèdent uniquement la scolarité requise) constituent en partie un rendement des compétences en littératie des travailleurs « surinstruits », qui sont supérieures à la moyenne pour les professions dans lesquelles ils travaillent. Les gains inférieurs des travailleurs sous-instruits (par rapport aux gains de ceux qui possèdent uniquement la scolarité requise) constituent une pénalité liée aux compétences en littératie des travailleurs « sous-instruits », qui sont inférieures à la moyenne pour les professions dans lesquelles ils travaillent. Leurs gains supérieurs (par rapport à ceux qui possèdent la même scolarité et qui travaillent dans une profession qui exige exactement cette scolarité) constituent en partie un rendement de leurs compétences en littératie supérieures.

À notre avis, nos résultats montrent de manière raisonnablement probante la pertinence de ce modèle en ce qui concerne la sous-instruction tant chez les femmes que chez les hommes, puisque, dans les régressions du tableau 5, le coefficient de la sous-instruction diminue lorsqu'on inclut la mesure des compétences en littératie comme variable explicative. Dans le cas de la surinstruction, l'appui empirique pour ce modèle est fort chez les hommes (lorsqu'on inclut la mesure des compétences en littératie, la valeur absolue du coefficient diminue) et faible chez les femmes (le coefficient reste inchangé). Tant chez les femmes que chez hommes, le rendement d'une année de surinstruction est nettement inférieur à celui d'une année de scolarité requise, même lorsqu'on tient compte des compétences en littératie et de l'utilisation des compétences en littératie.

Les compétences en littératie ne sont pas le seul type de compétences acquises au moyen de la scolarité et de la formation en cours d'emploi. Comme dans le cas des compétences en littératie, le niveau des autres compétences pourrait bien varier chez les personnes possédant des niveaux de scolarité et une expérience de travail semblables. Si nous disposions de mesures directes de ces autres compétences, comme dans le cas des compétences en littératie, peut-être pourrions-nous expliquer entièrement le rendement de la surinstruction et de la sous-instruction.

Dernier point, de nature plus spéculative : nos résultats nous semblent indiquer que les employeurs sont en mesure de déceler les différences dans les compétences en littératie à l'intérieur d'un niveau de scolarité (ou dans d'autres compétences souhaitées qui sont étroitement liées aux compétences en littératie). Autrement, comment expliquer que des personnes dont les compétences en littératie sont inférieures à la moyenne soient affectées à une profession dont les exigences professionnelles sont inférieures à la moyenne pour leur niveau de scolarité et soient moins bien payées que la moyenne pour le niveau de compétences exigé dans leur profession? Cette observation remet en question les lacunes qu'auraient les employeurs dans l'évaluation des compétences de leurs employés actuels ou éventuels. Nos résultats démentent l'importance de ces lacunes dans le cas des compétences en littératie.

PUBLICATIONS ÉLECTRONIQUES DISPONIBLES À
www.statcan.ca



Bibliographie

- Boothby, D. (1999), « Literacy Skills, the Knowledge Content of Occupations and Occupational Mismatch », document de travail n° W-99-E3, Ottawa, Direction générale de la recherche appliquée, Développement des ressources humaines Canada.
- Borghans, L. et de Grip, A. (2000), « The Debate in Economics about Skill Utilization » dans Borghans, L. et de Grip, A., *The Over-Educated Worker? The Economics of Skill Utilization*, Cheltenham, Edward Elgar, p. 3 à 23.
- Boyd, M. (1990), « Sex Differences in Occupational Skill : Canada, 1961-1986 », *La Revue canadienne de sociologie et d'anthropologie*, vol. 27, n° 3, p. 285 à 315.
- Carey, S., Bridgwood, A. et Thomas, M. (2000), *Measuring Adult Literacy: The International Adult Literacy Survey in the European Context*, Londres, Office for National Statistics.
- Charette, M. et Meng, R. (1998), « The determinants of literacy and numeracy, and the effect of literacy and numeracy on labour market outcomes », *Revue canadienne d'économique*, vol. 31, n° 3, p. 495 à 517.
- DiNardo, J. et J. Pischke, (1997), « The Returns to Computer Use Revisited: Have Pencils Changed the Wage Structure Too? », *Quarterly Journal of Economics*, vol. 112, n° 1, p. 291 à 303.
- Green, D. et Riddell, W. (2001), *Les capacités de lecture et de calcul et la situation sur le marché du travail au Canada* (89-552 au catalogue, n° 8), Ottawa, Statistique Canada.
- Green, F., McIntosh, S. et Vignola, A. (2000), « Overeducation: A tough nut to crack », *Centre Piece*, vol. 5, n° 1.
- Hartog, J. (2000), « Over-education and earnings: where are we, where should we go? », *Economics of Education Review*, vol. 19, n° 2, p. 131 à 147.
- Hunter, A. et Manley, M. (1986), « On the task content of work », *La Revue canadienne de sociologie et d'anthropologie*, vol. 23, n° 1, p. 47 à 71.
- Kapsalis, C. (2000), *Literacy Profile of Ontario's Immigrants*, Toronto, Imprimeur de la Reine pour l'Ontario.
- Krahn, H. et Lowe, G. (1998), *L'utilisation des capacités de lecture en milieu de travail au Canada* (89-552 au catalogue, n° 4), Ottawa, Statistique Canada.
- Livingstone, D. (1999), *The Education-Jobs Gap: Underemployment or Economic Democracy*, Toronto, Garamond Press.
- Main-d'œuvre et Immigration (1971), *Classification canadienne descriptive des professions. Volume 1, Classification et définitions*, Ottawa, Main-d'œuvre et Immigration.

- Main-d'œuvre et Immigration (1973), *Classification canadienne descriptive des professions. Volume 2, Normes professionnelles*, Ottawa, Main-d'œuvre et Immigration.
- Murray, T., Kirsch, I. et Jenkins, L. (1998), *Adult Literacy in OECD Countries: Technical Report on the First International Adult Literacy Survey*, Washington, U.S. Government Printing Office.
- Myles, J. (1988), « The expanding middle: some Canadian evidence on the deskilling debate », *La Revue canadienne de sociologie et d'anthropologie*, vol. 25, n° 3, p. 335 à 364.
- Osberg, L. (2000), *Scolarité, alphabétisation et revenus personnels* (89-552 au catalogue, n° 7), Ottawa, Statistique Canada.
- Pryor, F. et Schaffer, D. (1997), « Wages and the University Educated: A Paradox Resolved », *Monthly Labor Review*, vol. 120, n° 7, p. 3 à 14.
- Pryor, F. et Schaffer, D. (1999), *Who's Not Working and Why: Employment, Cognitive Skills, Wages and the Changing U.S. Labor Market*, Cambridge, Cambridge University Press.
- Sicherman, N. (1991), « "Overeducation" in the Labor Market », *Journal of Labor Economics*, vol. 9, n° 2, p. 101 à 122.
- Statistique Canada (sans date), *International Adult Literacy Survey: Canadian Data: Microdata User's Guide*, Ottawa, Statistique Canada.
- Statistique Canada (1981), *Classification type des professions, 1980* (n° 12-565F au catalogue), Ottawa, Statistique Canada.
- Vahey, S. (2000), « The great Canadian training robbery: evidence on the returns to educational mismatch », *Economics of Education Review*, vol. 19, n° 2, p. 219 à 227.
- Yamamoto, K. et Kirsch, I. (1998), « Proficiency Estimation », chapitre 11 de Murray et coll. (1998), p. 179 à 191.

Annexe 1

Restrictions applicables aux échantillons et effets sur la taille de l'échantillon

Échantillon initial de l'EIAA : 2 423 hommes + 3 237 femmes = 5 660 personnes

Restrictions applicables à tous les échantillons : étudiants et immigrants exclus :

Hommes : 2 423 hommes – (315 étudiants, non-immigrants + 183 immigrants, non-étudiants + 20 immigrants, étudiants) = 1 905 hommes

Femmes : 3 237 femmes – (397 étudiantes, non-immigrantes + 264 immigrantes, non-étudiantes + 28 immigrantes, étudiantes) = 2 548 femmes

Restrictions applicables à tous les échantillons dans le corps du texte :

Les répondants doivent occuper un emploi (au moment de l'EIAA ou au cours de l'année précédente) :

Hommes : 1 905 hommes – 600 inactifs = 1 305 hommes

Femmes : 2 548 femmes – 1 243 inactives = 1 305 femmes

Les répondants doivent déclarer un code professionnel CTP 1980 valide à 4 chiffres (groupe de base) pour leur emploi principal :

Hommes : 1 305 hommes – 131 sans code CTP valide = 1 174 hommes

Femmes : 1 305 femmes – 34 sans code CTP valide = 1 271 femmes

Le code CTP 1980 doit correspondre à un groupe de base de la CCDP (FG et PPS valides) :

Hommes : 1 174 hommes – 64 sans correspondance = 1 110 hommes

Femmes : 1 271 femmes – 15 sans correspondance = 1 256 femmes

Les répondants doivent travailler à temps plein dans l'emploi principal :

Hommes : 1 110 hommes – 61 travailleurs à temps partiel = 1 049 hommes

Femmes : 1 256 femmes – 341 travailleuses à temps partiel = 915 femmes

Les deux dernières lignes correspondent à l'échantillon du tableau 1. Dans cet échantillon, il manquait à trois hommes et à une femme la variable « nombre d'années de scolarité accomplie » et il manquait à 11 hommes et à 14 femmes la variable « plus haut niveau d'études ».

Il s'agit également de l'échantillon initial pour les régressions des tableaux 2 et 3. Dans ces régressions, les valeurs manquantes des variables explicatives ont entraîné respectivement 124 autres observations manquantes chez les hommes et 79 autres chez les femmes (ce qui ramène la taille de l'échantillon à 925 hommes et 836 femmes).

Il s'agit également de l'échantillon initial pour les régressions des tableaux 4 et 5. Toutes les suppressions dues aux valeurs manquantes dans les tableaux 2 et 3 s'appliquent aux tableaux 4 et 5. En outre, 197 hommes et 198 femmes compris dans les échantillons des tableaux 2 et 3 n'ont pas déclaré leurs gains, ce qui a ramené la taille de l'échantillon à 852 hommes et 717 femmes. Enfin, on n'a pu construire la variable « utilisation des compétences en littératie au travail » à cause des réponses manquantes de 78 hommes et de 33 femmes, ce qui a ramené la taille finale de l'échantillon, dans les tableaux 4 et 5, à 774 hommes et 684 femmes.

En ce qui concerne les estimations présentées dans l'annexe 3, nous commençons par la population de 1 905 hommes + 2 548 femmes = 4 453 personnes qui ne sont ni des étudiants ni des immigrants. Dans le tableau A.1 (annexe 3), nous présentons les corrélations dont les échantillons varient à cause des réponses manquantes. Parmi les 4 453 personnes de l'échantillon initial, il ne manque pas de réponses en ce qui concerne la note en littératie, il y a 788 personnes qui n'ont pas rempli la deuxième partie du test et 27 personnes dont il manque le nombre d'années de scolarité accomplie. De ces 27 personnes, 14 n'ont pas rempli la deuxième partie du test de lecture²². Sur le plan des corrélations, le plus petit groupe est donc celui de la corrélation du nombre d'années de scolarité avec le pourcentage de bonnes réponses ($4\,453 - 788 - (27 - 14) = 3\,652$). Il s'agit également de l'échantillon retenu pour les régressions présentées dans les tableaux A.2 et A.3 (annexe 3).

Annexe 2

Questions concernant la forme fonctionnelle de l'équation des gains

Non-linéarité du rendement de la scolarité

Comme le montrent les tableaux 1, 2 et 3, à mesure que le nombre d'années de scolarité accomplie augmente, la surinstruction augmente (la sous-instruction diminue). Supposons que le rendement de la scolarité diminue à mesure que le nombre d'années de scolarité augmente (diminution du rendement marginal). Supposons encore qu'on impose une spécification linéaire du rendement de la scolarité et qu'on inclue dans la spécification des mesures de la surinstruction et de la sous-instruction. Les coefficients estimatifs de la surinstruction et de la sous-instruction tiennent compte de la non-linéarité du rendement de la scolarité et ont ainsi tendance à montrer un rendement de la surinstruction et de la sous-instruction qui peut être factice.

Pour cette raison, les auteurs de recherches antérieures sur le rendement de la surinstruction et de la sous-instruction se sont penchés sur la non-linéarité du rendement de la scolarité. Hartog (2000, p. 135) cite plusieurs études dont les auteurs ont constaté un rendement croissant de la scolarité ou n'ont constaté aucune preuve de la non-linéarité du rendement de la scolarité.

Il en va autrement de l'échantillon que nous avons retenu de l'EIAA. Comme le montrent les tableaux 4 et 5, le terme linéaire est positif et le terme quadratique est négatif lorsque nous employons une spécification quadratique du rendement de la scolarité. Nous prenons donc explicitement en compte la non-linéarité du rendement de la scolarité en utilisant systématiquement une spécification quadratique de la scolarité.

Nous constatons néanmoins des effets statistiquement significatifs de la surinstruction et de la sous-instruction. Nous avons mis à l'essai des formes fonctionnelles des équations des gains avec des termes linéaires, quadratiques et cubiques du nombre d'années de scolarité. Lorsqu'on incluait dans la spécification des termes cubiques de la scolarité et des mesures linéaires de la surinstruction et de la sous-instruction, on continuait d'observer des effets

statistiquement significatifs de la surinstruction et de la sous-instruction, alors que le terme cubique (ou, à l'occasion, le terme linéaire) du nombre d'années de scolarité n'était pas statistiquement significatif.

Nous concluons qu'il y a, dans notre échantillon, des non-linéarités du rendement de la scolarité et du rendement de la surinstruction et de la sous-instruction. Évidemment, nous ne pouvons pas écarter complètement la possibilité que le rendement mesuré de la surinstruction et de la sous-instruction soit un artefact en raison de la forme fonctionnelle du rendement de la scolarité.

Biais dans la mesure de la surinstruction et de la sous-instruction

Considérons maintenant les effets de mesures biaisées de la surinstruction et de la sous-instruction. Supposons, par exemple, que nous sous-estimions régulièrement d'une année la scolarité requise. Supposons en outre que le modèle correct des gains comprenne des mesures de la surinstruction O et de la sous-instruction U avec leurs coefficients respectifs o et u . En utilisant O et U pour représenter les niveaux réels de la surinstruction et de la sous-instruction, nous pouvons reformuler le modèle de l'équation 3 comme suit :

$$3'. \ln Y = c + sS + rS^2 + [o(O-1) + od_o] + [u(U+1)-ud_u] + X'B + e$$

où d_o et d_u sont respectivement des variables indicatrices de la surinstruction et de la sous-instruction.

Notre sous-estimation de la scolarité requise nous porte à définir la surinstruction comme suit : $O' = \{S-(R-1), S-(R-1) > 0; 0$ dans le cas contraire} et la sous-instruction comme suit : $U' = \{(R-1)-S, (R-1)-S > 0; 0$ dans le cas contraire}. Aucun problème ne se pose si $o = -u$.

Si cette relation ne se vérifie pas, deux problèmes se posent. Premièrement, chez toutes les personnes correctement désignées surinstruites ($O' > 1$), la surinstruction mesurée O' dépasse d'une année la surinstruction réelle O , alors que chez toutes les personnes correctement désignées sous-instruites ($U' > 0$), la sous-instruction est également sous-estimée d'une année. Ce phénomène modifie la constante de la régression (selon une moyenne pondérée des estimations de u et de o dans la spécification correcte).

Le deuxième problème, plus grave, est le suivant : toutes les personnes pour lesquelles $1 > O' > 0$ sont désignées surinstruites, alors qu'elles sont en réalité sous-instruites. Ce phénomène a pour effet de biaiser les estimations de o , puisque l'effet estimatif de la différence entre la scolarité et la scolarité requise pour ces personnes est intégré à l'estimation de o plutôt qu'à celle de u .

Ces considérations nous ont incités à utiliser deux mesures de la surinstruction et de la sous-instruction. L'une d'elles, fondée sur la FG comme mesure de la scolarité requise, montre la moyenne de la surinstruction. L'autre, fondée sur la somme FG+PPS comme mesure de la durée de la formation nécessaire, montre la moyenne de la sous-instruction. Toutes deux donnent, sur le plan qualitatif, des résultats semblables dans le tableau 5. Ces résultats concordent avec les résultats habituels des estimations des modèles de la surinstruction et de la sous-instruction. Comme nous n'avons pas de moyen de savoir laquelle constitue une meilleure mesure de la scolarité requise, nous ignorons quelles sont les meilleures estimations ponctuelles des effets de la surinstruction et de la sous-instruction.

Tel qu'il est posé, le problème est que nous ignorons où finit la surinstruction et où commence la sous-instruction, mais notre forme fonctionnelle nous oblige à décider où passer de l'une à l'autre. Nous avons formulé d'autres estimations pour une spécification qui segmentait la différence $R-S$ et permettait différents coefficients pour chaque segment.

Si le modèle de l'équation 3 est exact, le rendement de R-S dans chaque segment de surinstruction devrait être $-o$, le rendement de R-S dans chaque segment de sous-instruction devrait être u et le rendement dans le segment où survient le passage de la surinstruction à la sous-instruction devrait se situer quelque part entre les deux.

Dans ces estimations, les valeurs élevées de R-S avaient des effets importants et statistiquement significatifs, alors que les valeurs faibles de R-S n'avaient habituellement pas de coefficients statistiquement significatifs. Nous n'avons pas pu cerner un segment où le passage de la surinstruction à la sous-instruction était survenu, car il y avait habituellement plusieurs segments adjacents où les valeurs de R-S étaient faibles et où le coefficient de R-S n'était pas statistiquement significatif. (Chez les hommes, par exemple, pour les mesures FG et FG+PPS de R, les coefficients n'étaient pas statistiquement significatifs pour les segments $[-2,-1)$, $[-1,0)$, $[0,1)$.)

D'après nous, ces résultats indiquent que la mesure de R-S comporte une erreur considérable, d'où un manque de précision dans l'estimation de ses effets pour les valeurs faibles de R-S. Il se peut que l'une de nos mesures de R-S soit biaisée, sinon les deux, mais les résultats de ces régressions ne nous permettent pas de cerner l'ampleur ni même le sens du biais.

La spécification de l'équation 3 pourrait présenter une autre difficulté : elle suppose que l'effet des différences entre la scolarité requise par l'emploi et la scolarité accomplie par la personne est linéaire par morceaux (jointes à 0). D'autres types de non-linéarité sont possibles. Nous avons mis cette hypothèse à l'essai en estimant des spécifications quadratiques dans R-S (et en estimant aussi le rendement de S sous forme quadratique). Dans cette spécification, les termes linéaires et quadratiques de S étaient toujours statistiquement significatifs.

Le terme quadratique dans (R-S) n'était statistiquement significatif que dans le cas des femmes, lorsque la mesure requise de la scolarité était FG+PPS. La dérivée partielle des gains à l'égard de R-S était alors positive pour $R-S < 12$ et négative par la suite.

Nous n'avons vu dans tous ces résultats aucune raison impérieuse de préférer l'une ou l'autre de ces formes fonctionnelles aux formes plus classiques que nous avons présentées dans les tableaux 4 et 5.

PUBLICATIONS ÉLECTRONIQUES DISPONIBLES À
www.statcan.ca



Annexe 3

La relation entre les études, la littératie et les notes en littératie

Dans les tableaux 3 et 5, nous avons inclus deux mesures des compétences en littératie, soit la note en littératie (moyenne des premières valeurs plausibles en compréhension de textes suivis, de textes schématiques et de textes au contenu quantitatif) et le pourcentage de bonnes réponses aux examens de lecture. Nous avons signalé que la note en littératie était conditionnée par les études, qu'il existait des différences entre les effets des deux mesures de la littératie sur la profession et que ces différences se traduisaient généralement par des effets plus faibles selon la mesure du pourcentage de bonnes réponses que selon la note en littératie. Comme, sur le plan qualitatif, les résultats des deux mesures étaient semblables, nous avons retenu la note en littératie, qu'on utilise plus couramment.

La présente annexe porte également sur la relation entre la littératie et les études et les effets du conditionnement de la note en littératie par des « variables de base ». Selon un examen de l'EIAA publié par le British Office of National Statistics pour le compte de la Commission européenne, « ...on craint que les relations observées entre les caractéristiques de base et les niveaux de compétence ne soient exagérées par la procédure de la théorie de la réponse d'item où l'on utilise les caractéristiques de base pour conditionner les estimations. On s'interroge sur la validité de l'utilisation de méthodes à plusieurs variables, par exemple dans les estimations des compétences, car on ignore dans quelle mesure la relation perçue est attribuable aux facteurs de l'analyse qui ont déjà été pris en compte dans la procédure de conditionnement et d'estimation. On se préoccupe surtout de la possibilité de découvrir des relations factices » (Carey et coll., 2000, p. 245) [traduction]. Dans le chapitre 7 du même rapport, Heady préconise d'utiliser les pourcentages de bonnes réponses, qui sont plus faciles à comprendre.

Construction des « valeurs plausibles » dans l'EIAA

Nous présentons ci-dessous un examen empirique des effets du conditionnement des notes en littératie par rapport à l'influence mesurée des études sur la littératie. Avant de présenter nos résultats, il convient d'expliquer comment ont été établies les notes en littératie. Ce qui suit est notre interprétation de Yamamoto et Kirsch (1998).

L'évaluation de la littératie menée dans le cadre de l'EIAA a produit cinq valeurs plausibles pour chacun des trois types de compétences en littératie : compréhension de textes suivis, de textes schématiques et de textes au contenu quantitatif. (On a également attribué à chaque répondant l'un de quatre niveaux pour chaque type de compétences en littératie, mais nous n'en tenons pas compte ici.) La mesure de la littératie que nous appelons la note en littératie est la moyenne de la première valeur plausible pour chacun des trois types de compétences en littératie. (Rien ne semble indiquer que la première valeur plausible soit plus plausible que les autres.)

Yamamoto et Kirsch (1998, p. 180) expliquent très clairement pourquoi on adopte la méthode des valeurs plausibles. « On ne saurait trop insister sur le fait que les valeurs plausibles ne sont pas des notes de test attribuées à des personnes dans le sens courant. En d'autres termes, elles se distinguent des estimations usuelles des capacités et du niveau d'instruction qui sont, d'une certaine façon, optimales pour chaque répondant (estimations selon la méthode du maximum de vraisemblance, selon la méthode de Bayes, etc.). Les estimations ponctuelles qui sont optimales pour des répondants individuels peuvent produire des estimations résolument non optimales (incohérentes) des caractéristiques d'une population (Little et Rubin, 1983). On construit des valeurs plausibles expressément pour fournir des estimations cohérentes des effets de la population même si, en général, il ne s'agit pas d'estimations non biaisées des compétences des personnes auxquelles elles sont associées (Mislevy, Beaton, Kaplan et Sheehan, 1992) » [traduction]²³. Comme notre étude a utilisé (à l'instar d'autres travaux) ces notes comme mesures individuelles, il semble indiqué d'examiner plus en détail comment elles ont été produites.

Dans la méthodologie décrite, chaque type de compétences en littératie est considéré comme une variable latente non observée. Les « valeurs des compétences sur l'échelle » (niveaux latents de littératie) n'étant pas observées pour les répondants échantillonnés, il n'est pas possible de construire une statistique t (par exemple, la moyenne d'une sous-population) des compétences sur l'échelle. On estime plutôt la valeur probable de t en fonction des données réellement observées, $t^*(x,y)$. Ici, les variables x sont des réponses d'item et les variables y sont des variables de base. Selon Yamamoto et Kirsch (1998, p. 180), « il est possible de calculer approximativement t^* en effectuant des prélèvements aléatoires dans la distribution conditionnelle des compétences sur l'échelle en fonction des réponses d'item x_j , des variables de base y_j et du paramètre modèle du répondant échantillonné j . Ces valeurs sont appelées imputations dans la documentation sur l'échantillonnage et valeurs plausibles dans l'EIAA » [traduction]²⁴. On répète le processus afin d'estimer l'incertitude due au fait que les compétences sur l'échelle ne sont pas observées.

Les valeurs plausibles pour chaque répondant j sont prélevées dans la distribution des valeurs latentes de littératie sur l'échelle, conditionnelles par rapport à x_j , y_j , une matrice de coefficients de régression G , et S , une matrice de variance ordinaire pour les résidus. On calcule cette distribution conditionnelle d'après le « produit sur les échelles des vraisemblances indépendantes induites par les réponses aux items [les x_j] à l'intérieur de chaque échelle » et « la densité dépendant de plusieurs variables des compétences sur les échelles, conditionnelles par rapport à la valeur observée y_j des réponses de base et des paramètres G et S . Dans ce calcul, les estimations des paramètres d'item sont fixes et considérées comme des valeurs de population... » (Yamamoto et Kirsch, 1998, p. 181) [traduction].

On a supposé que la distribution dépendant de plusieurs variables des trois compétences était une distribution normale à plusieurs variables « avec une variance ordinaire, S , et une moyenne donnée par un modèle linéaire avec des paramètres de pente, G , d'après les premières composantes principales approximatives de quelques centaines d'effets principaux choisis et les interactions à double entrée du vecteur complet des variables de base. Les variables de base comprenaient notamment le sexe, l'origine ethnique, la langue de l'interview, le niveau d'instruction du répondant, celui des parents, la profession et les habitudes de lecture. D'après la méthode des composantes principales, on a choisi les

composantes représentant 99 % de la variance présente dans les données. Ces composantes principales sont les variables de conditionnement utilisées dans l'analyse » (Yamamoto et Kirsch, 1998, p. 181) [traduction]. Les auteurs décrivent ensuite les algorithmes utilisés pour estimer G et S et la façon d'utiliser les valeurs plausibles pour formuler des estimations des statistiques des compétences latentes sur l'échelle.

Estimation des effets du conditionnement

Ici, deux observations s'imposent. Premièrement, toutes les variables de base (tout ce qui n'est pas une réponse au test) sont utilisées pour construire les mesures des compétences en littératie. On peut donc se demander dans quelle mesure la relation observée entre une « variable de base » et les valeurs plausibles pour un même répondant est attribuable à ce conditionnement. Deuxièmement, il serait très difficile de retracer l'influence d'une « variable de base » sur les valeurs plausibles individuelles.

Par conséquent, nous ne tentons pas de retracer l'influence des études sur la note en littératie (moyenne des trois premières valeurs plausibles) au moyen du conditionnement des valeurs plausibles par rapport aux variables de base. Nous comparons plutôt les effets des études sur la note en littératie et sur une autre mesure des compétences en littératie (le pourcentage de bonnes réponses) dans des modèles de régression simple. Nous montrons également que le niveau d'études a un effet important sur la note en littératie, même lorsqu'on tient compte du pourcentage de bonnes réponses.

Nous mesurons la proportion de bonnes réponses uniquement pour les personnes qui, d'après le code attribué, ont rempli le questionnaire de lecture. Sont donc exclues les personnes qui ont rempli le questionnaire de base mais qui ont refusé l'évaluation des compétences en littératie, celles qui n'ont pas répondu correctement à au moins deux items du questionnaire de sélection (livret principal) et celles qui ont refusé de terminer le test d'évaluation. Ainsi, la population dont nous avons calculé le pourcentage de bonnes réponses exclut toutes les personnes pour lesquelles il a fallu des méthodes d'imputation pour calculer les valeurs plausibles et toutes celles dont les compétences en littératie étaient si faibles qu'elles n'ont pas pu aller au-delà du niveau du questionnaire de sélection.

Nous avons calculé la proportion de bonnes réponses comme suit : nombre de bonnes réponses / (nombre de bonnes réponses + nombre de mauvaises réponses + nombre de questions refusées ou laissées sans réponse). Ce calcul a pour effet de traiter tous les items refusés et omis comme des mauvaises réponses. Il tenait compte du questionnaire de sélection et des questions contenues dans les livrets du test. On n'a retenu que la note du premier correcteur (l'EIAA a utilisé deux correcteurs pour 20 % des répondants).

La proportion moyenne de bonnes réponses (pour les personnes comprises dans notre échantillon et pour lesquelles nous avons pu effectuer ce calcul) était de 0,683. Selon Yamamoto et Kirsch (1998, p. 83), les pourcentages moyens de bonnes réponses sur les échelles de compréhension de textes suivis, de textes schématiques et de textes au contenu quantitatif s'établissaient respectivement à 70 %, 74 % et 67 % pour le Canada (testé en anglais) et à 64 %, 69 % et 60 % pour le Canada (testé en français).

Pour le même échantillon, la valeur moyenne de la mesure de la note en littératie (moyenne des premières valeurs plausibles) était de 282. Toujours selon Yamamoto et Kirsch, les valeurs moyennes des compétences en littératie sur les échelles de compréhension de textes suivis, de textes schématiques et de textes au contenu quantitatif s'établissaient respectivement à 284, 284 et 286 pour le Canada/testé en anglais et à 264, 265 et 266 pour le Canada/testé en français. Dans notre échantillon de non-étudiants et de non-immigrants, la valeur moyenne de la mesure de la note en littératie est de 272. La valeur inférieure de la note en littératie pour cet échantillon élargi témoigne probablement des faibles valeurs plausibles attribuées aux personnes qui ne sont pas allées au-delà des questions de sélection et à bon nombre de celles qui ont refusé de subir le test.

Le tableau A.1 présente les corrélations entre la note en littératie, la proportion de bonnes réponses et le nombre d'années de scolarité. Il y a deux points à signaler. Premièrement, la corrélation entre la note en littératie et la proportion de bonnes réponses (0,88) est très élevée. Deuxièmement, la corrélation entre la note en littératie et la scolarité est beaucoup plus élevée (0,70) que la corrélation entre la proportion de bonnes réponses et la scolarité (0,58).

Tableau A.1 Corrélations entre la note en littératie, la proportion de bonnes réponses et le nombre d'années de scolarité

	Note en littératie	Proportion de bonnes réponses	Années de scolarité
Note en littératie	—	0,88 n = 3 665	0,70 n = 4 426
Proportion de bonnes réponses	—	—	0,58 n = 3 652

Plusieurs auteurs ont tenté d'estimer l'influence de la scolarité sur les compétences en littératie en utilisant la note en littératie ou une variable semblable comme mesure des compétences en littératie. Le tableau A.2 présente les résultats de deux ensembles d'estimations de régression avec des mesures des compétences en littératie comme variable dépendante. Dans le premier ensemble, la variable dépendante est la note en littératie; dans le second, c'est la proportion de bonnes réponses. Les deux variables dépendantes ont été normalisées de sorte que la moyenne soit 0 et l'écart-type, 1. On peut ainsi comparer directement les coefficients estimatifs des deux régressions.

Il ne faut pas interpréter ces modèles comme des tentatives de mesure de la relation structurelle entre la scolarité et la littératie. Dans une enquête transversale comme l'EIAA, les compétences en littératie et la scolarité accomplie sont déterminées simultanément. La corrélation entre la scolarité et les compétences en littératie peut exister, en tout ou en partie, parce que les compétences en littératie masquent la scolarité complémentaire, et non parce que la scolarité complémentaire produit les compétences en littératie²⁵. À notre avis, c'est en effectuant des mesures répétées des compétences en littératie des personnes à différentes étapes de leur scolarité qu'on mesure le mieux les effets structurels de la scolarité sur la littératie.

Tableau A.2 Estimations de régression de la note en littératie et de la proportion de bonnes réponses

	Note en littératie		Proportion de bonnes réponses	
Constante	-0,0386	(,0019)	-0,0365	(,0020)
Années de scolarité	,0028	(,0001)	,0025	(,0001)
Femmes	,0005	(,0004)	,0011	(,0004)
Âge	,0004	(,0001)	,0006	(,0001)
Âge ²	-6x10⁻⁶	(8x10⁻⁷)	-8x10⁻⁶	(8x10⁻⁷)
Test subi en français	-,0032	(,0005)	-,0038	(,0005)
Test subi dans une autre langue que la langue maternelle	-,0039	(,0008)	-,0033	(,0008)
r² corrigé selon le degré de liberté	,45		,41	

n = 3 652

Entre parenthèses : erreurs-types estimatives.

Variables dépendantes normalisées de sorte que la moyenne = 0 et l'écart-type = 1. En gras et en italiques : chiffres significatifs à 1 %. En gras seulement : chiffres significatifs à 5 %. En italiques seulement : chiffres significatifs à 10 %.

Dans les deux modèles du tableau A.2, les coefficients des variables explicatives sont très semblables, à une exception près. L'effet du sexe sur la note en littératie n'est pas statistiquement significatif, contrairement à l'effet du sexe sur la proportion de bonnes réponses. Toutefois, cet effet n'est pas important, puisqu'il équivaut à moins de la moitié de l'effet d'une année de scolarité.

Le fait d'écrire le test en français ou dans une langue autre que la langue maternelle a des effets négatifs et statistiquement significatifs tant sur la note en littératie que sur le pourcentage de bonnes réponses. Presque toutes les personnes ayant écrit le test dans une langue autre que leur langue maternelle étaient des francophones qui ont subi le test en anglais.

S'il existe une certaine différence entre les coefficients des termes relatifs à l'âge, les différences entre les dérivées partielles des variables dépendantes à l'égard de l'âge sont faibles dans la plage d'âge de 20 ans à 60 ans (la plus grande est de $8,3 \times 10^{-5}$ à l'âge de 60 ans). La dérivée partielle de la note en littératie à l'égard de l'âge est positive jusqu'à l'âge de 33,5 ans et négative par la suite. La dérivée partielle de la proportion de bonnes réponses à l'égard de l'âge est positive jusqu'à l'âge de 34,1 ans et négative par la suite.

Dans ces deux modèles, toutes les variables explicatives font partie des variables de base utilisées pour établir la note en littératie. Dans ces régressions, rien n'indique que ces variables influencent les notes en littératie plus qu'elles n'influencent le pourcentage de bonnes réponses. En particulier, le conditionnement des notes en littératie ne semble pas avoir exagéré la relation entre la scolarité accomplie et les compétences en littératie.

Les corrélations présentées dans le tableau A.1 montrent que la variable « note en littératie » est étroitement liée au pourcentage de bonnes réponses. Nous examinons maintenant la relation entre le pourcentage de bonnes réponses et la note en littératie dans un contexte à plusieurs variables. Le tableau A.3 présente les résultats de la régression de la note en littératie sur une courbe cubique en fonction de la proportion de bonnes réponses, des variables explicatives comprises dans le tableau A.2 et d'une série de variables binaires représentant laquelle des sept versions du test la personne a subie.

Dans cette régression, les variables explicatives qui influent sur la note en littératie uniquement parce qu'elles influencent le pourcentage de bonnes réponses ne devraient pas avoir d'effets statistiquement significatifs. Nous avons entré le pourcentage de bonnes réponses sous forme de courbe cubique parce que les trois termes de l'expression cubique sont statistiquement significatifs et que nous cherchons à expliquer aussi complètement que possible l'effet des bonnes réponses sur la note en littératie. Les effets statistiquement significatifs des « variables de base » devraient constituer une approximation de leur effet par le biais du conditionnement par rapport aux notes en littératie des personnes comprises dans cet échantillon²⁶.

Dans la régression du tableau A.3, le sexe, les termes relatifs à l'âge et la langue du test n'ont pas d'effets statistiquement significatifs sur la note en littératie. L'effet du nombre d'années de scolarité accomplie est statistiquement significatif. Chaque année de scolarité supplémentaire devrait augmenter la note en littératie d'un peu plus de deux points (produit du coefficient par 3154, soit l'écart-type de la note en littératie dans l'échantillon avant normalisation). Le fait d'écrire le test dans une langue autre que la langue maternelle a également un effet statistiquement significatif : il réduit la note en littératie de quatre points en moyenne²⁷.

Tableau A.3 Estimations de régression relatives aux déterminants de la note en littératie

	Note en littératie	
Constante	<i>-,0103</i>	<i>(,0011)</i>
Années de scolarité	<i>,0007</i>	<i>(4x10⁻⁵)</i>
Femmes	- ,0003	(,0002)
Âge	-1x10 ⁻⁵	(4x10 ⁻⁵)
Âge ²	-9x10 ⁻⁸	(4x10 ⁻⁷)
Test subi en français	,0002	(,0003)
Test subi dans une autre langue que la langue maternelle	<i>-,0013</i>	<i>(,0004)</i>
Proportion de bonnes réponses	<i>,8289</i>	<i>(,0133)</i>
Proportion de bonnes réponses ²	<i>17,30</i>	<i>(,6869)</i>
Proportion de bonnes réponses ³	<i>277,1</i>	<i>(20,3500)</i>
Test 1	<i>-,0025</i>	<i>(,0004)</i>
Test 2	<i>,0019</i>	<i>(,0004)</i>
Test 3	<i>-,0023</i>	<i>(,0004)</i>
Test 4	<i>-,0024</i>	<i>(,0004)</i>
Test 5	<i>-,0014</i>	<i>(,0004)</i>
Test 6	<i>-,0022</i>	<i>(,0004)</i>
Test 7	—	—
r² corrigé selon le degré de liberté	0,84	
n = 3 652		

Variables dépendantes normalisées de sorte que la moyenne = 0 et l'écart-type = 1.

Entre parenthèses : erreurs-types estimatives.

En gras et en italiques : chiffres significatifs à 1 %. En gras seulement : chiffres significatifs à 5 %. En italiques seulement : chiffres significatifs à 10 %.

La version du test complété par le répondant a des effets statistiquement significatifs sur la note en littératie. Ce phénomène est certainement attribuable aux différences dans la difficulté des versions, dont la proportion de bonnes réponses ne tient pas compte. On pourrait améliorer la mesure de la proportion de bonnes réponses en la rajustant en fonction de la version du test.

La valeur r^2 dans le modèle (0,84) montre que les variables comprises dans l'équation expliquent la plus grande partie de la variance de la note en littératie.

Quelques conclusions

Deux des variables de base comprises dans le modèle du tableau A.3 ont des effets statistiquement significatifs sur la note en littératie lorsqu'on tient compte du pourcentage de bonnes réponses aux items du test et de l'effet de la version du test. L'une de ces variables est celle, couramment utilisée, du nombre d'années de scolarité accomplie. On doit conclure que pour des variables importantes du fichier de l'EIAA, il existe effectivement une « possibilité de découvrir des relations factices » (Carey et coll., 2000, p. 245) ou, du moins, celle d'exagérer l'ampleur des effets sur la relation.

Les effets des variables de base présentées dans le tableau A.3 sont faibles. De plus, les résultats présentés dans le tableau A.2 révèlent que diverses variables explicatives ont des effets très semblables sur la note en littératie et sur la proportion de bonnes réponses. Lorsqu'on estime les relations structurelles, il y a sans doute très peu de risques à utiliser les valeurs plausibles comme mesures des compétences en littératie si on limite l'échantillon aux personnes qui ont subi le test de lecture. Toutefois, les chercheurs continueront sans doute d'hésiter à utiliser les valeurs plausibles, surtout à cause des différences dans les résultats entre les valeurs plausibles et la proportion de bonnes réponses, comme celles qu'on trouve dans les tableaux 4 et 5.

On peut se demander pourquoi les valeurs plausibles sont les seules mesures à valeur continue des compétences en littératie qui figurent dans les fichiers de microdonnées à grande diffusion. L'objet d'un fichier de microdonnées consiste, pour une large part, à permettre la recherche au niveau individuel. Toutefois, Yamamoto et Kirsch (1998, p. 180) affirment explicitement que les valeurs plausibles, « en général, ne sont pas des estimations non biaisées des compétences des personnes auxquelles elles sont associées » [traduction]. À notre avis, les fichiers de microdonnées à grande diffusion devraient contenir des estimations non biaisées des compétences des personnes.

La proportion de bonnes réponses constitue une méthode rudimentaire pour construire une telle mesure. Cette mesure pourrait être produite dans une version dans laquelle l'effet de la version du test serait corrigée. Il serait sans doute préférable d'utiliser les cotes de difficulté des items de chaque échelle de compétences en littératie pour construire des compétences individuelles pour chaque type de compétences en littératie. Les cotes de difficulté des items figurent dans l'annexe du chapitre 11 (Yamamoto et Kirsch, 1998) dans Murray et coll. (1998).

On ne peut construire une mesure de ce type que pour les personnes ayant complété le test de lecture. À notre avis, le principal danger de l'utilisation des valeurs plausibles ne tient pas aux faibles effets des variables de base que nous avons décelés, mais plutôt à l'utilisation de la méthodologie pour produire des « valeurs plausibles » pour des personnes qui n'ont pas subi le test de lecture.

Yamamoto et Kirsch (1998, p. 190) justifient le traitement de la non-réponse en invoquant la nécessité d'estimations non biaisées des compétences pour des populations et des sous-populations. On peut mettre en doute la pertinence de ces méthodes. Par exemple, on présume que ces méthodes s'appliquent uniquement aux personnes ayant rempli le questionnaire de base. Pourquoi le refus de l'ensemble de l'évaluation (y compris des items de base) serait-il moins susceptible de produire des résultats biaisés que le refus de l'ensemble ou d'une partie des tests de lecture?

Quoi qu'on pense des méthodes utilisées à l'égard de la nécessité de produire des estimations non biaisées pour des populations, il n'y a aucune raison de penser que les valeurs plausibles produites par ces méthodes sont des mesures exactes pour des personnes. Toutefois, elles sont considérées comme telles par les chercheurs qui utilisent les fichiers de microdonnées à grande diffusion. À notre connaissance, il n'existe aucune façon de distinguer, dans ces fichiers, les personnes dont les valeurs plausibles reflètent à la fois les notes obtenues au test et les variables de base, de celles dont les valeurs plausibles sont entièrement fondées sur les variables de base.

À notre avis, le fait d'inclure, au niveau individuel, des estimations de personnes dont les valeurs plausibles sont entièrement fondées sur les variables de base constitue la démarche la plus susceptible de faire « découvrir des relations factices ». Nous croyons qu'à tout le moins, les fichiers de données à grande diffusion devraient permettre aux chercheurs de distinguer les valeurs plausibles fondées uniquement sur les « caractéristiques de base » de celles qui reflètent les résultats obtenus au test. De préférence, les données à grande diffusion devraient contenir une mesure valide des compétences individuelles, mais uniquement des personnes ayant subi le test de lecture.

PUBLICATIONS ÉLECTRONIQUES DISPONIBLES À
www.statcan.ca



Notes

1. Précisons que R, O et U décomposent S. Si tous les coefficients sont égaux, il existe donc un rendement égal de tous les niveaux de scolarité, sans égard aux exigences scolaires de l'emploi. Selon Hartog, divers auteurs ont établi que ce résultat n'était pas attribuable à la non-linéarité du rendement de la scolarité. Ces points sont abordés en détail dans l'annexe 2.
2. Selon Vahey (2000, p. 220), la *National Survey of Class Structure and Labour Process* est une enquête transversale au cours de laquelle environ 3 000 répondants ont été interviewés en 1982 par Canada Facts. Leurs réponses ont été décodées et transférées sur bande magnétique au Département de sociologie et d'anthropologie de l'Université Carleton. L'enquête ayant recueilli des données sur le plus haut niveau de scolarité atteint par le répondant et sur le niveau de scolarité exigé par l'emploi, les mesures de la surinstruction et de la sous-instruction utilisées par Vahey sont des indicateurs oui/non, plutôt que la mesure continue que nous avons décrite plus haut. D'autres auteurs utilisent aussi des indicateurs dichotomiques de la surinstruction et de la sous-instruction.
3. Cette enquête ne recueillait pas de données sur les gains des répondants. Si les répondants ont travaillé durant l'année précédente, Charette et Meng (1998) utilisent le revenu de cette année-là.
4. Selon la définition utilisée, le « défaut de complémentarité » consiste à travailler en dehors de la catégorie des « professions nécessitant des qualifications ».
5. L'annexe 1 présente le détail des restrictions applicables aux échantillons et de leur effet sur la taille de l'échantillon. Les restrictions applicables aux échantillons (notamment celle qui concerne les travailleurs à temps plein) sont probablement non aléatoires à l'égard des compétences en littératie.
6. Il existe de légères différences entre la CCDP et la CTP de 1980 au niveau des groupes de base. Nous avons réaffecté les groupes de base de la CCDP à ceux de la CTP de 1980 en fonction de la correspondance entre la CCDP et le Manuel de codage des professions (MCP), décrite dans le manuel de la CTP de 1980 (Statistique Canada, 1981, p. 14 à 18).
7. Ces niveaux de FG et de PPS en années pour les professions détaillées ont été établis par Wayne Roth, Développement des ressources humaines Canada, d'après les niveaux de FG et de PPS figurant dans la CCDP, au moyen de la méthode décrite dans le texte.
8. Si l'on demande à une centaine d'employés ayant le même titre de poste et le même employeur quelles sont les exigences scolaires de leur emploi, leurs réponses seront sans doute différentes, même si les exigences de l'emploi sont exactement les mêmes. De plus, rien ne garantit que la moyenne des réponses correspondra aux exigences réelles.
9. Les auteurs attribuent ce concept à J.B. Knight. L'idée du profil études-salaire (ou du profil compétences-salaire) d'une profession semble introduire une certaine incertitude à l'égard de ce qui constitue les « exigences scolaires » ou les « exigences professionnelles » de la profession. Il est possible que ces exigences représentent un niveau minimal de compétences nécessaires pour évoluer au sein de la profession. Borghans et de Grip envisagent une correspondance optimale entre les compétences et la profession, fondée sur la relation entre le profil études-salaire déterminé sur le marché du travail et le profil études-productivité propre à la profession. Au sein d'une profession, les travailleurs surinstruits seraient plus productifs et mieux payés que les travailleurs moins instruits.
10. Green et Riddell (2001) présentent cet argument de manière plus formelle dans le contexte d'une équation des gains comprenant des coefficients de la scolarité et de la littératie.
11. Les termes quadratiques de la scolarité et de l'expérience de travail n'ayant pas eu d'effets statistiquement significatifs dans les régressions préliminaires, ils ne figurent pas dans les modèles présentés dans le tableau 2.
12. Par contre, dans l'équation des gains, le coefficient de l'expérience de travail estimé de cette manière est plus grand chez les femmes que chez les hommes (tableau 4). Si la mesure de l'expérience est biaisée par excès chez les femmes, les résultats de ce biais sont amplement compensés par le rendement supérieur de l'expérience à l'égard des gains des femmes. Il serait manifestement préférable de compter sur une mesure directe de l'expérience de travail et sur une mesure de l'ancienneté professionnelle, mais l'EIAA ne fournit ni l'une ni l'autre.
13. Aux étapes préliminaires de notre étude, nous avons utilisé des variables provinciales dans la présente section et

dans les estimations salariales de la section suivante. Les différences entre les coefficients estimatifs des provinces à l'intérieur d'une région étaient faibles et l'échantillon de certaines provinces était restreint. C'est pourquoi nous avons préféré présenter les coefficients des variables régionales.

14. Le fichier de données de l'EIAA fournit cinq notes appelées première valeur plausible, deuxième valeur plausible, etc., pour chacun des trois types de compétences en littératie mesurées (compréhension de textes suivis, de textes schématiques et de textes au contenu quantitatif). Chaque note s'inscrit sur une échelle de compétences allant de 1 (minimum) à 500 (maximum). À l'aide d'arguments qui nous semblent convaincants, Green et Riddell (2001) préconisent d'utiliser la somme ou la moyenne des notes correspondant aux trois types de compétences en littératie. Ils utilisent la moyenne (sans doute des premières valeurs plausibles). Osberg (2001) utilise la somme des premières valeurs plausibles correspondant aux trois notes. Dans l'échantillon que nous avons prélevé pour les régressions présentées à l'annexe 3 (n = 3 652), la moyenne de la note moyenne en compétences en littératie est de 282, la moyenne de la proportion de bonnes réponses est de 0,68 et la corrélation de Pearson de ces deux mesures est de 0,88.
15. Les résultats pour les autres variables comprises dans le tableau 2 sont très semblables à ceux qui figurent dans ce tableau et ne figurent pas dans le tableau 3.
16. Nous évaluons le rendement marginal de la scolarité et celui de l'expérience de travail en prenant les dérivées partielles de notre modèle estimatif à l'égard de ces variables. Ainsi, le rendement marginal d'une année de scolarité est le coefficient du terme linéaire de la scolarité plus deux fois le coefficient du terme quadratique de la scolarité multiplié par le nombre d'années de scolarité; le calcul est le même pour l'expérience de travail.
17. Comme les coefficients des autres variables figurant dans le tableau 4 variaient très peu dans les régressions présentées dans le tableau 5, ils ne figurent pas dans ce tableau.
18. Sicherman (1991) se penche également sur la forme fonctionnelle de l'équation 2 et sur les identités qui lient les coefficients des formes fonctionnelles des équations 1 et 2. L'annexe 2 aborde plus longuement les questions de forme fonctionnelle.
19. Les activités de lecture constituent les réponses à la question « À quelle fréquence lisez-vous ou utilisez-vous de l'information provenant des types de document suivants dans le cadre de votre emploi principal? » : 1. Lettres ou notes de service, 2. Rapports, articles, revues ou bulletins, 3. Manuels ou livres de référence, y compris des catalogues, 4. Graphiques ou schémas, 5. Factures ou comptes, 6. Modes d'emploi de médicaments ou d'autres produits, recettes. Nous n'avons pas utilisé le septième item, « documents rédigés dans une langue autre que le français (ou l'anglais) ». Les activités d'écriture constituent les réponses à la question « À quelle fréquence rédigez-vous ou remplissez-vous chacun des types de document suivants dans le cadre de votre emploi principal? » : 1. Lettres ou notes de service, 2. Formulaires ou documents tels que factures, comptes ou budgets, 3. Rapports ou articles, 4. Devis ou fiches techniques. Selon la fréquence, nous avons noté les réponses à chaque item de 0 (rarement ou jamais) à 4 (chaque jour).
20. Nous avons tenté de déterminer si les effets des compétences en littératie et de l'utilisation des compétences en littératie au travail sur les gains se manifestaient uniquement lorsque la personne qui possède ces compétences les utilise. Lorsqu'on inclut une variable « interaction » dans la régression des gains, son coefficient est positif et elle présente un haut niveau de signification statistique. Lorsqu'on inclut à la fois la mesure des compétences en littératie, celle de l'utilisation des compétences en littératie et leur interaction, il est difficile d'interpréter les résultats. Nous avons donc choisi de présenter la forme fonctionnelle qui comprend les mesures des compétences en littératie et de l'utilisation des compétences en littératie, mais non leur interaction.
21. Plus exactement, comme le rendement de la scolarité diminue à mesure que la scolarité augmente, cette observation n'est vraie que jusqu'à la valeur des années de scolarité où le rendement marginal de la scolarité équivaut à la pénalité liée à la surinstruction. Chez les hommes, ce seuil est atteint avec 18 années de scolarité dans le modèle de la colonne 2 et avec 16 années de scolarité dans le modèle de la colonne 3. Chez les femmes, les niveaux correspondants sont de 18 et de 17 années de scolarité.
22. La première partie du test comportait six questions de sélection. Les personnes n'ayant pas réussi à répondre correctement à au moins deux de ces six questions n'ont pas subi la deuxième partie du test, qui comportait un plus grand nombre de questions présentant divers niveaux de difficulté.
23. Dans Murray et coll. (1998), ces sources sont citées comme suit : Little, R.J.A. et Rubin, D.B., 1983, « On jointly estimating parameters and missing data », *American Statistician*, n° 37, p. 218 à 220, et Mislevy, R.J., Beaton, A., Kaplan, B.A. et Sheehan, K., 1992, « Estimating population characteristics from sparse matrix samples of item responses », *Journal of Educational Measurement*, vol. 29, n° 2, p. 133 à 161.
24. La statistique t des compétences sur l'échelle, mentionnée dans la citation, représente toute statistique des compétences sur l'échelle, et non une statistique t (de Student) particulière.
25. On aurait un exemple de cet effet si les niveaux des types de compétences en littératie évaluées dans le cadre de l'EIAA étaient déjà complètement déterminés à la fin des études secondaires. Les diplômés de l'enseignement secondaire possédant de hauts niveaux de compétences en littératie seraient plus portés à poursuivre des études postsecondaires. On observerait alors une corrélation entre la scolarité postsecondaire et les compétences en littératie, mais cela ne supposerait pas que les études postsecondaires mènent à l'acquisition d'autres compétences en littératie.
26. Rappelons qu'il s'agit d'un échantillon de personnes ayant subi le test de lecture. Les effets des variables de base sur les notes en littératie des personnes qui n'ont pas subi le test seraient sans doute beaucoup plus importants.
27. Il se trouve que presque tous les francophones ayant subi le test de lecture en anglais provenaient de l'échantillon spécial franco-ontarien. Par conséquent, l'effet peut être ici attribuable à une autre caractéristique de cet échantillon, et non au fait d'avoir subi le test dans une autre langue que la langue maternelle.

PUBLICATIONS ÉLECTRONIQUES DISPONIBLES À
www.statcan.ca



Enquête internationale sur l'alphabétisation des adultes

Série de monographies

L'Enquête internationale sur l'alphabétisation des adultes (EIAA) est une initiative de sept pays qui s'est tenue à l'automne 1994. Son objectif visait à établir des profils d'alphabétisme comparables sans égard aux frontières nationales, linguistiques et culturelles. Les vagues successives de l'enquête incluent maintenant près de 30 pays partout dans le monde.

La série de monographies comprend des études détaillées découlant de la base de données de l'EIAA, qui ont été effectuées par des spécialistes de l'alphabétisme au Canada et aux États-Unis. Les recherches sont principalement financées par Développement des ressources humaines Canada. Les monographies mettent l'accent sur les questions actuelles en matière de politiques et portent sur des sujets comme la formation continue, la correspondance et la non-correspondance entre les capacités de lecture et le milieu de travail, les capacités de lecture et l'état de santé des personnes âgées, l'alphabétisme et la sécurité économique, pour ne nommer que ceux-là.

ISBN 0-66096-653-0



9 780660 966533