
L'Enquête sur la littératie et les compétences des adultes, 2003 et 2008

Fichier de microdonnées à grande diffusion

Manuel de l'utilisateur

Table des matières

1.0	Introduction	1
2.0	Aperçu de l'enquête	2
3.0	Objectifs de l'enquête	3
4.0	Concepts et définitions.....	4
4.1	Définir et mesurer les compétences : aperçu	4
4.2	Comprendre ce qui a été mesuré lors de l'Enquête internationale sur l'alphabétisation et les compétences des adultes	5
4.2.1	Introduction	5
4.2.2	Échelonner les tâches de littératie, de numératie et de résolution de problèmes dans le cadre de l'ELCA	6
4.3	Mesurer la compréhension de textes suivis et de textes schématiques dans le cadre de l'ELCA	7
4.3.1	Définir les caractéristiques des tâches.....	9
4.3.2	Type d'appariement	11
4.3.3	Type d'information demandée.....	11
4.3.4	Plausibilité des éléments de distraction	12
4.3.5	Caractériser les tâches de compréhension de textes suivis	12
4.3.6	Caractériser les tâches de compréhension de textes schématiques	17
4.4	Mesurer la numératie dans le cadre de l'ELCA	21
4.4.1	Définir les caractéristiques des tâches.....	23
4.4.2	Vie quotidienne	23
4.4.3	Travail	24
4.4.4	Vie sociale ou collective	24
4.4.5	Formation complémentaire	24
4.4.6	Quantité et nombre	25
4.4.7	Dimension et forme	26
4.4.8	Structure, fonctions et relations.....	26
4.4.9	Données et hasard.....	27
4.4.10	Variation	27
4.4.11	Caractériser les tâches de numératie	28
4.5	Mesurer la résolution de problèmes dans le cadre de l'ELCA	32
4.5.1	Définir les caractéristiques des tâches.....	34
4.5.2	Caractériser les tâches de résolution de problèmes	36
4.6	Conclusion.....	40
4.6.1	Quelques critères analytiques.....	41
5.0	Méthodologie de l'enquête	47
5.1	Plan d'évaluation	47
5.2	Population cible et base de sondage	48
5.3	Plan d'échantillonnage	50
5.4	Taille de l'échantillon	55
5.5	Collecte de données.....	55
5.6	Notation des tests.....	58
5.7	Réponse à l'enquête et pondération	61

Table des matières (suite)

6.0	Méthodes d'enquête et traitement des données	64
6.1	Introduction.....	64
6.2	Modèles de méthodes, de manuels et d'instruments	64
6.2.1	Questions de base	65
6.2.2	Items de tâche.....	65
6.2.3	Codage normalisé des non-réponses	66
6.3	Notation	68
6.3.1	Deuxième notation à l'intérieur d'un pays	68
6.3.2	Deuxième notation entre les pays.....	69
6.4	Saisie, traitement et codage des données	71
6.5	Variables dérivées	72
7.0	Totalisation et analyse des données	73
7.1	Utilisation de coefficients de pondération dans la totalisation	73
7.2	Définitions de l'estimation catégorique et de l'estimation quantitative	73
7.2.1	Totalisation des estimations catégoriques	74
7.2.2	Totalisation des estimations quantitatives.....	75
7.3	Estimations des niveaux de compétence	75
7.4	Arrondissement des estimations	76
8.0	Qualité des données	78
8.1	Erreurs d'échantillonnage.....	78
8.1.1	Lignes directrices en matière de diffusion du c.v.	79
8.1.2	Utilisation des valeurs plausibles et des poids de rééchantillonnage dans le calcul de l'erreur d'échantillonnage	80
8.1.3	Estimation de la variance d'erreur dans le cadre de l'ELCA	87
8.1.4	Analyse des données de l'ELCA sur SPSS	96
8.1.5	Analyse des données de l'ELCA sur SAS.....	108
8.2	Erreurs non liées à l'échantillonnage	114
8.2.1	Base de sondage	114
8.2.2	Non-réponse	114
8.2.3	Erreur dans la réponse.....	115
8.2.4	Notation.....	115
9.0	Dispositions d'enregistrement et fréquences univariées.....	116
10.0	Principaux participants au projet	117

1.0 Introduction

Le fichier de données qui fait l'objet du présent document est une compilation des ensembles de données de l'Enquête reçus de sept pays ou régions ayant recueilli des données en 2003 : les Bermudes, le Canada, les États-Unis, l'Italie, la Norvège, la Suisse et l'État du Nuevo León, au Mexique. La Hongrie, la Nouvelle-Zélande et les Pays-Bas sont un autre groupe de pays qui ont fait la collecte de données en 2006 ou 2008.

Le présent document résume les concepts et le déroulement de l'enquête internationale. Il importe que les utilisateurs se familiarisent avec le contenu du présent document avant de publier ou de diffuser des estimations tirées du fichier de microdonnées de l'Enquête.

Si vous avez des questions concernant ce fichier de microdonnées, veuillez contacter :

Services à la clientèle du CTCSE

Services à la clientèle du CTCSE
Statistique Canada
150, promenade du Pré Tunney
Ottawa ON K1A 0T6
Tél: 1-800-307-3382
Télécopieur:: 613-951-4441
Courriel: educationstats@statcan.ca

2.0 Aperçu de l'enquête

L'Enquête sur la littératie et les compétences des adultes (ELCA) est une étude conjointe à grande échelle menée par des gouvernements, des organismes statistiques nationaux, des établissements de recherche et des organismes multilatéraux. L'élaboration et la gestion de l'étude ont été coordonnées par Statistique Canada et par le Educational Testing Service (ETS) en collaboration avec le National Center for Education Statistics (NCES) du United States Department of Education, l'Organisation de coopération et de développement économiques (OCDE), le Bureau régional d'éducation pour l'Amérique latine et les Caraïbes (OREALC) et l'Institut de la statistique de l'Organisation des Nations unies pour l'éducation, la science et la culture (UNESCO).

Le matériel d'enquête a été mis au point par des équipes internationales de spécialistes et financé par les gouvernements du Canada et des États-Unis. Un groupe hautement diversifié de pays et de spécialistes provenant du monde entier a participé à la validation des instruments. Les gouvernements participants ont assumé, à l'échelle nationale, le coût de la collecte de données et, à l'échelle internationale, une partie des coûts indirects liés à la mise en oeuvre de l'enquête.

L'ELCA s'inspire de l'Enquête internationale sur l'alphabétisation des adultes (EIAA), première étude comparative internationale des compétences des adultes, menée entre 1994 et 1998 lors de trois cycles de collecte de données. Les compétences de base mesurées par l'ELCA comprennent la compréhension de textes suivis, la compréhension de textes schématiques, la numératie et la résolution de problèmes. Les compétences supplémentaires évaluées indirectement comprennent la connaissance et l'utilisation des technologies de l'information et des communications.

3.0 Objectifs de l'enquête

L'établissement de l'ELCA avait deux objectifs fondamentaux :

- 1) Le premier objectif consistait à approfondir l'évaluation des compétences mesurées dans le cadre de l'EIAA de 1994 en ajoutant de nouveaux domaines d'évaluation avec des cadres théoriques robustes et des paramètres d'item stables d'un pays à l'autre et d'une langue à l'autre. Par le fait même, cet objectif consistait aussi à lier directement l'EIAA à l'ELCA selon les deux domaines de littératie afin de permettre une comparaison des profils de compréhension de textes suivis et de textes schématiques mesurés en 1994, puis en 2003.
- 2) Le deuxième objectif consistait à mener une analyse internationale, nationale et infra-nationale des corrélats et des antécédents possibles des compétences mesurées en rassemblant une banque de données assez importante à partir d'un nombre suffisant de répondants.

L'enquête reposait avant tout sur l'évaluation directe des compétences des répondants en littératie, en numératie et en résolution de problèmes au moyen de tâches courantes présentant divers degrés de difficulté et relevant de diverses sphères de connaissances et de compétences. Cette information était soutenue par la collecte de renseignements de base sur les répondants. De plus, le questionnaire de référence comprenait des questions sur l'auto-évaluation des compétences des répondants en littératie et en numératie, sur la formation suivie au cours de l'année précédant l'enquête et sur les obstacles perçus à l'amélioration des niveaux de compétence en littératie ou en numératie.

4.0 Concepts et définitions

4.1 Définir et mesurer les compétences : aperçu

Dans le cadre de l'ELCA, chaque échelle de compétence va de zéro jusqu'à un maximum théorique de 500 points. Une personne dont les connaissances la situent à un échelon donné de l'échelle a 80 % de probabilité de réussir une tâche qui comporte ce même niveau de difficulté. Par exemple, une personne dont le rendement a été évalué à 250 points a 80 % de probabilité de réaliser correctement une tâche dont le niveau de difficulté a été estimé à 250. Cette même personne aurait « 80 % et plus » de probabilité de réaliser correctement une tâche plus simple (environ 95 % pour une tâche dont le niveau de complexité a été établi à 200) et moins de probabilité (moins de 80 %) de réaliser une tâche plus difficile (environ 40 % pour une tâche dont le niveau de complexité a été établi à 300)¹.

En vue de faciliter l'analyse, on a regroupé ces notes continues en cinq niveaux de compétence (seulement quatre niveaux ont été définis pour l'échelle de la résolution de problèmes), le niveau 1 étant le niveau le plus faible. Les niveaux de compétence utilisés dans l'ELCA sont utiles pour résumer les résultats, mais comportent également des limites. D'abord, la proportion relativement faible de répondants qui atteignent le niveau 5 ne permet pas toujours d'établir des constatations exactes. C'est pourquoi, chaque fois que les résultats sont présentés par niveau de capacité, les niveaux 4 et 5 sont combinés. Ensuite, les niveaux indiquent des ensembles particuliers de compétences et, par conséquent, les seuils de ces niveaux ne sont pas équidistants. Par conséquent, les fourchettes de résultats pour chaque niveau ne sont pas identiques. En fait, pour les quatre domaines, le niveau 1 englobe presque la moitié de l'échelle. Les seuils du domaine de la résolution de problèmes sont établis de façon quelque peu différente, et le niveau 1 couvre exactement la moitié de l'échelle. Ce niveau comprend toutes les compétences de base requises pour atteindre des niveaux plus élevés. En d'autres mots, la capacité de lire peut se situer quelque part au niveau 1, mais la capacité de comprendre et d'utiliser ce qui a été lu affiche un accroissement de complexité allant du niveau 1 au niveau 5. Les fourchettes relativement larges des résultats du niveau 1 pour chacune des échelles entraînent donc des sous-niveaux multiples de compétences à ce niveau. Ainsi, la fourchette comprend les personnes qui peuvent à peine lire de même que celles qui lisent mal ou avec un manque d'attention².

Littératie et analphabétisme

Il est intéressant de noter que, si la probabilité d'une bonne réponse se rapproche de zéro à mesure que la tâche devient plus difficile, cette probabilité ne peut jamais vraiment atteindre ce niveau parce qu'il y a toujours des chances, même minimes, de répondre correctement, quel que soit le niveau de capacité.

C'est pourquoi les résultats de l'EIACA mesurent le rendement selon un continuum de compétences. Les échelles ne mesurent pas l'absence d'une compétence et ne distinguent pas les personnes qui possèdent une compétence donnée de celles qui ne la possèdent pas.

Le présent chapitre propose un bref aperçu des cadres de travail adoptés pour mettre au point et interpréter les échelles qui ont servi à mesurer la compréhension de textes suivis et de textes

¹ Kirsch, Jungeblut et Campbell (1992), p. 14-15.

² L'Enquête internationale sur les compétences en matière de lecture est un suivi de l'EIACA de 2003 qui fournira davantage de renseignements sur les répondants du niveau 1. Les résultats sont attendus pour 2006.

schématiques, la numératie et la résolution de problèmes lors de l'Enquête internationale sur l'alphabétisation et les compétences des adultes (ELCA). L'élaboration d'un cadre de travail semble revêtir une importance capitale dans une approche conceptuelle de la mesure. Entre autres éléments, le cadre de travail doit comprendre une définition convenue de ce qu'il faut mesurer et le choix des caractéristiques à retenir dans la conception et l'interprétation des tâches. Dans le présent chapitre, nous décrivons ces caractéristiques pour chaque mesure, mais nous donnons également des exemples d'item et précisons les caractéristiques qui, à la lumière des résultats, contribuent à la difficulté des items. Ensemble, ces renseignements fournissent un moyen pour passer d'une interprétation des résultats d'enquête centrée sur des tâches prises isolément ou sur un seul nombre, à une interprétation permettant de définir des niveaux de capacité assez généralisés pour être valables à l'égard de l'ensemble des évaluations et des groupes.

4.2 Comprendre ce qui a été mesuré lors de l'Enquête internationale sur l'alphabétisation et les compétences des adultes

4.2.1 Introduction

En 1992, l'Organisation de coopération et de développement économiques (OCDE) a conclu que la faiblesse des niveaux de littératie constituait, à l'échelle internationale, une grave menace pour les résultats économiques et la cohésion sociale (OCDE, 1992). Or, l'absence de données internationales comparables a compliqué l'analyse des problèmes de littératie observés dans les pays industrialisés, et empêché les décideurs d'en tirer des leçons. Statistique Canada et Educational Testing Service (ETS) ont collaboré pour mettre au point et mener une étude comparative internationale de la littératie.

L'Enquête internationale sur l'alphabétisation des adultes (EIAA) était la première enquête comparative menée auprès d'adultes pour étudier la répartition de la littératie parmi les pays participants. En 2000, on a publié dans un rapport final (OCDE et Statistique Canada, 2000) les résultats de trois cycles d'évaluations auxquels avaient participé quelque 23 pays ou groupes linguistiques représentant un peu plus de 50 % du PIB mondial. Si l'EIAA a jeté une base importante pour les enquêtes comparatives internationales menées auprès des adultes, on a cependant exprimé le besoin d'ajouter des éléments à mesurer. Les administrations publiques et les décideurs tenaient de plus en plus à savoir quelles autres compétences étaient nécessaires à une personne pour participer pleinement et fructueusement à une société moderne, et à la société pour relever les défis d'un monde en évolution rapide. Un projet conçu à cette fin, intitulé *Définition et sélection des compétences* (DeSeCo), a été mené sous l'égide de la Suisse. Son objectif consistait à définir, d'un point de vue théorique, un ensemble de compétences clés qui sont essentielles à une vie réussie et à la bonne marche de la société (Rychen et Salganik, 2003).

En réponse au souci d'élargir la mesure des compétences, les responsables de la mise au point de l'ELCA ont entrepris de définir des cadres de travail qui serviraient à effectuer de nouvelles mesures lors des évaluations comparatives des adultes. Ils étaient conscients du fait que la conception d'un instrument valide et fiable reposait avant tout sur des bases théoriques solides,

représentées par un cadre de travail qui reflète la pensée courante dans le domaine. Selon Messick (1994), un cadre de travail qui adopte une approche conceptuelle du plan d'une évaluation doit commencer par une définition ou un énoncé de principe général, qui précise la raison d'être de l'enquête et ce qu'elle vise à mesurer en fait de connaissances, de compétences ou d'autres attributs, puis définir divers rendements ou comportements qui reflètent ces concepts et, enfin, définir diverses caractéristiques des tâches et indiquer comment on utilisera ces caractéristiques pour concevoir les tâches qui feront ressortir ces comportements.

4.2.2 Échelonner les tâches de littératie, de numératie et de résolution de problèmes dans le cadre de l'ELCA

Les résultats de l'ELCA sont présentés sur quatre échelles – deux échelles de littératie (textes suivis et textes schématiques), une échelle de numératie et une échelle de la résolution de problèmes –, chacune allant de 0 à 500 points. On peut imaginer ces tâches disposées le long de leur échelle respective en fonction de leur difficulté pour les adultes et du niveau de capacité nécessaire pour accomplir correctement chaque tâche. Dans l'ELCA, la méthode employée pour modéliser ces continuums de difficulté et de capacité est la théorie de la réponse d'item (TRI). La TRI est un modèle mathématique servant à estimer la probabilité, pour une personne donnée, d'accomplir correctement une tâche donnée tirée d'une banque de tâches (Murray, Kirsch et Jenkins, 1998).

Sur l'échelle, la valeur attribuée à chaque item est liée au rendement d'échantillons représentatifs d'adultes des pays participants à l'égard de cet item. Elle repose sur la théorie selon laquelle une personne qui se situe à un point donné de l'échelle est capable d'accomplir, avec la même compétence, toutes les tâches qui se situent à ce point de l'échelle. Pour les besoins de l'ELCA, comme pour ceux de l'EIAA, on a déterminé qu'une personne qui se situe à un point donné de l'échelle de capacités avait 80 % des chances de répondre correctement aux items qui se situent à ce point.

Tout comme les adultes au sein de chaque pays participant à l'ELCA sont échantillonnés à partir de la population d'adultes vivant dans des ménages, chaque tâche conçue et utilisée dans l'évaluation représente un type de tâche échantillonnée à partir du domaine ou du concept défini ici. Elle est donc représentative d'un type donné de tâche de littératie, de numératie ou de résolution de problèmes qui est lié à des contextes pour adultes.

Lorsqu'on regarde la répartition des tâches le long de chacune des échelles, une question évidente se pose : qu'est-ce qui distingue les tâches qui se situent au bas de chaque échelle de celles qui se situent au milieu et au sommet de l'échelle? Les tâches qui se situent à peu près au même endroit sur chaque échelle ont-elles en commun un ensemble de caractéristiques qui leur confère un niveau de difficulté semblable? Même un examen superficiel des items révèle

Et les connaissances en matière de santé?

On peut également analyser selon ce même continuum une cinquième échelle, étroitement liée à la compréhension de textes suivis et de textes schématiques, celle des connaissances en matière de santé. Toutefois, comme elle a été créée après la collecte de données, son cadre de travail n'intègre pas encore systématiquement les détails empruntés à chacun des domaines donneurs. Comme point de départ, on entend par connaissances en matière de santé la capacité d'utiliser et de comprendre l'information imprimée liée à la santé et provenant de diverses sources. Par conséquent, les facteurs de difficulté des items de compréhension de textes suivis, de textes schématiques ou de numératie réutilisés dans cette nouvelle échelle contribuent également à sa difficulté.

que les tâches qui se situent au bas de chaque échelle diffèrent de celles qui se situent au sommet.

Afin de représenter cette progression de la complexité et de la difficulté, on a divisé chaque échelle de capacités en niveaux. Les échelles de littératie et de numératie comportent cinq niveaux de capacités allant du niveau 1 (le plus faible) au niveau 5 (le plus élevé). Ces niveaux sont définis comme suit : niveau 1 (0 à 225), niveau 2 (226 à 275), niveau 3 (276 à 325), niveau 4 (326 à 375) et niveau 5 (376 à 500). L'échelle de la résolution de problèmes comporte quatre niveaux de capacités allant du niveau 1 (le plus faible) au niveau 4 (le plus élevé). Ces quatre niveaux sont définis comme suit : niveau 1 (0 à 250), niveau 2 (251 à 300), niveau 3 (301 à 350) et niveau 4 (351 à 500).

Comme chaque niveau représente une progression des connaissances et des compétences, les personnes qui se situent à un niveau donné possèdent non seulement les connaissances et les compétences liées à ce niveau, mais aussi les compétences liées aux niveaux inférieurs. En pratique, cela signifie que les personnes dont le rendement est de 250 (le milieu du niveau 2 sur l'une des échelles de littératie ou de numératie) devraient pouvoir accomplir les tâches moyennes des niveaux 1 et 2 avec un degré de compétence élevé. Sur l'échelle de la résolution de problèmes, un point comparable serait 275. Dans l'ELCA, comme dans l'EIAA, on définit un degré de compétence élevé en fonction d'une probabilité de réponse de 80 (PR80). Cela signifie que les personnes qui se situeraient à un niveau précis de l'échelle réaliseraient les tâches à ce niveau avec une probabilité de réussite de 80 %. Cela signifie également que la probabilité qu'elles exécutent les tâches situées à un niveau de capacité plus faible sur l'échelle est supérieure à 80 %, mais pas que les personnes possédant un niveau de capacité donné ne pourront jamais effectuer des tâches dont le niveau de difficulté est plus élevé; elles pourraient réussir de temps à autre. Cela signifie que la probabilité de succès est « relativement » faible, c'est-à-dire que plus le degré de difficulté de la tâche est élevé par rapport à leur niveau de capacité, plus la probabilité de bonne réponse est faible.

On peut expliquer ce principe à l'aide d'une analogie. La relation qui existe entre la difficulté de la tâche et la compétence de la personne est comparable à l'épreuve du saut en hauteur dans une compétition d'athlétisme : l'athlète essaie de sauter par-dessus une barre placée de plus en plus haut. Chaque athlète est compétent à une hauteur particulière. À cette hauteur, la probabilité de réussite est élevée et il peut aussi sauter presque tout le temps par-dessus la barre placée à des niveaux plus bas. Cependant, lorsque la barre est placée à un niveau plus élevé que celui où l'athlète est compétent, on ne s'attend pas à ce qu'il puisse réussir de manière constante.

4.3 Mesurer la compréhension de textes suivis et de textes schématisés dans le cadre de l'ELCA

Financée par le National Center for Education Statistics (NCES) dans le cadre de son programme global d'évaluation de la littératie des adultes, la National Adult Literacy Survey (NALS) est l'étude la plus vaste et la plus complète de la littératie des adultes jamais menée aux États-Unis (Kirsch *et coll.*, 1993). Comme toutes les évaluations à grande échelle financées par le NCES, la NALS a été conçue par un comité composé d'universitaires, de praticiens et

d'administrateurs reconnus à l'échelle nationale, qui ont adopté la définition suivante de la littératie :

« La littératie est la capacité d'utiliser des imprimés et des écrits nécessaires pour fonctionner dans la société, atteindre ses objectifs, parfaire ses connaissances et accroître son potentiel. »

Cette définition résulte du travail initial du comité d'orientation de l'évaluation et peut servir de base pour créer d'autres aspects du cadre de travail à aborder. Approuvée et adoptée par les pays qui ont participé au premier cycle de l'EIAA, elle a également été retenue pour les besoins de l'ELCA . Elle comporte certaines hypothèses formulées par les membres du comité; il importe donc d'envisager tour à tour les divers éléments de cette définition.

Tout d'abord, « **La littératie est...** » : ici, on emploie le terme « littératie », de préférence à « lecture », parce qu'il est susceptible de communiquer plus précisément à un public profane ce que l'enquête mesure. Par « lecture », on entend souvent le simple fait de décoder ou lire à voix haute, alors que les enquêtes menées auprès des adultes visent à mesurer une réalité plus vaste et plus profonde. Les chercheurs qui étudient la littératie dans certains contextes ont observé que, selon les cultures et les groupes, on semblait privilégier différentes sortes de pratiques en matière de littératie (Sticht, 1975; Heath, 1980; Szwed, 1981). Heath, par exemple, observe que la pratique de la lecture peut constituer une fin en soi ou servir de mode d'interaction sociale, de moyen d'information, d'aide-mémoire, de substitut de messages oraux, de moyen d'archivage ou de confirmation personnelle. Le fait qu'on lise différents documents dans des buts différents suppose un éventail de capacités dont ne témoigne pas nécessairement le fait de signer son nom, de compter un certain nombre d'années de scolarité ou d'obtenir un résultat du niveau de la 8^e année lors d'un test scolaire de compréhension de lecture.

L'expression « ... **la capacité d'utiliser des imprimés et des écrits** » attire l'attention sur le fait que les membres du comité ne considèrent pas la littératie comme un ensemble de capacités isolées liées à la lecture et à l'écriture, mais – ce qui est plus important – comme l'application de ces capacités à des fins précises dans des contextes précis. Lorsqu'on étudie la littératie dans des contextes divers, elle devient marquée au sceau de la diversité. Premièrement, les gens pratiquent la littératie pour répondre à diverses fins ou à divers besoins (Sticht, 1978; Heath, 1980; Cook-Gumperz et Gumperz, 1981; Mikulecky, 1982). Ces utilisations varient selon les contextes (Heath, 1980; Venezky, 1983) et selon les personnes à l'intérieur du même contexte (Kirsch et Guthrie, 1984a). Cette variation de l'utilisation amène les gens à lire un vaste éventail de documents ayant des formes linguistiques qualitativement très différentes (Diehl, 1980; Jacob, 1982; Miller, 1982). Dans certains cas, on a lié ces différents types de tâches de littératie à différentes stratégies cognitives ou habitudes de lecture (Sticht, 1978, 1982; Crandall, 1981; Scribner et Cole, 1981; Kirsch et Guthrie, 1984b).

L'expression « ... **pour fonctionner dans la société, atteindre ses objectifs, parfaire ses connaissances et accroître son potentiel** » vise à englober toutes les situations dans lesquelles la littératie joue un rôle dans la vie des adultes, en privé et en public, de l'école au travail, puis à l'acquisition continue du savoir et à la participation active à la vie de la collectivité. « Atteindre ses objectifs, parfaire ses connaissances et accroître son potentiel » exprime le

point de vue selon lequel la littératie permet de combler les aspirations personnelles – celles qui sont définies, comme l’obtention d’un diplôme ou d’un emploi, et celles qui sont moins définies et moins immédiates, qui enrichissent la vie d’une personne. L’expression « dans la société » tient compte du fait que la littératie offre à chacun le moyen de contribuer à la vie de la société tout en en tirant des avantages. On s’accorde à reconnaître que les compétences en littératie sont importantes pour permettre à un pays de maintenir ou d’améliorer son niveau de vie et de s’imposer sur un marché concurrentiel de plus en plus mondial. Pourtant, elles sont tout aussi importantes pour la participation individuelle à la vie d’une société caractérisée par l’évolution technologique avec ses institutions structurées, un système juridique complexe et de vastes programmes gouvernementaux.

4.3.1 Définir les caractéristiques des tâches

Les caractéristiques des tâches représentent des variables qu’on peut utiliser de diverses façons pour mettre au point une évaluation et en interpréter les résultats. Selon Almond et Mislevy (1998), les variables peuvent jouer l’un des cinq rôles suivants : limiter la portée de l’évaluation, définir les caractéristiques à utiliser pour concevoir les tâches, contrôler l’assemblage des tâches en livrets ou en formulaires de test, caractériser le rendement des répondants ou leurs réponses aux tâches ou caractériser les aspects des compétences ou des capacités. L’EIAA s’est concentrée sur des variables pouvant servir à concevoir des tâches ainsi qu’à caractériser le rendement selon une ou plusieurs échelles de capacités.

Chaque tâche de l’évaluation témoigne d’un aspect de la littératie d’une personne (Mislevy, 2000). Si le but de l’évaluation consiste à brosser un tableau aussi fidèle que possible des compétences et des capacités d’une personne, le test ne peut pas comprendre un nombre infini de tâches et l’on ne peut pas manipuler un nombre infini de caractéristiques de ces tâches. Il faut donc faire des choix au sujet des caractéristiques à inclure dans le processus de mise au point du test. Pour les besoins de la construction des tâches de l’EIAA, on a retenu les trois caractéristiques suivantes :

Contextes ou contenu pour adultes. Comme les adultes ne lisent pas d’écrits ou d’imprimés au hasard, mais plutôt dans un contexte donné ou pour répondre à un besoin particulier, on choisit, en vue de l’évaluation de la littératie, des documents qui représentent divers contextes et divers contenus. On s’assure ainsi qu’aucun groupe d’adultes n’est avantagé ou désavantagé en raison du contexte ou du contenu retenu pour l’évaluation. On a retenu les six catégories suivantes de contexte ou de contenu pour adultes :

- ❖ **Vie familiale** : documents portant sur les relations interpersonnelles, les finances personnelles, le logement et l’assurance.
- ❖ **Santé et sécurité** : documents portant sur les drogues et l’alcool, la prévention et le traitement des maladies, la sécurité et la prévention des accidents, les premiers soins, les urgences et les mesures à prendre pour rester en santé.
- ❖ **Vie sociale et collective** : documents portant sur les ressources communautaires et les moyens d’information.
- ❖ **Économie de la consommation** : documents portant sur le crédit et les opérations bancaires, l’épargne, la publicité, les achats et les biens personnels.

- ❖ **Travail** : documents portant sur diverses professions en général (mais pas sur des cas particuliers), la recherche d'un emploi, les finances et la vie professionnelle.
- ❖ **Loisirs et détente** : documents portant sur les voyages, les activités récréatives et les restaurants.

Documents ou textes. S'il ne fait aucun doute qu'une évaluation de la littératie doit porter sur des documents divers, la diversité et les caractéristiques spécifiques des textes qui entrent dans la conception des tâches sont essentielles à la conception et à l'interprétation des résultats qui sont produits. Pour les besoins de l'ELCA, on a établi une distinction essentielle entre les textes continus et les textes non continus. Ordinairement, les textes continus sont formés de phrases organisées en paragraphes. Dans ces textes, l'organisation tient à la disposition des paragraphes, aux alinéas et à la ventilation du texte en une hiérarchie signalée par des rubriques qui permettent au lecteur de reconnaître l'organisation du texte. Les textes sont classés le plus souvent selon le but de l'auteur ou le type de texte. Pour l'EIAA, ces catégories sont les suivantes : exposition, description, argumentation et instructions.

Les textes non continus sont organisés différemment des textes continus, de sorte qu'ils permettent au lecteur d'employer des stratégies différentes pour y accéder et en extraire de l'information. En surface, ces textes semblent présenter de nombreux principes d'organisation ou formats différents – tableaux, horaires, diagrammes, graphiques, cartes géographiques, formulaires, etc. Toutefois, on dit que le principe d'organisation de ces types de texte, que Mosenthal et Kirsch (1998) appellent textes schématiques, présente l'une des quatre structures de base suivantes : liste simple, liste combinée, liste croisée ou liste imbriquée. Ensemble, ces quatre types de document constituent ce que les auteurs appellent des documents matriciels, soit des textes non continus présentant des rangées et des colonnes bien définies. Ils sont aussi étroitement liés à d'autres textes non continus que ces auteurs appellent documents graphiques, documents locatifs et documents à remplir.

Deux des trois échelles de littératie utilisées dans l'EIAA reposaient sur la distinction entre textes continus et textes non continus. Les tâches situées le long de l'échelle des textes suivis étaient axées sur des textes continus, alors que celles situées le long de l'échelle des textes schématiques étaient axées sur des textes non continus. L'échelle des textes au contenu quantitatif comprenait à la fois des textes continus et des textes non continus. La caractéristique distinctive de cette échelle tenait au fait que les répondants devaient cerner et exécuter une ou plusieurs opérations arithmétiques fondées sur l'information contenue dans les textes. Dans l'ELCA, on a remplacé cette échelle par l'échelle de la numératie, dont nous traiterons en détail plus loin dans la présente annexe.

Processus ou stratégies. Cette caractéristique des tâches a trait à la façon dont le répondant traite le texte pour répondre correctement à une question ou à une directive. Elle comprend les processus utilisés pour rapprocher l'information énoncée dans la question (l'information donnée) de l'information nécessaire contenue dans le texte (la nouvelle information), ainsi que les processus nécessaires pour repérer ou formuler la bonne réponse à partir de l'information disponible. Nous envisageons ici trois variables retenues pour examiner les tâches de lecture ou de littératie utilisées dans les enquêtes nationales et internationales : le type d'appariement, le type d'information demandée et la plausibilité des éléments de distraction.

4.3.2 Type d'appariement

On distingue quatre types de stratégie d'appariement : le repérage, le regroupement des caractéristiques, l'intégration et la formulation. Pour les tâches de **repérage**, le répondant doit appairer un ou plusieurs éléments d'information énoncés dans la question à des renseignements identiques ou synonymiques présents dans le texte. Les tâches de **regroupement des caractéristiques** invitent également le répondant à appairer un ou plusieurs éléments d'information; elles diffèrent cependant des tâches de repérage du fait que le répondant doit appairer une série de caractéristiques pour satisfaire aux conditions énoncées dans la question.

Les tâches **d'intégration** demandent au répondant de rapprocher deux ou plusieurs éléments d'information tirés du texte selon un type de relation donnée. Par exemple, cette relation peut amener le répondant à repérer des similitudes (établir une comparaison), des différences (mettre en opposition), un degré (trouver plus petit ou plus grand) ou des relations de cause à effet. Cette information peut se trouver dans un seul paragraphe ou figurer dans différents paragraphes ou différentes sections du texte. Lorsqu'il intègre l'information, le répondant s'inspire des catégories d'information données dans une question pour repérer l'information correspondante dans le texte. Il peut alors rapprocher l'information énoncée dans le texte à ces différentes catégories en fonction du terme de relation précisé dans la question. Dans certains cas, toutefois, le répondant doit **formuler** ces catégories ou relations avant d'intégrer l'information énoncée dans le texte.

Outre qu'il demande au répondant d'appliquer l'une de ces quatre stratégies, le type d'appariement entre une question et le texte est influencé par plusieurs autres conditions de traitement qui contribuent à la difficulté globale d'une tâche. La première de ces conditions est le nombre d'expressions à utiliser dans la recherche. La difficulté d'une tâche augmente avec la quantité d'information énoncée dans la question et que le répondant doit chercher dans le texte. Par exemple, les questions qui comportent une seule proposition indépendante sont habituellement plus simples, en moyenne, que celles qui contiennent plusieurs propositions, indépendantes ou non. La difficulté augmente aussi avec le nombre de réponses qu'on demande au répondant de donner. Les questions qui demandent une seule réponse sont plus simples que celles qui en demandent trois ou plus. En outre, les questions qui précisent le nombre de réponses demandées sont habituellement plus simples que celles qui ne le font pas. Par exemple, une question qui énonce « Énumérez les trois raisons... » est plus simple qu'une question qui énonce « Énumérez les raisons... ». Les tâches sont également influencées par la mesure dans laquelle le répondant doit faire des déductions pour appairer l'information donnée dans une question à l'information correspondante dans le texte et pour trouver l'information demandée.

4.3.3 Type d'information demandée

Il s'agit de la sorte d'information que le lecteur doit repérer pour répondre correctement à une question du test. Plus l'information demandée est concrète, plus la tâche est jugée facile. Dans les travaux de recherche antérieurs portant sur les évaluations à grande échelle de la littératie des adultes et des enfants (Kirsch et Mosenthal, 1994; Kirsch, Jungeblut et Mosenthal, 1998), la variable « type d'information » était notée sur une échelle à cinq points. La cote 1

représentait l'information la plus concrète, donc la plus simple à traiter, et la cote 5, l'information la plus abstraite, donc la plus difficile à traiter.

Par exemple, les questions où le répondant devait trouver une personne, un animal ou une chose (soit des noms imaginables) demandaient une information très concrète; on leur attribuait donc une valeur de 1. Celles où le répondant devait trouver des objectifs, des conditions ou des buts demandaient des types d'information plus abstraits. Comme elles étaient jugées plus difficiles, on leur attribuait une valeur de 3. Les questions qui demandaient au répondant de trouver un « équivalent » étaient jugées les plus abstraites; on leur attribuait donc une valeur de 5. Dans ce dernier cas, l'équivalent était habituellement un terme ou une expression que le répondant ne connaissait pas et pour lequel il devait déduire une définition ou une interprétation à partir du texte.

4.3.4 Plausibilité des éléments de distraction

Il s'agit de la mesure dans laquelle l'information contenue dans le texte possède une ou plusieurs caractéristiques en commun avec l'information demandée dans la question, mais sans satisfaire entièrement aux exigences de la question. Les tâches sont jugées le plus simples lorsque le texte ne comporte aucun élément de distraction. Elles ont tendance à devenir plus difficiles à mesure que le nombre d'éléments de distraction augmente, que les éléments de distraction possèdent plus de caractéristiques en commun avec la bonne réponse et qu'ils figurent plus près de la bonne réponse. Par exemple, les tâches sont habituellement jugées plus difficiles lorsqu'un ou plusieurs éléments de distraction répondent à une partie, mais non à la totalité, des conditions énoncées dans la question et qu'ils figurent dans un paragraphe ou une section du texte autre que celui ou celle qui contient la bonne réponse. Les tâches sont jugées le plus difficiles lorsque deux ou plusieurs éléments de distraction possèdent la plupart des caractéristiques de la bonne réponse et qu'ils figurent dans le même paragraphe ou bloc d'information que la bonne réponse.

4.3.5 Caractériser les tâches de compréhension de textes suivis

L'échelle de compréhension de textes suivis comporte 55 tâches ordonnées le long de l'échelle de 500 points; de ce nombre, 19 tâches sont tirées de l'EIAA et 36 nouvelles tâches ont été conçues et mises au point pour les besoins de l'ELCA. L'indice de difficulté de ces tâches varie de 169 à 439. L'une des tâches les plus simples (indice de difficulté : 188; niveau 1) consiste, pour le répondant, à lire une étiquette de médicament afin de déterminer le nombre maximal de jours de prise du médicament. En fonction des variables retenues, le type d'appariement est jugé simple parce que le lecteur devait repérer un seul élément d'information qui est énoncé textuellement sur l'étiquette du médicament. L'étiquette ne fait qu'une seule référence au nombre de jours, et ce renseignement figure sous la rubrique « Posologie ». Le type d'information est jugé simple parce qu'on demandait un certain nombre de jours et la plausibilité des éléments de distraction est jugée simple parce qu'il n'y a pas d'autre mention d'un nombre de jours sur l'étiquette du médicament.

ASPIRIN MEDCO

500

INDICATIONS THÉRAPEUTIQUES : Maux de tête, douleurs musculaires et rhumatismales, maux de dents et d'oreilles. SOULAGE LES SYMPTÔMES COMMUNS DU RHUME.

POSOLOGIE : ORALE. Prendre 1 ou 2 comprimés toutes les 6 heures, de préférence en mangeant, pendant au plus 7 jours. Conserver en un endroit frais et sec.

MISE EN GARDE : Ne pas prendre en cas de gastrite ou d'ulcère gastro-duodénal. Éviter de prendre en même temps qu'un anticoagulant, ou en cas de maladie du foie ou d'asthme bronchique grave. Si ce médicament est pris à grosse dose pendant une période prolongée, il peut affecter les reins. Avant d'administrer à un enfant atteint de varicelle ou de grippe, consulter un médecin au sujet du syndrome de Reyes, maladie rare mais grave. Les femmes enceintes et celles qui allaitent doivent consulter leur médecin avant d'utiliser ce produit, surtout durant le troisième trimestre de la grossesse. En cas de symptômes persistants ou de surdose accidentelle, consulter un médecin. Garder hors de la portée des enfants.

INGRÉDIENTS : Chaque comprimé contient 500 mg d'acide acétylsalicylique.

Excipient c.b.p. 1 comprimé.

N° d'enreg. 88246

Fabrique au Canada par STERLING PRODUCTS, INC.
1600, boul. Industriel, Montréal (Québec) H3J 3P1



Reproduit avec autorisation

Une deuxième tâche de compréhension de textes suivis demande au lecteur de lire un article sur les impatientes. Cette tâche se situe au milieu du niveau 2 et comporte un indice de difficulté de 254. Elle demande au lecteur ce qu'il peut déduire de la surface lisse des feuilles et des tiges de cette plante. Encore une fois, la tâche amenait le lecteur à repérer une information contenue dans le texte; elle est donc jugée simple en ce qui concerne le type d'information. La dernière phrase du deuxième paragraphe, sous la rubrique *Apparence*, énonce : « La surface lisse des feuilles et la nature des tiges montrent que cette plante a grand besoin d'eau. » Le type d'information est jugé moyen parce que la tâche demandait au lecteur de trouver un état. La plausibilité des éléments de distraction est également jugée moyenne parce que le même paragraphe contenait une phrase qui pouvait distraire un certain nombre de lecteurs. Cette phrase énonce : « Les tiges (...) sont ramifiées et très juteuses, ce qui, en raison de l'origine tropicale de cette plante, la rend très fragile au froid. »

CHOIX D'UN CADRE APPROPRIÉ

LE CYCLISTE DOIT POUVOIR ENFOURCHER SA BICYCLETTE ET, EN POSITION DEBOUT, AVOIR AU MOINS 2 cm DE JEU

AU MOINS 2 cm

AU MOINS 2 cm

NOTA : Pour les femmes, déterminer la mesure à partir d'un modèle pour hommes.

GRANDEUR APPROPRIÉE DE LA BICYCLETTE	
GRANDEUR DU CADRE	LONGUEUR DE JAMBE DU CYCLISTE
430mm	660mm-760mm
460mm	690mm-790mm
480mm	710mm-790mm
530mm	760mm-840mm
560mm	790mm-860mm
580mm	810mm-890mm
635mm	860mm-940mm

AU-DESSUS DE LA TIGE HORIZONTALE.

RESPONSABILITÉS DU PROPRIÉTAIRE

1. Choix et achat de la bicyclette : Voir si la bicyclette convient au futur propriétaire. Les bicyclettes sont de grosseur variable. Pour assurer une sécurité et un confort optimaux, la selle et les guidons doivent être ajustés. Comme les bicyclettes sont dotées d'une grande variété d'équipement et d'accessoires ..., il faut veiller à ce que le cycliste sache comment s'en servir.
2. Assemblage : Suivre attentivement les directives d'assemblage. Veiller à ce que tous les écrous, boulons et vis soient bien serrés.
3. Réglage de la bicyclette : Pour que le cycliste puisse rouler en toute sécurité et de façon confortable, la bicyclette doit être bien ajustée. La hauteur de la selle doit être réglée de manière que, lorsqu'il a le pied posé à plat sur la pédale au plus bas de sa course, le cycliste ait le genou légèrement plié.

Nota : Le tableau de gauche permet de déterminer la grandeur de cadre appropriée.

Le fabricant n'est pas responsable des défaillances, blessures ou dommages causés par un assemblage incomplet ou par un mauvais entretien après expédition.

Les tâches qui se situent à des niveaux plus élevés sur l'échelle présentent au lecteur des exigences plus variées en fonction du type d'appariement demandé ainsi que du nombre et de la nature des éléments de distraction présents dans le texte. L'une de ces tâches (indice de difficulté : 281; début du niveau 3) demande au lecteur de consulter une page d'un manuel d'entretien de bicyclette et de trouver comment s'assurer que le siège est dans la bonne position. Le type d'information est jugé moyen parce que le lecteur devait trouver et énoncer par écrit deux conditions à respecter. De plus, on ne précisait pas combien de caractéristiques il devait trouver parmi celles qui étaient énoncées. Le type d'information est également jugé moyen parce qu'il s'agissait de trouver une condition; enfin, on a attribué à la plausibilité des éléments de distraction une cote indiquant qu'elle était relativement simple.

Une tâche un peu plus difficile (318), qui se situe près du sommet du niveau 3, porte sur un article concernant les couches en coton. On demande au lecteur d'énumérer trois raisons pour lesquelles l'auteur préfère les couches en coton aux couches jetables. Cette tâche est relativement difficile en raison de plusieurs variables. Premièrement, le type d'appariement est

jugé difficile parce que le lecteur devait donner plusieurs réponses en faisant des déductions à partir du texte. L'auteur ne dit nulle part dans le texte : « Je préfère les couches en coton parce que... ». Ces déductions sont un peu plus difficiles parce que le type d'information demandée est une « raison » plutôt qu'une réponse concrète. Cette variable est également jugée difficile à cause de son côté abstrait. Enfin, la plausibilité des éléments de distraction est jugée moyenne parce que le texte contient de l'information pouvant distraire le lecteur.

Une autre tâche, qui se situe au niveau 4 sur l'échelle des textes suivis (338), demande au lecteur d'utiliser l'information contenue dans un dépliant sur les entrevues d'emploi et de décrire dans ses mots une différence entre l'entrevue par jury et l'entrevue en groupe. Ici, la difficulté ne tient pas au repérage de l'information dans le texte. Au lieu de simplement repérer un fait concernant chaque type d'entrevue, le lecteur doit intégrer ce qu'il a lu pour trouver une caractéristique qui distingue les deux types d'entrevue. L'expérience retenue d'autres évaluations de ce genre révèle que les tâches demandant au lecteur de mettre en opposition des éléments d'information sont plus difficiles, en moyenne, que celles pour lesquelles on lui demande de trouver des similitudes. Le type d'appariement est donc jugé complexe et difficile. Le type d'information est également jugé difficile parce qu'on demandait au lecteur de trouver une différence. Les différences sont habituellement plus abstraites puisqu'il s'agit de repérer des caractéristiques distinctives ou opposées liées, dans ce cas, à un processus d'entrevue. La plausibilité des éléments de distraction est jugée simple parce que le texte ne contenait aucun élément de distraction. On n'estimait donc pas que cette variable contribuait à la difficulté globale de la tâche.

L'entrevue d'emploi

Avant l'entrevue

Essayez de bien vous renseigner sur l'entreprise. Quels produits fabrique-t-elle ou quels services offre-t-elle? Quelles méthodes ou procédés utilise-t-elle? Ces renseignements peuvent être trouvés dans des annuaires professionnels, des répertoires de chambre de commerce ou d'industrie, ou à votre bureau d'emploi local.

Renseignez-vous aussi sur le poste. S'agit-il de remplacer une personne ou de combler un poste nouvellement créé? Dans quels services ou ateliers travailleriez-vous? Vous pouvez obtenir de la plupart des bureaux d'emploi locaux des conventions collectives qui décrivent divers postes et fonctions uniformisés. Vous pouvez aussi communiquer avec l'organisation syndicale appropriée.

À l'entrevue

Posez des questions au sujet du poste et de l'entreprise. Répondez clairement et avec précision à toutes les questions qui vous sont posées. Apportez un bloc-notes ainsi que vos documents de travail et de formation.

Les genres d'entrevue les plus courants

Individuelle : Se passe d'explication.

Par jury : Un certain nombre de personnes vous posent des questions et comparent ensuite leurs opinions sur votre candidature.

En groupe : Après avoir écouté un exposé sur le poste et les fonctions en compagnie d'autres candidats, vous participez à une discussion en groupe.

Après l'entrevue

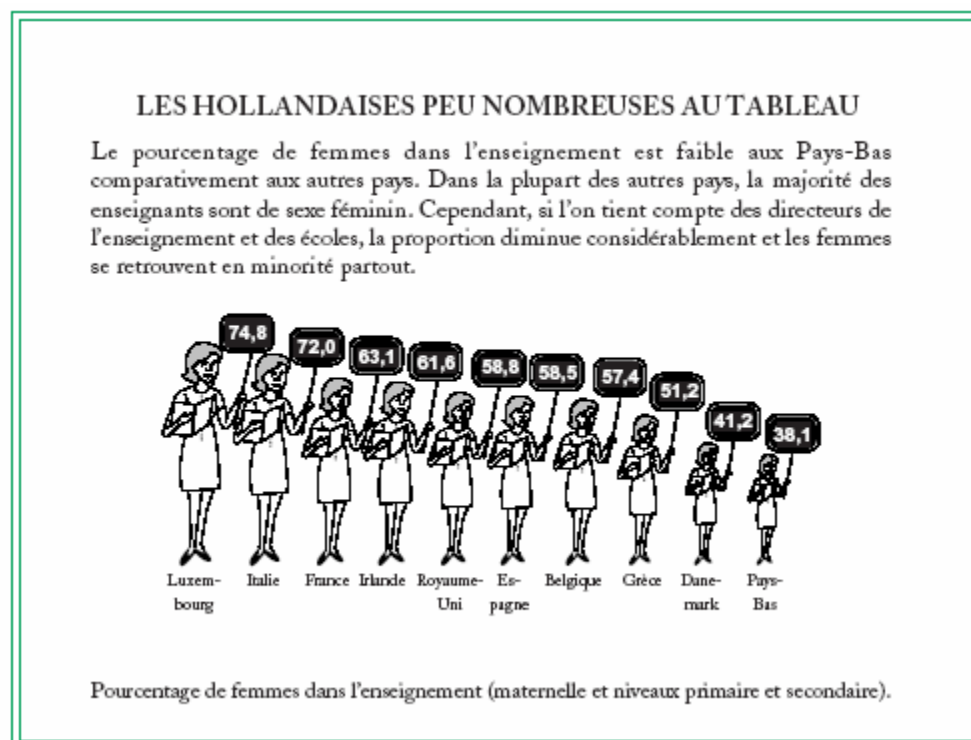
Notez les principaux points qui ont été discutés. Comparez les questions qui vous ont causé des difficultés et celles qui vous ont permis de faire valoir vos points forts. Une telle revue vous aidera à vous préparer à de futures entrevues. Si vous le désirez, vous pouvez en discuter avec l'agent de placement ou l'orienteur professionnel de votre bureau d'emploi local.

La tâche la plus difficile sur l'échelle des textes suivis (377) se situe au bas du niveau 5 et demande au lecteur de lire une annonce publiée par un service du personnel et de citer deux façons dont le CIEM (un groupe d'aide aux salariés d'une entreprise) peut aider les personnes qui perdent leur emploi par suite d'une restructuration de leur service. Le type d'appariement est jugé difficile parce que la question contenait plusieurs éléments d'information que le lecteur devait garder à l'esprit en lisant le texte. De plus, le lecteur devait donner plusieurs réponses et faire des déductions de faible niveau à partir du texte. Le type d'information est jugé moyen parce que le lecteur cherchait un but ou une fonction; la plausibilité des éléments de distraction est jugée relativement difficile. Cette tâche est un peu plus difficile du fait que l'annonce est

centrée sur des renseignements qui sont différents de ceux qui sont demandés dans la question. Ainsi, bien que les renseignements corrects se situent dans une seule phrase, l'information est cachée sous une série de rubriques décrivant les activités du CIEM pour les employés à la recherche d'un autre emploi. Cette liste de rubriques représente un excellent élément de distraction pour le lecteur qui ne cherche pas ou ne repère pas le renseignement conditionnel énoncé dans la question, soit celui qui concerne les personnes qui perdent leur emploi en raison d'une réorganisation de leur service.

4.3.6 Caractériser les tâches de compréhension de textes schématiques

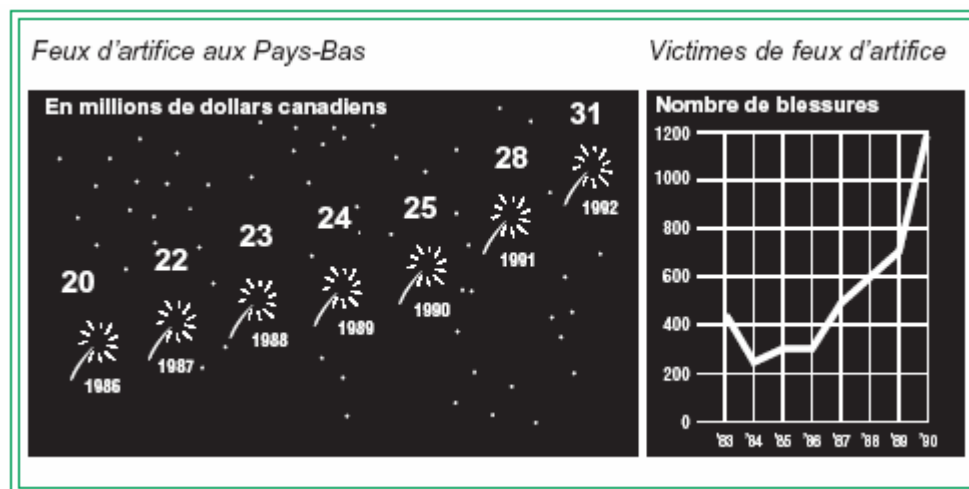
L'échelle de compréhension de textes schématiques comporte 54 tâches ordonnées le long de l'échelle de 500 points. Ces 54 tâches comprennent 19 items tirés de l'EIAA et 35 nouvelles tâches mises au point pour les besoins de l'ELCA. L'indice de difficulté de ces tâches varie de 157 à 444. L'une des tâches de niveau 1 (indice de difficulté : 188) demande au lecteur de déterminer, à partir d'un diagramme, le pourcentage d'enseignants en Grèce qui sont des femmes. Le diagramme présente le pourcentage d'enseignants de divers pays qui sont des femmes. En fonction des variables retenues, le type d'appariement est jugé simple parce que le lecteur devait repérer un seul élément d'information qui était énoncé littéralement dans le diagramme; le type d'information est jugé relativement simple parce qu'il s'agissait d'un montant; enfin, la plausibilité des éléments de distraction est aussi jugée relativement simple parce que des éléments de distraction entourent l'information demandée.



Une deuxième tâche de compréhension de textes schématiques relative à ce même diagramme demande au lecteur de trouver le pays autre que les Pays-Bas dans lequel les femmes sont en

minorité parmi les enseignants (indice de difficulté : 234; milieu du niveau 2). Cette tâche était un peu plus difficile que la première car, au lieu de chercher un pays et de repérer un pourcentage, le lecteur devait savoir qu'une minorité signifie moins de 50 %. Puis, il devait passer en revue les pourcentages pour trouver les pays dans lesquels le pourcentage de femmes parmi les enseignants était inférieur à 50 %. De plus, il devait se rappeler la condition « autre que les Pays-Bas »; autrement, il aurait pu choisir ce pays au lieu de la bonne réponse. Par conséquent, le type d'appariement est jugé moyennement difficile; le type d'information, simple (parce que l'information demandée est un pays ou un lieu); et la plausibilité des éléments de distraction, relativement simple (parce que des éléments de distraction entourent l'information demandée).

Une tâche un peu plus difficile (indice de difficulté : 295; milieu du niveau 3) demande au lecteur de regarder des diagrammes concernant les feux d'artifice aux Pays-Bas et de rédiger une brève description du lien qui existe entre les ventes et les blessures à partir de l'information fournie. Ici, le lecteur doit examiner et comparer l'information contenue dans les deux diagrammes et l'intégrer en faisant une déduction concernant la relation entre les deux ensembles d'information. La tâche est jugée relativement difficile en fonction du type d'appariement. Le type d'information est aussi jugé relativement difficile parce que l'information demandée est une tendance ou une similitude dans les données. La plausibilité des éléments de distraction est jugée moyennement difficile, principalement parce que l'information donnée et l'information demandée sont présentes dans la tâche. Par exemple, l'un des facteurs pouvant avoir contribué à la difficulté de cette tâche est le fait que le graphique des ventes va de 1986 à 1992, alors que celui des blessures va de 1983 à 1990. Le lecteur devait comparer l'information des deux diagrammes pour la période comparable.



Un autre ensemble de tâches de difficulté variable sur l'échelle de compréhension de textes schématiques concerne un document plutôt compliqué tiré d'une revue de consommateurs dans laquelle on évalue les radios-réveils. La plus simple des trois tâches (indice de difficulté : 287; niveau 3) demande au lecteur de trouver deux caractéristiques qui n'appartiennent à aucun radio-réveil de base. Le lecteur doit parcourir le document pour trouver la liste des radios-réveils de base, puis déterminer qu'un tiret représente l'absence d'une caractéristique. Il doit alors repérer les deux caractéristiques indiquées par l'ensemble de tirets. Le type d'appariement est donc jugé relativement difficile parce qu'il s'agit d'une tâche de regroupement des caractéristiques nécessitant plusieurs réponses avec une condition ou une déduction de

faible niveau. Le type d'information est jugé relativement simple parce qu'il s'agit de trouver des caractéristiques du radio-réveil, et la plausibilité des éléments de distraction est aussi relativement simple parce que certaines caractéristiques ne sont pas liées aux autres radios-réveils.

Une tâche un peu plus difficile liée à ce document se situe au bas du niveau 4 et comporte un indice de difficulté de 327. Elle demande au lecteur : « Quel radio-réveil possédant toutes les caractéristiques est le mieux coté à l'égard du rendement? » Ici, le lecteur doit apparier trois caractéristiques (« possédant toutes les caractéristiques », « le mieux coté » et « rendement »), dont l'une l'amène à traiter un renseignement conditionnel. Il est possible, par exemple, qu'un lecteur ait trouvé les radios-réveils possédant toutes les caractéristiques et la colonne « Rendement », mais qu'il ait choisi le premier radio-réveil de la liste en supposant qu'il était le mieux coté. Dans ce cas, il n'a pas compris le renseignement conditionnel, qui est une légende énonçant la signification des symboles. Un autre lecteur peut être allé à la colonne « Cote globale » et, ayant trouvé le nombre le plus élevé, avoir choisi le radio-réveil correspondant. Pour ces raisons, la plausibilité des éléments de distraction est jugée moyennement difficile. Le type d'information est jugé simple parce que l'information demandée est une chose concrète.

La tâche la plus difficile relative à ce document (indice de difficulté : 408; niveau 5) demande au lecteur de trouver le prix moyen annoncé pour le radio-réveil de base ayant obtenu la meilleure cote globale. Cette tâche est plus difficile parce que le lecteur devait apparier quatre caractéristiques au lieu de trois; il devait aussi traiter un renseignement conditionnel; enfin, il y avait un élément de distraction hautement plausible dans le même bloc que la bonne réponse. En raison de ces facteurs, le type d'appariement est jugé relativement difficile et le type d'information, relativement simple, tandis que la plausibilité des éléments de distraction présente le plus haut niveau de difficulté.

ÉVALUATIONS

Radios-réveils

Présentée par catégories et, à l'intérieur de chacune, par ordre de cote globale. Les cotes de 4 points ou moins au niveau de la cote ont été sous-évaluées.

1 Marque et modèle. Si vous ne pouvez trouver un modèle donné, priez d'appeler le fabricant. Les cotes de Marque et Modèle du fabricant figurent à la page 738.

2 Prix. Prix de vente conseillé par le fabricant ou prix de détail approximatif, s'il n'y a pas de moyen annoncé.

3 Dimensions. Au centimètre près.

4 Cote globale. Réfète les résultats de tous les tests et évaluations. Une radio « parfaite » aurait une cote de 100 points.

5 Commodité. L'évaluation de cet aspect complexe tient compte d'éléments tels la lisibilité de l'affichage, la facilité de régler le son et de régler le réveil, ainsi que la présence ou l'absence de caractéristiques utiles.

6 Rendement. Appréciation globale reflétant les résultats des tests de sensibilité et de sélectivité; de la fidélité de reproduction; du rendu de l'écoute; la capacité de recevoir la station qui a la plus grande puissance de diffusion en onde moyenne et la même fréquence; du rejet d'images; la capacité d'écarter les signaux parasites de la bande de signal; et de la résistance aux interférences provenant de signaux réfléchis par des avions ou par d'autres surfaces réfléchissantes.

7 Sensibilité. Qualité de la réception d'une même station, par les serrades, sans trop d'interférence.



8 Sélectivité. Qualité de la réception, par toutes les ondes, d'une station faible, voisine d'une forte sur la bande.

9 Qualité du son. Fondée principalement sur l'analyse informative des sons produits par le haut-parleur et sur des tests d'écoute de musique enregistrée sur disques compacts. Aucun modèle n'a donné de son haut-fidélité.

10 Réglage par avance ou retard. Cette caractéristique facilite le réglage de l'horloge et du réveil, car si on se réveille trop tôt, il suffit de reculer.

11 Double sonnerie. Permet de sélectionner deux heures de réveil distinctes.

1	2	3	4	5	6	7	8	9	10	11	Avantages	Inconvénients	Observations
Radios-réveils possédant toutes les caractéristiques													
RCA RP-36	50\$/40\$	8x25x18	86	●	●	●	●	●	✓	✓	12	A,B,D,H,J,L,O,T,U	A
Sony ICF-C303	50/45	5x20x15	84	●	●	●	●	●	✓	✓	12	C,E,F,I,N,T	C
Panasonic RC-X220	50/45	10x28x13	82	●	●	●	●	●	✓	✓	12	A,G,K,M,O,S,T,U	b,c A
Realistic 272	50/30	5x28x15	79	●	●	●	●	●	✓	✓	3	A,G,H,K,O,T	D
Magnevox AJ3900	65/—	15x35x13	78	●	●	●	●	●	—	—	3	D,G,K,M,O,R,T	b,g B
Emerson AK2745	39/20	8x28x15	70	●	●	●	●	●	✓	✓	3	G,O	g K
Soundesign 3753	20/20	8x23x13	62	●	●	●	●	●	✓	✓	3	J,Q	d,h J
Radios-réveils de base													
Realistic 263	28/18	10x20x10	74	●	●	●	●	●	—	—	3	A,D,H,O,P,U	h —
Soundesign 3622	12/10	5x20x13	68	●	●	●	●	●	—	—	3	U	d L
Panasonic RC-6064	18/15	5x20x13	67	●	●	●	●	●	—	—	12	—	b,c —
General Electric 7-4612	13/10	5x20x13	66	●	●	●	●	●	—	—	12	A,D	a,g —
Lloyds CR001	20/15	5x18x13	64	●	●	●	●	●	—	—	3	U	—
Sony ICF-C240	15/13	5x18x15	63	●	●	●	●	●	—	—	12	—	f,g —
Emerson AK2720	19/10	5x20x13	61	●	●	●	●	●	—	—	3	O,T	e K
Gran Prix D507	15/10	5x18x10	54	●	●	●	●	●	—	—	3	—	d —
Radios-réveils avec lecteur de cassettes													
General Electric 7-4965	60/50	10x30x15	85	●	●	●	●	●	✓	✓	12	A,D,G,H,K,O,S,T	— B,E
Panasonic RC-X250	11	10x33x13	78	●	●	●	●	●	✓	✓	12	A,G,K,O,R,U	b,c A,H
Sony ICF-CS650	75/65	15x28x15	74	●	●	●	●	●	✓	✓	12	G,R,T,U	c,f,i A,F,H
Soundesign 3844MGY	40/30	13x30x13	62	●	●	●	●	●	—	—	3	G,K,J,S,U	F,G,I,M

11. Discontinué et remplacé par RC-X260, prix courant de 79 \$ et prix de vente moyen annoncé de 69 \$.

Caractéristiques communes

A. sans. • Permet de sélectionner environ 5 min. • Conserve les heures mémorisées pendant les pauses de courant de source d'alarme.

Son/lecture continue, nouveau. • Une alternation de sonnerie pour la mémoire de la pendule et de réveil. • Affichage en range de chiffres de 1 cm de haut. • Période maximale de 60 min. de sonnerie pour vous réveiller, avec arrêt automatique. • Rappel de sonnerie.

Légende des avantages

A. Fonctionnement du réveil même en cas de panne de courant.

B. Affiche l'heure et deux heures distinctes de réveil.

C. Capable possible d'une station d'écoute à distance des deux heures de réveil.

D. Sonnerie à volume réglable.

E. Mémoire de réveil par de pile.

F. Système de réveil avec pré-réglage de stations.

G. Permet d'écouter en direct la bande FM.

H. Indicateur de puissance des piles.

I. Cordon de réveil en direct.

J. Réveille lumineuse.

K. Pour d'écouter.

L. Clavier de réveil.

M. Entrée audio pour lecteur de cassettes ou de CD.

N. Affichage de la date et de l'heure.

O. Affichage à deux intensités lumineuses.

P. Affichage à chiffres plus gros que la plupart des radios.

Q. Taille d'écoute maximale la plus grande de la pile.

R. Fonctionnement des heures.

S. Fonctionnement des signaux.

T. Fonctionnement plus facile que la plupart des radios.

U. Réglage d'écoute maximale la plus grande de la pile.

Légende des inconvénients

a. Pour le chargeur d'alarme par accident.

b. Les boutons de réglage de l'heure ou le réveil sont placés à des endroits peu commodes; en conséquence, il faut les régler à l'aveugle.

c. Affichage plus difficile que la plupart d'autres avec piles à recharge interne.

d. Le volume doit être réglé plus bas pour la sonnerie et l'écoute.

e. Sans sonnerie, la radio et le son réveil.

f. Sans sonnerie de source.

g. Pas de rappel de sonnerie.

h. Pas de réveil rapide pour le réglage de l'heure.

i. Pas d'alarme lumineuse de réveil rapide pour le réglage de l'heure.

Légende des observations

A. Affichage des chiffres en vert.

B. Affichage des chiffres en bleu.

C. Affichage à cristaux liquides.

D. Bonne d'alarme externe.

E. Égalisateur graphique à trois positions.

F. Le lecteur de cassettes ne permet pas d'écouter.

G. Le lecteur de cassettes ne permet pas de sélectionner.

H. Nécessite d'être réglé par le lecteur de cassettes.

I. Le lecteur de cassettes est plus bruyant que la plupart.

J. Frais de maintenance de 3 \$ pour les réparations garanties.

K. Frais de maintenance de 3,50 \$ pour les réparations garanties.

L. Frais de maintenance de 6 \$ pour les réparations garanties.

M. Frais de maintenance de 10 \$ pour les réparations garanties.

4.4 Mesurer la numératie dans le cadre de l'ELCA

La conception de la numératie adoptée pour les besoins de l'ELCA repose sur des travaux de recherche récents, menés dans plusieurs pays, sur les exigences fonctionnelles de différents contextes de vie, sur la nature des connaissances et des compétences mathématiques et statistiques des adultes, ainsi que sur l'application ou l'utilisation de ces compétences dans différentes circonstances. Vu l'intention générale de l'ELCA de produire des renseignements sur un ensemble diversifié de compétences utiles à la vie, le présent cadre de travail définit la numératie comme suit :

Connaissances et compétences nécessaires pour gérer efficacement les exigences mathématiques de diverses situations.

Cette définition suppose que le concept de la numératie est plus large que celui de la compréhension de textes au contenu quantitatif, défini par l'EIAA. En outre, la numératie des adultes doit être considérée comme différente d'une « connaissance scolaire des mathématiques ». S'il n'existe pas de définition universellement reconnue de la « numératie » (Baker et Street, 1994), un examen de certaines perspectives concernant le sens de la numératie des adultes montre qu'elles renferment plusieurs points communs. Voici deux exemples, tous deux tirés de travaux menés en Australie :

La numératie consiste à utiliser les mathématiques pour fonctionner efficacement au sein d'un groupe et d'une collectivité, ainsi qu'à utiliser ces compétences pour favoriser l'évolution d'une personne et celle de sa collectivité (Beazley, 1984). [Traduction libre]

La numératie consiste à interpréter, appliquer et communiquer une information mathématique dans des situations courantes afin de jouer pleinement, de manière essentielle et efficace, un vaste éventail de rôles dans la vie (Département de l'Éducation du Queensland, 1994). [Traduction libre]

Sur le plan de la portée générale, toutes ces définitions sont très semblables à celles de la compréhension de textes suivis et de la compréhension de textes schématiques formulées pour les besoins de l'ELCA et présentées dans une section antérieure. Bon nombre de conceptions de la numératie mettent l'accent sur l'application et l'utilisation pratiques ou fonctionnelles des connaissances et des compétences mathématiques pour composer avec la présence d'éléments mathématiques dans des situations réelles. Les adultes sont censés maîtriser plusieurs façons de réagir avec souplesse à une situation mathématique en vue d'atteindre un objectif, selon les besoins et les intérêts d'une personne dans un contexte donné (maison, collectivité, travail, etc.) ainsi que son attitude et ses croyances à l'égard de la numératie (Gal, 2000; Coben, O'Donoghue et FitzSimons, 2000).

La numératie ne se limite donc pas à l'application des compétences arithmétiques à l'information contenue dans des imprimés, qui était évaluée dans le cadre de l'EIAA. La numératie des adultes englobe également le sens des nombres, ainsi que des compétences en estimation, en mesure et en statistique. Étant donné le rôle envahissant de la numératie dans le

monde moderne, ce ne sont pas nécessairement que les situations courantes qui font appel aux pratiques de calcul, mais aussi de **nouvelles** situations.

Un autre élément important de la définition de la numératie tient au rôle des processus de communication. La numératie englobe non seulement la capacité d'une personne d'utiliser et d'appliquer ses compétences mathématiques de manière efficace et critique, mais aussi la capacité d'interpréter des messages textuels ou symboliques, ainsi que de communiquer au sujet de l'information mathématique et des processus de raisonnement (Marr et Tout, 1997; Gal, 1997).

Les définitions de la numératie énoncent explicitement que la numératie ne consiste pas seulement à effectuer des opérations avec des nombres, comme le terme le laisse entendre, surtout pour ceux qui connaissent bien les conceptions de la numératie des enfants, mais qu'elle englobe une foule de compétences et de connaissances mathématiques. En outre, au cours des dernières années, il a beaucoup été question de la relation entre les mathématiques et la numératie et du caractère « essentiel » de la numératie (Frankenstein, 1989; Steen, 2001). Johnston, par exemple, soutient que :

Savoir calculer, ce n'est pas qu'être capable de manipuler des nombres, ni même « posséder » les mathématiques apprises à l'école ou à l'université. La numératie suppose une connaissance essentielle qui constitue un lien entre les mathématiques et le monde réel, dans toute sa diversité (Johnston, 1994).
[Traduction libre]

Bon nombre d'auteurs soutiennent qu'en abordant les compétences fonctionnelles, on doit aussi tenir compte d'attitudes et de croyances « habilitantes ». Dans le domaine des compétences mathématiques des adultes, on s'attend à l'« aisance » avec les nombres ou à la « confiance » dans les compétences mathématiques, car ces caractéristiques ont une incidence sur la mise en pratique effective des compétences et des connaissances (Cockroft, 1982; Tobias, 1993).

À la brève définition de la numératie formulée pour les besoins de l'ELCA et présentée plus haut s'ajoute une définition plus large des **pratiques de calcul**, formulée par l'équipe d'évaluation de la numératie de l'ELCA pour servir de base à la mise au point des items de numératie de l'ELCA :

On met en œuvre les pratiques de calcul lorsqu'on gère une situation ou qu'on résout un problème dans un contexte réel; il s'agit de réagir à l'information sur des notions mathématiques pouvant être représentées de diverses façons et de mettre en œuvre une gamme de connaissances, de facteurs et de processus habilitants.

Cette conception des pratiques de calcul suppose que pour évaluer la numératie des adultes, il est nécessaire de produire des tâches et des items qui varient selon les contextes, les réactions demandées, la nature de l'information mathématique en question et les représentations de cette information. Ces caractéristiques des tâches sont définies ci-dessous. Cette conception est

beaucoup plus large que la définition de la compréhension de textes au contenu quantitatif formulée pour les besoins de l'EIAA. Ses concepts clés concernent, de façon générale, la gestion d'une situation et un éventail de réactions efficaces (et pas seulement l'application des compétences arithmétiques). Elle englobe une vaste gamme de compétences et de connaissances (et pas seulement les opérations de calcul) et une foule de situations qui présentent différents types d'information mathématique (et pas seulement ceux qui comprennent des **nombres** inclus dans des textes **imprimés**).

La mise au point des items visait à assurer qu'une certaine proportion de la banque d'items imposerait aux répondants un fardeau de lecture minimal, c'est-à-dire que certains stimuli comporteraient peu ou pas de texte, ce qui permettrait même aux répondants possédant une maîtrise limitée de la langue du test de comprendre la situation décrite. D'autres sections de la banque d'items comprenaient des items comportant des textes essentiels plus ou moins longs, selon les exigences de la situation représentée par l'item.

Comme le laissent entendre la documentation et les notions évoquées plus haut, la nature des réactions d'une personne aux exigences mathématiques et autres d'une situation dépend essentiellement de la mise en application de diverses bases de connaissances habilitantes (compréhension du contexte, connaissances et compétences dans les domaines des mathématiques, de la statistique et de la littératie), de processus de raisonnement, ainsi que de son attitude et de ses croyances à l'égard de la numératie. Les pratiques de calcul dépendent en outre de l'intégration des connaissances et des compétences mathématiques avec des compétences et des stratégies plus vastes en littératie et en résolution de problèmes ainsi qu'avec l'expérience et les comportements acquis qu'un adulte apporte à chaque situation. Il est clair que les pratiques de calcul consistent à tenter d'accomplir une tâche de manière autonome, sans la déléguer ni la gérer en faisant volontairement abstraction de son contenu mathématique.

4.4.1 Définir les caractéristiques des tâches

Pour mettre au point et représenter les tâches de numératie conçues pour les besoins de l'ELCA, on a retenu quatre caractéristiques clés des pratiques de calcul : le type d'objet ou de contexte, le type de réaction, le type d'information mathématique ou statistique et le type de représentation de cette information. Chacune d'elles est décrite ci-dessous.

Type d'objet ou de contexte. Une personne tente de gérer une situation de numératie ou d'y réagir parce qu'elle veut satisfaire un besoin ou atteindre un objectif. Quatre types de buts et d'objectifs sont décrits ci-dessous. Force est de constater qu'ils ne sont pas mutuellement exclusifs et qu'ils peuvent correspondre aux mêmes thèmes mathématiques sous-jacents.

4.4.2 Vie quotidienne

Les tâches de numératie qui se présentent dans des situations courantes sont souvent des tâches de gestion auxquelles une personne fait face dans sa vie personnelle et familiale. D'autres sont liées à des passe-temps, ainsi qu'au perfectionnement et aux intérêts personnels. Il peut s'agir des tâches suivantes : régler des questions d'argent et de budget, comparer des prix, gérer son temps personnel, prendre des décisions en matière de déplacements, planifier

des vacances, faire des calculs dans le cadre de passe-temps comme la courtépointe ou le travail du bois, jouer à des jeux de hasard, comprendre des statistiques et des résultats sportifs, lire des cartes géographiques et utiliser des mesures dans le cadre d'activités ménagères comme la cuisine ou le bricolage.

4.4.3 Travail

Au travail, on fait face à des situations à composante quantitative qui sont souvent plus spécialisées que celles qui se présentent dans la vie quotidienne. Dans ce contexte, une personne peut acquérir de bonnes compétences pour gérer des situations plus « pointues » dans l'application de thèmes mathématiques. Il peut s'agir des tâches suivantes : remplir des bons de commande, totaliser des recettes, calculer la monnaie, gérer des horaires, utiliser des feuilles de calcul, disposer et emballer des articles de formes différentes, remplir et interpréter des cartes de contrôle, effectuer et consigner des mesures, lire des plans, contrôler des dépenses, prévoir des coûts et appliquer des formules.

4.4.4 Vie sociale ou collective

Les adultes ont besoin de connaître les processus qui existent dans le monde qui les entoure, par exemple les tendances de la criminalité, des salaires et de l'emploi, de la pollution et des risques pour la santé ou l'environnement. Ils sont parfois appelés à prendre part à des activités sociales, communautaires ou politiques. Ils doivent donc savoir lire et interpréter l'information quantitative présentée dans les médias, y compris les messages et graphiques statistiques. Ils peuvent être appelés à gérer les situations suivantes : organiser une campagne de financement, calculer l'incidence budgétaire d'un programme communautaire ou interpréter les résultats d'une étude sur la plus récente tendance en matière de santé.

4.4.5 Formation complémentaire

Les compétences en numératie permettent à une personne de participer à des études complémentaires, que ce soit dans un but scolaire ou dans le cadre d'une formation professionnelle au niveau des métiers. Dans un cas comme dans l'autre, il importe de connaître certains aspects formels des mathématiques – symboles, règles et formules – et de comprendre certaines conventions régissant l'application des règles et des principes mathématiques.

Type de réaction. Dans différents types de situation réelle, une personne peut avoir à réagir de l'une ou de plusieurs des façons suivantes (la première va pratiquement de soi; les autres dépendent de l'interaction entre les exigences de la situation et les objectifs, les compétences, les dispositions et les connaissances acquises de la personne) :

Identifier ou repérer, dans la tâche ou la situation qui se présente, une information mathématique qui est pertinente à l'égard d'un besoin ou d'un objectif.

Réagir à l'information présente dans la situation. Bishop (1988), par exemple, distingue six modes d'actions mathématiques qui sont communs à toutes les cultures : compter, repérer,

mesurer, concevoir, jouer et expliquer. D'autres types d'actions ou de réactions peuvent entrer en jeu, comme faire des calculs (mentalement ou à l'aide d'une calculatrice), ordonner ou trier, estimer, mesurer ou modéliser (par exemple, en utilisant une formule).

Interpréter l'information comprise dans la situation (et les résultats d'une intervention antérieure) et comprendre ce qu'elle signifie ou suppose. Il peut s'agir de porter un jugement sur la façon dont une information mathématique ou des faits connus s'appliquent à la situation ou au contexte. Il peut être nécessaire de faire preuve de jugement pour déterminer qu'une réponse a du sens ou non dans le contexte donné (savoir, par exemple, qu'un résultat de « 2,35 voitures » n'est pas une solution valide lorsqu'on se demande combien de voitures sont nécessaires pour transporter un groupe). La situation peut aussi comporter un aspect critique si la personne remet en question l'objet de la tâche, la validité des données ou de l'information présentée, ainsi que le sens et la portée des résultats, pour elle-même sur le plan individuel et, peut-être, pour la collectivité dans son ensemble.

Communiquer à quelqu'un d'autre au sujet de l'information mathématique donnée ou des résultats d'une intervention ou d'une interprétation. On peut le faire oralement ou par écrit (il peut s'agir simplement d'un nombre ou d'un mot ou encore d'une explication ou d'une analyse détaillée) ou au moyen d'un dessin (diagramme, carte géographique, graphique).

Type d'information mathématique ou statistique. On peut classer l'information mathématique d'un certain nombre de façons et selon différents niveaux d'abstraction. Une démarche consiste à se reporter aux « grands principes » fondamentaux du domaine des mathématiques. Steen (1990), par exemple, distingue les six grandes catégories suivantes : quantité, dimension, structure, forme, incertitude et variation. Rutherford et Ahlgren (1990) décrivent des réseaux de notions connexes : nombres, formes, incertitude, données récapitulatives, échantillonnage et raisonnement. Dossey (1997) catégorise comme suit les comportements mathématiques en matière de compréhension de textes au contenu quantitatif : représentation et interprétation de données, sens des nombres et des opérations, mesure, variables et relations, formes géométriques et visualisation spatiale, et hasard. L'équipe d'évaluation de la numératie de l'ELCA s'est inspirée de ces catégorisations et d'autres catégorisations étroitement liées (par ex., National Council of Teachers of Mathematics, 2000) pour en arriver à un ensemble de cinq notions fondamentales qui, selon elle, caractérisent les exigences mathématiques que les adultes doivent satisfaire dans diverses situations en ce début du XXI^e siècle.

4.4.6 Quantité et nombre

Selon Fey (1990), la notion de **quantité** procède du besoin de quantifier le monde qui nous entoure en utilisant des attributs comme les suivants : longueur, surface et volume des cours d'eau ou des masses continentales; température, humidité et pression atmosphérique; populations et taux de croissance des espèces; amplitudes des marées; revenus ou bénéfices d'entreprises.

La notion de **nombre** est fondamentale pour la quantification, et différents types de nombres déterminent la quantification de diverses façons : les nombres entiers servent de compteurs ou d'estimateurs; les fractions, décimales et pourcentages expriment des quantités plus précises ou représentent les parties d'un tout, ce qui permet de comparer des proportions. Les nombres

positifs et négatifs servent d'indicateurs directionnels. En plus de la quantification, on utilise les nombres pour ordonner et pour identifier (par exemple, dans le cas de numéros de téléphone ou de codes postaux). L'aisance avec la quantité, les nombres et les opérations sur les nombres exige un bon « sens » de la grandeur et de la signification de nombres très grands ou très petits, et parfois le sens de la grandeur relative de différentes proportions.

La gestion de l'argent et du temps, aspect omniprésent des mathématiques dans la vie des adultes, dépend d'un sens aigu des nombres et des quantités. On doit faire preuve de jugement contextuel pour déterminer dans quelle mesure il faut être précis ou quel outil utiliser (calculatrice, calcul mental, ordinateur). Une tâche de numération de base peut consister à calculer le coût d'une boîte de soupe en sachant que quatre boîtes coûtent 2,00 \$; une tâche dont l'exigence cognitive est élevée peut nécessiter des nombres plus « pointus », par exemple si l'on calcule le coût d'un kilo de fromage en sachant que 0,783 kg coûte 12,95 euros.

4.4.7 Dimension et forme

La **dimension** repose sur les « grands principes » liés à une, deux et trois dimensions des « choses ». On doit faire appel à la connaissance des dimensions lorsqu'il s'agit de lire ou de produire des descriptions spatiales ou numériques d'objets, de faire des projections ou de travailler avec des longueurs, des périmètres, des plans, des surfaces, un emplacement, etc. L'aisance avec chaque dimension exige un sens des mesures « repères », des mesures directes et des mesures estimatives.

La **forme** est une catégorie qui décrit des images et des entités réelles ou imaginaires qu'on peut visualiser (maisons, immeubles, motifs d'art ou d'artisanat, enseignes de sécurité, emballages, nœuds, cristaux, silhouettes, plantes, etc.). L'orientation et l'emplacement sont des qualités fondamentales auxquelles on a recours pour lire ou tracer des cartes géographiques et des diagrammes. Une tâche de numération de base liée à cet aspect fondamental pourrait consister à identifier des formes, et une tâche complexe, à décrire la variation de la taille ou du volume d'un objet dont une dimension est modifiée, par exemple lorsqu'on choisit entre différentes boîtes pour emballer certains objets.

4.4.8 Structure, fonctions et relations

On écrit souvent que les mathématiques sont l'étude des structures et des relations. La structure apparaît comme un vaste concept qui englobe les structures qu'on trouve tout autour de nous, par exemple sous forme musicale, dans la nature et dans les mouvements de la circulation. Senechal (1990) soutient que notre capacité de reconnaître, d'interpréter et de créer des structures est essentielle pour appréhender le monde qui nous entoure. La capacité humaine d'identifier des relations et d'analyser sous-tend le raisonnement mathématique. L'algèbre – au-delà de la manipulation symbolique – offre un outil servant à représenter des relations entre des montants au moyen de tableaux, de graphiques, de symboles et de mots. La capacité de généraliser et de caractériser des fonctions et des relations entre des variables est essentielle à la compréhension des analyses économiques, politiques ou sociales, même les plus simples. Une tâche de reconnaissance de structure relativement simple pourrait consister à demander à quelqu'un de décrire la structure d'une suite de nombres ou de formes données ou, dans un contexte fonctionnel, à comprendre la relation entre des listes de variables (par.

ex., le poids et le volume de certains objets); élaborer une formule pour un tableur imposerait à la personne un niveau d'exigence plus élevé.

4.4.9 Données et hasard

Les données et le hasard recouvrent deux sujets connexes mais distincts. Les **données** reposent sur des « grands principes » tels que la variabilité, l'échantillonnage, l'erreur ou la prévision et sur des sujets statistiques connexes tels que la collecte et l'analyse de données, les mesures courantes d'une valeur centrale ou d'une fourchette de valeurs, ou la notion d'une inférence statistique. La société moderne exige des adultes qu'ils interprètent (et parfois même qu'ils produisent) des tableaux de fréquence, des diagrammes et des graphiques de base, des renseignements sur des moyennes et des médianes, et qu'ils reconnaissent des affirmations statistiques douteuses (Gal, 2002).

Le **hasard** repose sur des « grands principes » liés à la probabilité et aux concepts et outils statistiques pertinents. Peu de choses au monde sont absolument certaines; la capacité d'associer un nombre à la probabilité d'un événement (et de ses risques ou effets secondaires) constitue donc un outil précieux, qu'il s'agisse du temps qu'il fait, de la bourse des valeurs ou de la décision de prendre un certain médicament. Dans cette catégorie, une tâche de numératie simple pourrait consister à interpréter un diagramme circulaire simple ou à comprendre un énoncé relatif à une moyenne; une tâche plus complexe consisterait à déduire la probabilité d'un événement d'après l'information fournie.

4.4.10 Variation

Il s'agit des mathématiques de l'évolution du monde autour de nous. Les organismes croissent, les populations varient, les prix fluctuent, les objets en mouvement accélèrent et ralentissent. La variation et les taux de variation nous aident à décrire l'état du monde au fil du temps. Les structures additives, multiplicatives et exponentielles de la variation peuvent caractériser des tendances stables; les variations périodiques dénotent des cycles, et les courbes de variation irrégulières évoquent la théorie du chaos. Décrire une perte de poids constitue une tâche relativement simple, alors que calculer un intérêt composé est une tâche relativement complexe.

Type de représentation de l'information mathématique. L'information mathématique inhérente à une activité ou à une situation peut exister ou être représentée sous plusieurs formes. Il peut s'agir d'objets concrets à compter (moutons, personnes, immeubles, voitures, etc.) ou d'illustrations de ces objets. Elle peut être représentée au moyen d'une notation symbolique (chiffres, lettres et signes d'opération). Parfois, l'information mathématique figure dans des formules, qui constituent un modèle de relations entre des entités ou des variables.

En outre, l'information mathématique peut être codée de manière visuelle, comme dans un **diagramme**; on peut utiliser des **graphiques** et des **tableaux** pour présenter une information statistique ou quantitative agrégée. De même, une **carte géographique** d'une entité réelle (par exemple, le plan d'une ville ou d'un chantier) peut contenir des données numériques, mais aussi une information qu'on peut quantifier ou mathématiser.

Enfin, une personne peut être appelée à extraire une information mathématique de divers types de textes suivis ou schématiques de formes précises (par exemple, des déclarations de revenus). Les tâches de numératie fonctionnelle portent sur deux sortes différentes de texte. La première comporte une information mathématique représentée sous forme textuelle, c'est-à-dire accompagnée de mots ou d'expressions ayant un sens mathématique. On peut citer comme exemples l'utilisation de chiffres en toutes lettres (par exemple, « cinq » au lieu de « 5 »), de termes mathématiques de base (par exemple, fraction, multiplication, pourcentage, moyenne, proportion) ou d'expressions plus complexes (par exemple, « le taux de criminalité a chuté de moitié ») qu'il faut interpréter. Dans le deuxième type de texte, l'information mathématique est exprimée en notations ou symboles réguliers (nombres, signes plus ou moins, symboles d'unités de mesure, etc.), mais elle est entourée d'un texte qui, malgré sa nature non mathématique, doit aussi être interprété afin de fournir une information et un contexte supplémentaires. On peut citer l'exemple d'un bordereau de dépôt bancaire comportant un texte et des directives dans lesquels sont inclus des nombres correspondant à des montants monétaires.

4.4.11 Caractériser les tâches de numératie

Pour les besoins de l'ELCA, on a retenu au total 40 tâches de numératie. Ces tâches s'échelonnent de 174 à 380 le long de l'échelle de la numératie, leur emplacement étant déterminé par la façon dont les adultes des pays participants ont réagi à chaque tâche. On trouvera ci-dessous quelques exemples de tâches qui reflètent certains aspects conceptuels de la numératie et les principes de conception de l'échelle décrits plus haut, soit le calcul, le raisonnement spatial et proportionnel, la mesure et les connaissances statistiques.

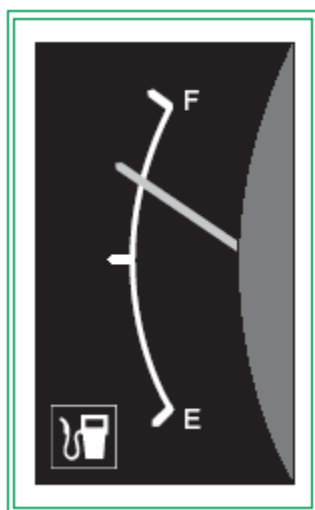
La tâche la plus simple sur l'échelle de la numératie (174) présentait aux adultes une photographie montrant deux cartons de bouteilles de Coca-Cola. On leur demandait de trouver le nombre total de bouteilles que contenaient les deux caisses pleines. Si cette tâche était simple, c'est d'abord parce qu'elle était empruntée à la vie quotidienne et que les objets en question étaient relativement familiers pour la plupart des gens. Ensuite, ce qu'on demandait aux répondants était évident et explicite : cette tâche faisait appel à une photographie représentant des objets concrets et ne comportait aucun texte à traiter. Un troisième facteur tenait au fait que les répondants pouvaient accomplir la tâche de diverses façons plus ou moins complexes, par exemple en multipliant des rangées par des colonnes, mais aussi par simple comptage. Cette tâche amène le répondant à faire une conjecture, car la totalité des bouteilles de la caisse du dessous n'est pas visible, mais, comme l'indique le faible niveau de difficulté de la tâche, cette caractéristique ne présentait pas un problème pour la vaste majorité des adultes de tous les pays participants.



Une deuxième tâche, également très simple, exigeait des adultes qu'ils lisent un court texte rapportant les résultats d'une élection qui opposait trois candidats et de déterminer le nombre total de votes exprimés. Cette tâche (indice de difficulté : 192) se situe au niveau 1 de l'échelle de la numératie. Encore une fois, on mettait les répondants dans un type de situation réaliste en leur présentant une information numérique simple, sous forme de colonnes montrant le nom de chaque candidat et le nombre de votes obtenus par chacun. La tâche ne comportait aucune autre information numérique pouvant constituer un élément de distraction. Pour trouver le nombre total de votes exprimés lors de l'élection, il suffit d'effectuer une seule addition, ce que la question rend explicite par l'emploi du mot-clé « total », et le calcul porte sur des nombres entiers relativement peu élevés.

Une tâche de numératie plus complexe, située au milieu du niveau 2 (indice de difficulté : 248) amène les adultes à regarder un indicateur de niveau d'essence. Cet indicateur comporte trois lignes ou coches dont l'une est marquée « F » pour « full » (plein), une autre est marquée « E » pour « empty » (vide) et la troisième se situe au milieu, entre les deux. Une ligne représentant l'aiguille de l'indicateur montre un niveau qui se trouve à peu près à mi-chemin entre la coche du milieu et la coche marquée « F », montrant ainsi que le réservoir est environ aux trois quarts plein. La directive énonce que le réservoir peut contenir 48 gallons et demande au répondant de déterminer combien de gallons il reste dans le réservoir. Cette tâche est empruntée à un contexte quotidien et amène le répondant à interpréter une présentation comportant une information quantitative, mais pratiquement pas de texte ni de nombres. Aucune information mathématique n'est présente, à part celle qui est fournie dans la question.

Ce qui rend cette tâche plus difficile que les deux précédentes est le fait que le répondant doit d'abord estimer le niveau d'essence restant dans le réservoir en convertissant l'emplacement de l'aiguille en une fraction. Puis, il doit déterminer combien de gallons cette fraction représente par rapport à la capacité de 48 gallons énoncée dans la question ou la directive. Cette tâche l'amène donc à effectuer plus d'une opération pour arriver à une bonne réponse, sans préciser le type d'opération. Néanmoins, comme bien d'autres tâches de numératie courantes, cette tâche n'exige pas qu'on effectue un calcul exact, mais une approximation qui devrait se situer à l'intérieur de limites raisonnables.

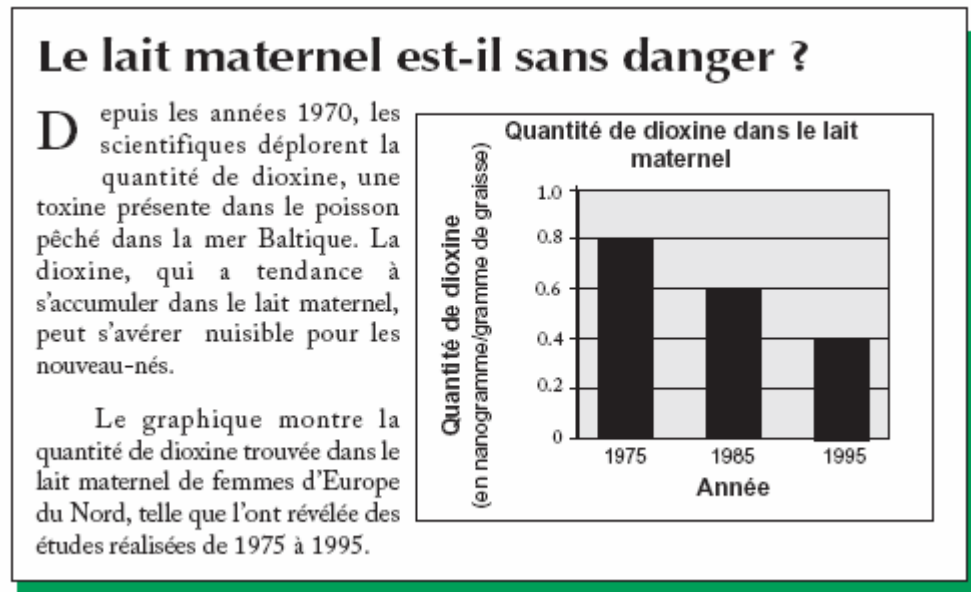


Une tâche de numératie un peu plus difficile, située au sommet du niveau 2 (indice de difficulté : 275) présente aux adultes le diagramme d'un conteneur sur lequel se trouvent quatre marques ou lignes; on demande aux répondants de tracer sur le conteneur une ligne à la hauteur du tiers. La ligne du haut est marquée « 1 », alors que celle du milieu est marquée « $\frac{1}{2}$ ». Les deux autres lignes sont sans marque; l'une est située à mi-chemin entre « 1 » et « $\frac{1}{2}$ » et l'autre, à mi-chemin entre la ligne marquée « $\frac{1}{2}$ » et le fond du conteneur. Pour répondre correctement, le répondant doit tracer sur le conteneur une ligne entre la ligne marquée « $\frac{1}{2}$ » et la ligne au-dessous à la hauteur du quart (mais cette ligne n'est pas marquée « $\frac{1}{4}$ » – il faut le déduire). Ici, le contexte peut être moins familier pour le répondant mais, encore une fois, l'image visuelle est simple et réaliste et ne comporte pratiquement pas de texte. Il ne s'agit pas d'écrire un symbole ou un texte, mais simplement de tracer une ligne à un certain endroit sur le dessin du conteneur. Pour accomplir correctement cette tâche, le répondant doit posséder une certaine connaissance pratique des fractions et le sens des proportions : il doit connaître les symboles « $\frac{1}{2}$ » et « $\frac{1}{3}$ », savoir ordonner les fractions en fonction de leur taille relative et les faire correspondre aux marques existantes sur le conteneur.

On a mis au point quelques tâches de numératie axées sur un bref article de journal intitulé « Le lait maternel est-il sans danger? », au sujet des risques environnementaux et de la salubrité des aliments. L'article comportait deux brefs paragraphes concernant une toxine, la dioxine, qu'on trouve dans le poisson de la mer Baltique, ainsi qu'un graphique à barres indiquant les niveaux de dioxine observés à trois moments – 1975, 1985 et 1995 – dans le lait maternel des femmes de l'Europe du Nord. Une question demandait au répondant de décrire comment la quantité de dioxine avait varié de 1975 à 1995, c'est-à-dire d'interpréter directement les données présentées dans le graphique. On ne lui demandait pas de calculer la variation exacte observée au cours de chacune des périodes, mais simplement de décrire en ses propres mots la variation des niveaux de dioxine (avaient-ils diminué, augmenté ou étaient-ils restés inchangés?).

On a attribué à cette tâche un indice de difficulté de 280, au bas du niveau 3. Le graphique montre clairement que la quantité de dioxine a diminué au cours de chacune des trois périodes; pourtant, certains adultes ont du mal à accomplir cette tâche, qui est fondée sur un stimulus

dont la structure est courante dans les journaux, soit un texte bref accompagné d'un graphique. Le niveau de difficulté accru de cet item est sans doute attribuable en partie au fait que le répondant doit formuler sa propre description, au degré moyen de dépendance au texte pour comprendre le contexte auquel se rapporte le graphique, ou à la nécessité de comprendre l'orientation des valeurs décimales sur l'axe vertical (mode courant de représentation des concentrations de produits chimiques polluants).



Dans une deuxième tâche axée sur le même stimulus, mais plus difficile, on demandait aux adultes de comparer le pourcentage de variation du niveau de dioxine de 1975 à 1985 à celui de 1985 à 1995, de déterminer lequel était le plus élevé et d'expliquer leur réponse. Cette tâche s'étant avérée beaucoup plus difficile pour les adultes des pays participants, on lui a attribué un indice de difficulté de 377 sur l'échelle de la numératie. Ici, l'information nécessaire est intégrée au graphique et exige une certaine part de transformation et d'interprétation. Pour arriver à une bonne réponse, le répondant doit prendre en compte le taux de variation exprimé en pourcentage, et pas seulement la valeur absolue de la variation. En outre, il doit composer avec des pourcentages d'entités inférieures à un (c'est-à-dire les valeurs décimales sur l'axe vertical) et savoir que la base de calcul de la variation en pourcentage varie d'une paire à l'autre. Il semble que la nécessité de prendre en compte ces caractéristiques de la tâche, d'employer des méthodes mathématiques formelles ou de composer avec la notion abstraite de taux de variation ajoute considérablement à la difficulté de ce genre de tâche.

La tâche de numératie la plus difficile de cette évaluation (indice de difficulté : 380; niveau 5) présentait aux adultes une publicité selon laquelle il est possible, pour un investisseur, de doubler en sept ans un montant investi au taux d'intérêt fixe de 10 % par année. On demandait aux répondants s'il était possible de doubler en sept ans un montant de 1 000 \$ investi à ce taux et d'étayer leur réponse par leurs calculs. On a accepté différentes réponses à condition qu'elles soient accompagnées d'une justification raisonnable et des calculs pertinents. Les répondants étaient libres d'effectuer le calcul à leur manière, mais pouvaient aussi utiliser un « tuyau financier » qui accompagnait la publicité et présentait une formule pour estimer la

valeur d'un investissement après un certain nombre d'années. Ceux qui employaient la formule devaient assimiler l'information énoncée dans le texte aux variables de la formule (principal, taux d'intérêt et période), puis effectuer les calculs nécessaires et comparer le résultat au montant qu'on obtiendrait si le montant de 1 000 \$ était doublé.

Tous les répondants pouvaient utiliser une calculatrice de poche fournie au moment de l'évaluation. Cette tâche s'est avérée difficile parce qu'elle portait sur des pourcentages et que le calcul, avec ou sans la formule, exigeait qu'on intègre plusieurs étapes et plusieurs types d'opération. Pour effectuer les calculs sans la formule, il était nécessaire de comprendre la méthode de calcul des intérêts composés. Cette tâche permettait aux répondants de recourir à différentes stratégies de raisonnement, y compris des méthodes non structurées ou de leur cru. Pourtant, à l'instar de la tâche qui consistait à comparer des taux de variation, elle exigeait le recours à une information mathématique formelle et à une compréhension approfondie de méthodes de calcul peu courantes, ce qui n'était peut-être pas familier ou accessible à bien des adultes.

4.5 Mesurer la résolution de problèmes dans le cadre de l'ELCA

La résolution de problèmes fait depuis longtemps l'objet de recherches psychologiques universitaires et de recherches appliquées en ressources humaines. Voici une définition très générale de la résolution de problèmes telle qu'on l'entend habituellement dans le domaine de la recherche psychologique (Hunt, 1994; Mayer, 1992; Mayer et Wittrock, 1996; Smith, 1991) :

La résolution de problèmes correspond à la pensée et à l'action orientées vers les buts dans une situation où il n'existe aucune procédure courante de résolution. La personne qui résout des problèmes a un but plus ou moins bien défini, mais ne sait pas immédiatement comment l'atteindre. La non-congruence des buts et des opérateurs admissibles constitue un problème. La compréhension de la situation problématique et sa transformation progressive fondée sur la planification et le raisonnement constituent le processus de la résolution de problèmes.

Lorsqu'on élabore un cadre de travail de la résolution de problèmes en vue d'une enquête comme l'ELCA, l'une des principales difficultés consiste à adapter le mieux possible la documentation psychologique aux contraintes imposées par une étude comparative internationale à grande échelle. À cette fin, on a décidé de mettre l'accent sur un sous-ensemble essentiel de résolution de problèmes : la résolution de problèmes analytiques. Il faut éviter de confondre notre notion de la résolution de problèmes analytiques avec le sens courant et intuitif de l'expression et avec la notion clinico-psychologique selon laquelle la résolution de problèmes est associée à la résolution de conflits sociaux et émotifs. Pourtant, le contexte social est tout à fait pertinent pour notre définition de la résolution de problèmes analytiques, par exemple lorsqu'il faut aborder un problème interactivement et le résoudre par la collaboration. Des facteurs motivants comme l'intérêt pour le sujet et l'orientation des tâches influencent également le processus de résolution des problèmes. En revanche, la qualité de la résolution est déterminée avant tout par la compréhension de la situation problématique, les processus de pensée utilisés pour aborder le problème et l'à-propos de la solution retenue.

On peut faire appel à différents aspects pour caractériser le **problème** lui-même :

- ❖ Le **contexte** peut refléter divers domaines, de nature théorique ou pratique, rattachés à une situation intellectuelle ou au monde réel. Dans ces domaines, un problème peut s'avérer plus ou moins authentique.
- ❖ **L'ampleur** d'un problème est variable; une personne peut s'en tenir à des aspects limités et concrets d'une tâche, planifier et exécuter des actions complexes ou évaluer des séquences d'actions multiples.
- ❖ Le problème peut comporter un but bien ou mal défini, des contraintes transparentes (explicites) ou non, un nombre restreint d'éléments indépendants ou un grand nombre d'éléments interdépendants. Ces caractéristiques définissent la **complexité** du problème.

La connaissance du contexte que possède la population cible, le fait que le problème porte sur des tâches concrètes ou des actions complexes, la mesure dans laquelle le but est bien défini, la transparence des contraintes, le nombre d'éléments que la personne qui résout des problèmes doit prendre en compte et l'ampleur de leur interdépendance – voilà autant de caractéristiques qui déterminent le niveau de compétence nécessaire pour résoudre un problème donné. La difficulté empirique, c'est-à-dire la probabilité de trouver une solution correcte, dépend de la relation entre ces caractéristiques du problème, d'une part, et le niveau de compétence des sujets, de l'autre.

Les **processus cognitifs** actionnés au cours de la résolution d'un problème sont divers et complexes, et sont plutôt organisés de manière non linéaire. Ces processus comportent notamment les cinq éléments suivants :

1. Chercher l'information, puis la structurer et l'intégrer en une représentation intellectuelle du problème (« modèle situationnel »).
2. Raisonner à la lumière du modèle situationnel.
3. Planifier des actions et d'autres étapes de la solution.
4. Exécuter et évaluer les étapes de la solution.
5. Traiter l'information externe et la rétroaction sur une base soutenue.

Baxter et Glaser (1997) présentent une liste semblable d'activités cognitives, désignées les « composantes générales de la compétence en résolution de problèmes » : représentation du problème, stratégies de résolution, autorégulation et explications. La résolution de problèmes analytiques dans un contexte courant, mesurée par l'instrument de résolution de problèmes de l'ELCA, met l'accent sur les trois premières composantes (et, dans une certaine mesure, sur la quatrième).

L'un des constats les plus importants des récents travaux de recherche en psychologie cognitive est que la résolution de problèmes difficiles exige au moins une connaissance élémentaire du domaine en question. La notion d'un espace de problème dans lequel la personne qui résout des problèmes évolue au moyen de stratégies de recherche indépendantes du domaine (Newell et Simon, 1972) s'est avérée trop simple pour décrire la manière d'appréhender une situation problématique et la recherche d'une solution. Les efforts

en vue d'identifier une compétence générale, indépendante du domaine et susceptible de guider les systèmes dynamiques (l'intelligence opérative) au sein du cadre de la recherche en résolution de problèmes complexes ont également été infructueux; seule une partie du rendement de ces systèmes peut être transférée à un autre système (Funke, 1991). En revanche, des recherches réalisées auprès des élèves de la troisième à la douzième années ont révélé que l'aptitude à résoudre des problèmes s'améliorait manifestement dans des conditions de formation bien rodées et qu'il était possible de réaliser un transfert appréciable entre différents problèmes (Reeff *et coll.*, 1989, 1992, 1993; Regenwetter, 1992; Regenwetter et Müller, 1992; Stirner, 1993).

La résolution de problèmes dépend de la connaissance des notions et des faits (connaissance déclarative) ainsi que des règles et des stratégies (connaissance procédurale) dans un domaine donné. S'il est évident, à la lumière des recherches passées, que la connaissance déclarative du domaine problème peut favoriser considérablement la réussite des stratégies de résolution de problèmes, la connaissance procédurale est tout aussi essentielle. La quantité de connaissances préalables pertinentes dont on dispose peut également expliquer le rapport entre l'intelligence et le rendement en résolution de problèmes, présenté dans les travaux de Raaheim (1988) et Leutner (1999). Les personnes n'ayant absolument aucune connaissance préalable pertinente sont incapables d'approfondir la situation problématique et de planifier systématiquement une solution; elles doivent plutôt procéder par tâtonnements. Celles qui connaissent déjà très bien la tâche sont en mesure de l'aborder de manière systématique. L'aptitude mentale générale, mesurée par des tâches de raisonnement, n'intervient dans aucun de ces cas. Lorsque la tâche est moyennement familière à la personne qui résout des problèmes, celle-ci peut mettre en œuvre des stratégies de raisonnement analytiques.

L'approche adoptée pour l'évaluation de la résolution de problèmes dans le cadre de l'ELCA est fondée sur la notion de tâches (moyennement) familières. Dans un contexte assez familier, les problèmes à résoudre sont suffisamment peu « explicites » pour ne pas être perçus comme des tâches purement routinières. En revanche, les connaissances préalables spécifiques au domaine sont assez limitées pour faire des techniques de raisonnement analytiques le principal outil cognitif de résolution des problèmes.

4.5.1 Définir les caractéristiques des tâches

Comment définir des problèmes réels contextualisés et les transformer en un ensemble de tâches d'évaluation? Après avoir passé en revue les diverses approches adoptées lors de recherches antérieures pour mesurer la résolution de problèmes, les concepteurs de l'ELCA ont opté pour une approche par projet. L'approche par projet peut s'avérer un instrument puissant lorsqu'il s'agit d'évaluer les compétences en résolution de problèmes analytiques dans des contextes réels et courants, et ce, pour plusieurs raisons. Il est à la fois important et pertinent, pour la vie professionnelle et privée de l'adulte, que celui-ci puisse résoudre des problèmes dans un contexte de projet. De plus, l'approche par projet a été appliquée avec succès dans d'autres évaluations à grande échelle et elle peut se réaliser en tant qu'outil de type crayon-papier, ce qui a une importance déterminante pour les enquêtes contemporaines à grande échelle. En outre, l'approche par projet utilise différentes étapes de la résolution de problèmes pour orienter la mise au point des items de test proprement dits. Dans la foulée de Pólya (1945, 1980), le processus de résolution de problèmes a souvent été réduit aux étapes suivantes :

- ❖ Définir le but.
- ❖ Analyser la situation et construire une représentation mentale.
- ❖ Concevoir une stratégie et planifier les mesures à prendre.
- ❖ Exécuter le plan, en y incorporant le contrôle et – s'il y a lieu – la modification de la stratégie.
- ❖ Évaluer le résultat.

Les diverses étapes définissent le cours de l'action d'un projet « courant ». À chacune de ces étapes correspondent un ou plusieurs items ou tâches. Ainsi, les répondants exécutent les tâches individuelles qui ont été identifiées comme des étapes à franchir dans le cadre de leur projet (par exemple, « planifier une réunion de famille » ou « rénover un pavillon »). En incorporant les tâches individuelles dans le contexte d'une action, on obtient une grande authenticité contextuelle. Même si elles font toutes partie d'un projet cohérent et exhaustif, les tâches individuelles sont conçues pour être résolues indépendamment les unes des autres et sont appelées à varier en complexité et en difficulté globale pour les adultes.

Comme l'évaluation des compétences en résolution de problèmes a des origines relativement récentes, il convient de rendre compte en détail du processus de construction. Le tableau A1 présente un aperçu des étapes de la résolution de problèmes qui correspondent aux diverses étapes susmentionnées. Il énumère divers aspects et composantes de chacune des étapes de la résolution de problèmes.

Tableau A1 Résolution de problèmes : étapes et mesures à prendre

Définir les buts	<ul style="list-style-type: none">• Fixer les buts.• Reconnaître les buts à atteindre et préciser les motifs importants de la décision.• Reconnaître les buts/souhaits contradictoires et ceux qui sont compatibles.• Attribuer des priorités aux buts/souhaits.
Analyser la situation	<ul style="list-style-type: none">• Sélectionner, obtenir et évaluer l'information.<ul style="list-style-type: none">⇒ De quelle information a-t-on besoin, quelle information est déjà disponible, laquelle manque toujours et laquelle est superflue?⇒ Où et quand peut-on obtenir l'information?⇒ Comment faut-il interpréter l'information?• Identifier les personnes (possédant les connaissances et les compétences nécessaires) qui doivent participer à la résolution du problème.• Choisir les outils à utiliser.• Reconnaître les conditions (contraintes temporelles, par exemple) dont il faut tenir compte.
Planifier la solution	<ul style="list-style-type: none">• Reconnaître les étapes à suivre.• Décider de la séquence des étapes (par exemple, les questions à l'ordre du jour).• Coordonner le travail et les délais.• Procéder à une analyse comparative de différents plans (reconnaître le plan qui convient à la réalisation des buts).• Adapter le plan aux changements de conditions.• Sélectionner un plan.
Exécuter le plan	<ul style="list-style-type: none">• Réaliser chacune des étapes (par exemple, rédiger une lettre, remplir un formulaire, faire des calculs).
Évaluer les résultats	<ul style="list-style-type: none">• Évaluer si, et dans quelle mesure, on a atteint la cible.• Reconnaître les erreurs.• Identifier pourquoi les erreurs ont été commises.• Évaluer les conséquences des erreurs.

La construction d'une banque de tâches d'évaluation qui correspondraient à ces cinq étapes comportait elle-même plusieurs étapes. Premièrement, il fallait choisir des projets pertinents qui conviendraient aux adultes ayant des antécédents scolaires variables et qui s'appliqueraient au plus grand nombre possible de personnes au sein du groupe cible. Puis, les concepteurs devaient établir et définir la situation problématique et une séquence d'actions correspondant au modèle. Troisièmement, ils devaient constituer une banque d'items correspondant à chaque étape et faisant appel à certains processus; il fallait notamment mettre au point les bonnes réponses et les éléments de distraction pertinents pour les items à choix multiple et établir une clé de correction et des guides de notation pour les tâches ouvertes.

4.5.2 Caractériser les tâches de résolution de problèmes

L'ELCA comportait au total quatre projets, dont vingt tâches consacrées à l'évaluation de la résolution de problèmes. On a ainsi produit 19 items à noter qui s'échelonnaient de 199 à 394 le long de l'échelle et, à l'instar des tâches de littératie et de numératie, leur emplacement était déterminé par les tendances des bonnes et des mauvaises réponses chez les adultes des pays participants. Au lieu de publier l'un des quatre projets qui ont servi à l'ELCA, nous allons décrire une échelle hypothétique de compétences en résolution des problèmes analytiques qui a été testée à partir des données de l'étude pilote et présenter un exemple, tiré de l'étude pilote, qui n'a pas servi à l'évaluation principale. On a décrit des modèles semblables dans le cadre de

travail d'autres évaluations à grande échelle des compétences en résolution de problèmes, comme le test projet pour Hambourg, en Allemagne (Ebach, Klieme et Hensgen, 2000) et l'évaluation de l'étude PISA 2003 de la résolution de problèmes transdisciplinaires (OCDE, sous presse).

Pour les besoins de l'ELCA, quatre niveaux de compétence en résolution de problèmes sont postulés :

Niveau 1

À un niveau très élémentaire, une personne peut maîtriser des tâches concrètes et limitées en appliquant un raisonnement pratique axé sur le contenu. À ce niveau, la personne fait appel à des schémas spécifiques axés sur le contenu afin de résoudre les problèmes.

Niveau 2

Le deuxième niveau exige au minimum un niveau rudimentaire de raisonnement systématique. Les problèmes à ce niveau se caractérisent par des buts unidimensionnels bien définis; ils appellent l'évaluation de certaines solutions de rechange au regard de contraintes transparentes et explicites. À ce niveau, la personne fait appel à des opérations logiques et concrètes.

Niveau 3

Au troisième niveau de compétence en résolution de problèmes, la personne est en mesure de faire appel à des opérations formelles (l'ordonnancement, par exemple) afin d'intégrer des buts multidimensionnels ou mal définis et de composer avec des contraintes dépendantes multiples ou non transparentes.

Niveau 4

Au niveau supérieur de compétence, la personne est à même de comprendre, dans son ensemble, un système d'états des problèmes et des solutions envisageables. Ainsi, la cohérence de certains critères, l'interdépendance de séquences multiples d'intervention et d'autres « métacaractéristiques » d'une situation problématique peuvent être considérées systématiquement. C'est également à ce stade que la personne est en mesure d'expliquer comment et pourquoi elle est arrivée à une solution donnée. Ce niveau de compétence exige une pensée critique et une certaine part de métacognition.

L'exemple ci-dessous illustre la réalisation concrète d'un projet. Le projet en question, exclu de l'instrument final de l'ELCA, consiste à exécuter une tâche typique de la résolution de problèmes. Il s'agit de planifier un voyage et une réunion de famille.

Dans l'introduction au projet, le répondant prend connaissance du résumé suivant, où sont décrits le scénario et le problème global :

« Imaginez que vous habitez la ville A. Vos parents sont éparpillés aux quatre coins du pays, et vous aimeriez organiser une réunion de famille, qui durera une journée. Vous décidez de vous réunir dans la ville B, facilement accessible pour tous. Comme vous et

vos parents aimez tous la randonnée pédestre, vous décidez de prévoir une longue randonnée dans un parc provincial près de la ville B. Vous avez accepté de prendre en charge la plus grande partie de l'organisation. »

Le répondant reçoit ensuite la liste suivante d'étapes à franchir :

- ❖ Fixez la date de la réunion
- ❖ Évaluez les recommandations de vos parents en vue de la randonnée
- ❖ Planifiez la marche à suivre avant de réserver votre billet d'avion
- ❖ Répondez aux questions de vos parents concernant le voyage en avion
- ❖ Réservez votre billet d'avion
- ❖ Veillez à ce que votre billet ne comporte aucune erreur
- ❖ Planifiez le déplacement entre la ville B et l'aéroport

La première tâche de ce projet, soit « Fixez la date de la réunion », constitue un bon exemple d'une tâche de résolution de problèmes, et nous la présentons maintenant comme elle figurerait dans un cahier de test.

Exemple de tâche : Fixez la date de la réunion

La réunion de famille devrait se tenir au mois de juillet.

Vous avez demandé à tous vos parents de vous faire part des dates qui leur conviendraient. Après leur avoir parlé, vous avez dressé la liste des engagements de vos parents pendant le mois de juillet. Votre propre agenda est ouvert devant vous. Vous constatez que certaines personnes devront arriver un jour avant la réunion et qu'elles ne pourront rentrer chez elles que le lendemain de la réunion.

Veillez consulter la liste des engagements de vos parents et votre propre agenda.

Liste des engagements de vos parents en juillet 1999

Henri	Karine	Pierre	Janette	Anne	François
Vacances à E à compter du 26 juillet Rendez-vous le 11 juillet	N'importe quel jour de la semaine sauf le jeudi et le 16 juillet	Rendez-vous d'affaires les 2 et 13 juillet et entre les 27 et 29 juillet	Aucun engagement	Ne peut se libérer les 5, 20 et 24 juillet	Doit faire un voyage d'affaires la première semaine complète de juillet, mais n'en connaîtra les dates exactes que la veille de son départ

Henri, Karine et Pierre pourraient arriver le jour prévu pour la réunion, alors que Janette, Anne et François ne peuvent arriver que la veille, en après-midi, et ne rentrer chez eux que le lendemain de la réunion.

Exemple de tâche (suite)

Votre agenda – juillet 1999

Juillet 1999

Jeudi	1	Rendez-vous avec David
Vendredi	2	
Samedi	3	
Dimanche	4	
Lundi	5	
Mardi	6	
Mercredi	7	
Jeudi	8	
Vendredi	9	
Samedi	10	Randonnée pédestre à C
Dimanche	11	
Lundi	12	
Mardi	13	
Mercredi	14	
Jeudi	15	
Vendredi	16	
Samedi	17	
Dimanche	18	
Lundi	19	
Mardi	20	
Mercredi	21	
Jeudi	22	
Vendredi	23	
Samedi	24	
Dimanche	25	
Lundi	26	
Mardi	27	
Mercredi	28	Vacances
Jeudi	29	Vacances
Vendredi	30	Vacances
Samedi	31	

Question 1. Lesquelles des dates suivantes conviendraient à la réunion de famille?

Veillez cocher toutes les dates possibles.

- a 4 juillet
- b 7 juillet
- c 14 juillet
- d 18 juillet
- e 25 juillet
- f 29 juillet

Ce projet illustre bien comment la logique des actions requises est effectivement « traduite » en un cheminement concret d'actions thématiques. L'intrigue sous-jacente – la planification d'un voyage et d'une réunion de famille – constitue une action courante très représentative dans laquelle une vaste majorité des habitants de différents pays peuvent vraisemblablement se reconnaître. Les actions requises elles-mêmes et l'ordre dans lequel elles sont accomplies peuvent s'écarter du modèle normatif d'action complète, comme c'est le cas ici. Le modèle normatif sert de balise, qui est adaptée à chaque contexte. En l'occurrence, par exemple, la tâche intitulée « Évaluez les recommandations de vos parents en vue de la randonnée » correspond à peu près à l'étape « Analyser la situation », la tâche intitulée « Planifiez la marche à suivre avant de réserver votre billet d'avion », à l'étape « Planifier la solution » et la tâche « Réservez votre billet d'avion », à l'étape « Exécuter le plan ».

L'exemple de tâche donne une première idée des structures et des formats des items. La tâche s'amorce habituellement par une brève introduction à la situation, suivie de divers types et quantités de renseignements à assimiler. Dans l'exemple, le répondant qui veut fixer la date de la réunion de famille doit traiter, comparer et incorporer l'information fournie dans la liste des engagements de ses parents, y compris les renseignements complémentaires, ainsi que son propre agenda. En l'occurrence, l'information est surtout textuelle et sous forme de tableaux. Il faut donner plus d'une réponse à la question à choix multiple, bien que le nombre de bonnes réponses ne soit pas précisé.

4.6 Conclusion

Le présent chapitre propose un bref aperçu des cadres de travail adoptés pour mettre au point les tâches qui ont servi à mesurer la compréhension de textes suivis et de textes schématiques, la numératie et la résolution de problèmes lors de l'ELCA ainsi qu'à comprendre le sens des résultats à l'égard des compétences en littératie comparatives des adultes. Ces cadres tiennent compte d'une série de variables qui, à la lumière des résultats, sous-tendent l'accomplissement fructueux d'une vaste gamme de tâches. Ensemble, elles fournissent un moyen pour passer d'une interprétation des résultats d'enquête centrée sur des tâches prises isolément ou sur un seul nombre, à une interprétation permettant de définir des niveaux de capacités assez généralisés pour valoir à l'égard de l'ensemble des évaluations et des groupes. Au fur et à mesure qu'on s'intéresse moins aux comportements discrets ou aux observations isolées et davantage à des interprétations significatives du rendement, on obtient un niveau de mesure supérieur (Messick, 1989).

4.6.1 Quelques critères analytiques

Les niveaux de compétence présentés dans l'ensemble de données de l'ELCA permettent non seulement d'étudier la progression des exigences de l'information à traiter sur chacune des échelles, mais aussi d'expliquer comment les compétences démontrées par les répondants reflètent leur probabilité de répondre correctement au vaste éventail de tâches utilisées dans cette évaluation ainsi qu'à d'autres tâches présentant les mêmes caractéristiques. En pratique, cela signifie que les personnes dont le rendement est de 250 sur chaque échelle devraient pouvoir accomplir les tâches moyennes des niveaux 1 et 2 avec un degré de compétence élevé, c'est-à-dire avec une probabilité moyenne d'au moins 80 % de répondre correctement. Cela ne signifie pas qu'elles sont incapables d'effectuer les tâches des niveaux 3 et plus. On s'attendrait à ce qu'elles y arrivent parfois, mais pas toujours.

Où sont passés les analphabètes?

À l'instar de l'enquête antérieure, l'EIACA de 2003 conceptualise les compétences selon un continuum qui indique dans quelle mesure les adultes utilisent l'information afin de fonctionner dans la société et dans l'économie. Rappelons que l'EIACA ne mesure pas l'absence de compétences, mais plutôt les connaissances et les compétences liées aux quatre domaines par rapport à une gamme étendue de capacités. Par conséquent, les résultats ne peuvent servir à classer les groupes de population en « personnes alphabétisées » et en « analphabètes ».

D'après les résultats de l'EIAA de 1994 pour les deux échelles communes à l'ELCA de 2003, les tableaux 4.1 et 4.2 montrent la probabilité que des personnes dont les résultats se situent à certains points des échelles de compréhension de textes suivis et de textes schématiques répondent correctement à des tâches de difficulté variable. Par exemple, un lecteur dont la note en compréhension de textes suivis est de 150 a moins de 50 % des chances de répondre correctement aux tâches de niveau 1. Les personnes dont le rendement est de 200, par contre, ont une probabilité d'environ 80 % de répondre correctement à ces tâches.

Sur le plan des exigences des tâches, on peut supposer que les adultes qui obtiennent une note de 200 sur l'échelle des textes suivis sont probablement capables de repérer un élément d'information dans un texte bref en l'absence d'éléments distracteurs ou si le texte contient des informations plausibles mais incorrectes qui sont éloignées de la bonne réponse. Toutefois, ces personnes risquent d'éprouver beaucoup plus de difficulté avec des tâches des niveaux 2 à 5. Par exemple, elles n'auraient que 40 % de chances d'effectuer correctement une tâche moyenne de niveau 2, 18 % de chances de réussir des tâches de niveau 3 et seulement 7 % de chances avec des tâches des niveaux 4 et 5.

À l'opposé, les répondants qui obtiennent un résultat de 300 sur l'échelle des textes suivis ont à peu près 80 % de chances ou plus de réussir des tâches des niveaux 1, 2 et 3. Cela signifie qu'ils sont capables d'accomplir des tâches qui les amènent à faire des inférences de niveau peu élevé ou à tenir compte de renseignements conditionnels. Ils peuvent également intégrer ou comparer et distinguer des éléments d'information facilement identifiables dans le texte. Par contre, ils risquent d'éprouver de la difficulté avec les tâches qui les obligent à faire des inférences complexes fondées sur le texte ou à traiter des types d'information plus abstraits. Ces tâches plus difficiles peuvent également les obliger à aller au-delà du texte pour faire appel à des connaissances moins familières ou plus spécialisées. En moyenne, ils ont une probabilité d'environ 50 % d'effectuer correctement les tâches de niveau 4; avec les tâches de niveau 5, leurs chances de répondre correctement tombent à 40 %.

On peut interpréter de manière semblable les données concernant les échelles de compréhension de textes schématiques et de textes au contenu quantitatif. Par exemple, une personne qui se situe à 200 sur l'échelle des textes au contenu quantitatif a, en moyenne, 67 % de chances de répondre correctement aux tâches de niveau 1. Sa probabilité de réussite tombe à 47 % pour les tâches de niveau 2, à 21 % pour les tâches de niveau 3, à 6 % pour les tâches de niveau 4 et à seulement 2 % pour les tâches de niveau 5. De même, les lecteurs qui ont une note de 300 sur l'échelle des textes au contenu quantitatif auraient une probabilité de 92 % ou plus de répondre correctement aux tâches des niveaux 1 et 2. Leur probabilité moyenne tomberait à 81 % pour les tâches de niveau 3, à 57 % pour les tâches de niveau 4 et à 20 % pour les tâches de niveau 5.

Tableau 4.1 Probabilités moyennes de réussite, échelle des textes suivis

Textes suivis	Pour certains résultats				
	150	200	250	300	350
			%		
1	48	81	95	99	100
2	14	40	76	94	99
3	6	18	46	78	93
4	2	7	21	50	80
5*	2	6	18	40	68

* En fonction d'une seule tâche

Source : Enquête sur l'alphabétisation des adultes (1994).

Tableau 4.2 Probabilités moyennes de réussite, échelle des textes schématiques

Textes schématiques	Pour certains résultats				
	150	200	250	300	350
			%		
1	40	72	94	99	100
2	20	51	82	95	99
3	7	21	50	80	94
4	4	13	34	64	85
5*	<1	1	3	13	41

* En fonction d'une seule tâche

Source : Enquête sur l'alphabétisation des adultes (1994).

Les compétences dans chaque domaine sont mesurées selon une échelle continue. Chaque échelle de compétence va de zéro jusqu'à un maximum théorique de 500 points (avec quatre décimales de précision). Une personne dont les connaissances la situent à un échelon donné de l'échelle a 80 % de probabilité de réussir une tâche qui comporte ce même niveau de difficulté.

D'un point de vue analytique, des statistiques sommaires utiles peuvent être produites en vue de décrire les compétences des populations, notamment une note moyenne globale. Les populations affichant des notes moyennes similaires peuvent cependant avoir des nombres relativement différents d'adultes à rendement faible ou élevé. Par conséquent, il est également possible d'examiner la façon dont les notes sont distribuées dans les populations en utilisant les

rangs-centiles. Les rangs-centiles sont les niveaux sous lesquels se situe un pourcentage particulier d'adultes. Par exemple, le résultat au cinquième centile est celui sous lequel se situent 5 % des adultes d'une population donnée. Les écarts entre les rangs-centiles nous renseignent sur les différences qui existent dans les populations par rapport aux compétences. Les lecteurs trouveront au chapitre 5 des renseignements plus détaillés sur l'utilisation des valeurs plausibles dans leur analyse.

Les résultats de l'ELCA sont également groupés selon des niveaux de compétence représentant un ensemble de tâches à difficulté croissante. Pour les domaines de la compréhension de textes suivis et de textes schématiques, de même que pour le domaine de la numératie, les spécialistes ont défini cinq grands niveaux de difficulté, chacun correspondant à une fourchette similaire (mais non équidistante) de résultats. En ce qui concerne le domaine de la résolution de problèmes, quatre grands niveaux de difficulté ont été définis. Dans chaque domaine, le niveau 1 indique le niveau le plus faible de compétence, et les niveaux 4 (pour la résolution de problèmes) et 5 (pour les autres domaines), les niveaux les plus élevés.

Il importe, pour des motifs tant analytiques qu'opérationnels, de définir un « niveau souhaité » de compétence pour faire face aux demandes grandissantes de compétences dans une économie axée sur le savoir et l'information. Le niveau 3 de performance est généralement utilisé comme point de repère minimum parce que dans les pays développés, une performance au-delà du niveau 2 est généralement associée à un certain nombre de rendements positifs. Ceux-ci comprennent une participation accrue à la vie sociale, des chances accrues de succès économique et d'indépendance, une amélioration des chances de participer à l'apprentissage continu et d'améliorer son niveau personnel de littératie (Kirsch, I. *et coll.*, 1993; Murray, T.S. *et coll.*, 1997; Tuijnman, A., 2001), alors que les individus qui performent aux niveaux 1 et 2 ne maîtrisent généralement pas encore les compétences minimales de base permettant d'atteindre des niveaux supérieurs de performance (Strucker, J., Yamamoto, K., 2005).

Une analyse secondaire des données de l'EIAA de 1994 a permis de constater que l'écart des rendements entre les niveaux 2 et 3 des échelles de compréhension de textes suivis, de textes schématiques et de textes au contenu quantitatif est considérable et correspond à une différence importante quant au nombre d'avantages mesurables dont bénéficient les citoyens des pays de l'OCDE (OCDE et DRHC, 1997). Les résultats de l'analyse préliminaire des données de l'ELCA, y compris la nouvelle échelle mesurant la numératie, sont conformes à cette constatation. C'est pourquoi il est parfois utile de fixer les échelles à un point de séparation situé entre les niveaux 2 et 3, mettant ainsi l'accent sur la répartition des compétences au-dessus et au-dessous de ce seuil pour les domaines de la compréhension de textes suivis et de textes schématiques ainsi que pour celui de la numératie. À l'opposé, l'interprétation du domaine de la résolution de problèmes est plus complexe, et aucun seuil « souhaitable » n'a encore été établi; il conviendrait donc de situer le point de séparation au niveau 1 en attendant de trouver un seuil plus précis.

Le chapitre 8 présente d'autres outils et d'autres techniques d'utilisation des valeurs plausibles et des poids de rééchantillonnage nécessaires pour produire des estimations exactes des erreurs-types liées à chaque estimation ponctuelle.

Références

- Almond, R.G., and Mislevy, R.J. (1998). Graphical models and computerized adaptive testing. (TOEFL Tech. Rep. No. 14). Princeton, NJ: Educational Testing Service.
- Baker, D., and Street, B. (1994). Literacy and numeracy: Concepts and definitions. In T. Husen and E. A. Postlethwaite (Eds.), *Encyclopedia of education*. New York: Pergamon Press.
- Beazley, K. (1984). *Education in Western Australia: Report of the Committee of Inquiry into Education in Western Australia*. Education Department of Western Australia.
- Coben, D., O'Donoghue, J., and FitzSimons, G.E. (Eds.)(2000). *Perspectives on adults learning mathematics: Theory and practice*. London: Kluwer Academic Publishers.
- Cockcroft, W.H. (1982). *Report of the Committee of Inquiry into the Teaching of Mathematics in Schools*. London: HMSO.
- Cook-Gumperz, J., and Gumperz, J. (1981). From oral to written culture: The transition to literacy. In M. Whitman (Ed.), *Writing: The nature, development and teaching of written communication: Vol. 1*. Hillsdale, NJ: Erlbaum.
- Crandall, J. (1981, December). Functional literacy of clerical workers: Strategies for minimizing literacy demands and maximizing available information. Paper presented at the annual meeting of the American Association for Applied Linguistics, New York.
- Diehl, W. (1980). *Functional literacy as a variable construct: An examination of the attitudes, behaviours, and strategies related to occupational literacy*. Unpublished doctoral dissertation, Indiana University.
- Dossey, J.A. (1997). "Defining and measuring quantitative literacy". In L.A. Steen (Ed.), *Why numbers count: Quantitative literacy for tomorrow's America*. New York: College Entrance Examination Board.
- Fey, James T. (1990). "Quantity" In L.A. Steen (Ed.) *On the shoulders of giants: New approaches to numeracy*. Washington, DC: National Academy Press.
- Frankenstein, M. (1989). *Relearning mathematics: A different third 'R' – Radical maths*. London: Free Association Books.
- Gal, I. (1997). Numeracy: Imperatives of a forgotten goal. In L.A. Steen (Ed.), *Why numbers count: quantitative literacy for tomorrow's America* (pp. 36-44). New York: The College Board.
- Gal, I. (2000). The numeracy challenge. In I. Gal (Ed.), *Adult numeracy development: Theory, research, practice* (pp. 1-25). Cresskill, NJ: Hampton Press.
- Gal, I. (2002). Adult Statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, 70(1), 1-25.
- Jacob, E. (1982). *Literacy on the job: Final report of the ethnographic component of the industrial literacy project*. Washington, DC: Center for Applied Linguistics.
- Johnston, B. (1994, Summer). Critical numeracy? In *Fine print*, Vol. 16, No. 4.
- Heath, S.B. (1980). The functions and uses of literacy. *Journal of Communication*, 30, 123–133.
- Kirsch, I.S., and Guthrie, J.T. (1984a). Adult reading practices for work and leisure. *Adult Education Quarterly*, 34(4), 213–232.
- Kirsch, I.S., and Guthrie, J. T. (1984b). Prose comprehension and text search as a function of reading volume. *Reading Research Quarterly*, 19, 331–342.

- Kirsch, I. (2001). *The International Adult Literacy Survey (IALS): Understanding What Was Measured* (ETS Research Report RR-01-25). Princeton, NJ: Educational Testing Service.
- Marr, B., and Tout, D. (1997). *A numeracy curriculum: Australian Association of Mathematics Teachers (AAMT) conference proceedings*. Melbourne: AAMT.
- Messick, S. (1989). *Validity*. In R. Linn (Ed.), *Educational measurement* (3rd ed.). New York: Macmillan.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Education Researcher*, 32(2), 13-23.
- Mikulecky, L. (1982). Job literacy: The relationship between school preparation and workplace actuality. *Reading Research Quarterly*, 17(3), 400–419.
- Miller, P. (1982). Reading demands in a high-technology industry. *Journal of Reading*, 26(2), 109–115.
- Mislevy, R.J. (September, 2000). Leverage points for improving educational assessment. Paper submitted to National Center for Research on Evaluation, Standards, and Student Testing (CRESST) as part of award #R305B60002 from the US Department of Education, Office of Educational Research and Improvement.
- Mosenthal, P.B., and Kirsch, I.S. (1998). A new measure for assessing document complexity: The PMOSE/IKIRSCH document readability formula. *Journal of Adolescent and Adult Literacy*, 41(8), 638–657.
- Murray, T.S., Clermont, Y. and Binkley, M. (Eds.) *The Adult Literacy and Life Skills Survey: Aspects of Design, Development and Validation*. Canada: Statistics Canada, in press.
- Murray, T.S., Kirsch, I.S., and Jenkins, L. (1998). *Adult Literacy in OECD Countries: Technical report on the First International Adult Literacy Survey*. Washington, DC: National Center for Education Statistics.
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
- Organization for Economic Co-operation and Development. (1992). *Adult illiteracy and economic performance*. Paris, France: Author.
- Rutherford, F.J and Ahlgren, A. (1990). *Science for all Americans*. New York: Oxford University Press.
- Rychen, D.S. and Salganik, L.H. (Eds.) *Key Competencies for a Successful Life and a Well-Functioning Society*. Cambridge, MA: Hogrefe and Huber Publishers, 2003.
- Senechal, Majorie (1990) "Shape" In L.A. Steen (Ed.) *On the shoulders of giants: New approaches to numeracy*. Washington, DC: National Academy Press.
- Scribner, S., and Cole, M. (1981). *The psychology of literacy*. Cambridge, MA: Harvard University Press.
- Steen, L.A. (Ed). (1990). *On the shoulders of giants: New approaches to numeracy*. Washington, DC: National Research Council.
- Steen, L.A. (2001). *Mathematics and democracy: the case for quantitative literacy*. USA: National Council on Education and the Disciplines.
- Sticht, T.G. (Ed.). (1975). *Reading for working: A functional literacy anthology*. Alexandria, VA: Human Resources Research Organization.
- Sticht, T.G. (1978). *Literacy and vocational competency* (Occasional Paper 39, National Center for Research in Vocational Education). Columbus, OH: Ohio State University.

-
-
- Sticht, T.G. (1982, January). Evaluation of the reading potential concept for marginally literate adults. (Final Report FR–ET50–82–2). Alexandria, VA: Human Resources Research Organization.
- Szwed, J. (1981). The ethnography of literacy. In M. Whitman (Ed.), *Writing: The nature, development, and teaching of written communication*: Vol. 1. Hillsdale, NJ: Erlbaum.
- Tobias, S. (1993). *Overcoming math anxiety*. New York: Norton.
- Venezky, R.L. (1983). The origins of the present-day chasm between adult literacy needs and school literacy instruction. *Visible Language*, 16, 113–136.

5.0 Méthodologie de l'enquête

Chaque pays participant devait concevoir et mettre en œuvre l'Enquête sur la littératie et les compétences des adultes (ELCA) selon les normes énoncées dans les lignes directrices du document « Normes et règles d'utilisation pour l'élaboration et la mise en place de l'Enquête sur la littératie et les compétences des adultes ». Ces normes établissaient les exigences minimales concernant le plan d'enquête et la mise en œuvre de l'enquête à l'égard des points suivants :

1. Planification de l'enquête	13. Stratégie relative aux taux de réponse
2. Population cible	14. Embauche, formation et supervision des enquêteurs
3. Méthode de collecte de données	15. Saisie des données
4. Base de sondage	16. Codage
5. Plan d'échantillonnage	17. Notation
6. Prélèvement d'un échantillon	18. Fichier de données de l'ELCA - Format et vérification
7. Plan d'évaluation de la littératie	19. Pondération
8. Questionnaire de référence	20. Estimation
9. Livrets de tâches	21. Confidentialité
10. Exigences relatives aux instruments pour faciliter le traitement des données	22. Documentation relative à l'enquête
11. Collecte de données	23. Enquête pilote
12. Stratégie de prise de contact avec les répondants	

5.1 Plan d'évaluation

Les pays participants, à l'exception de l'État du Nuevo León au Mexique, ont mis en œuvre le plan d'évaluation de l'ELCA. Le Nuevo León a mesuré les compétences en littératie en utilisant les instruments d'évaluation de l'Enquête internationale sur l'alphabétisation des adultes (EIAA).

Dans le cas de l'ELCA et de l'EIAA, un plan d'évaluation de blocs incomplets équilibrés (BIE) a été utilisé pour mesurer les différents domaines de compétences. Le plan BIE comprend un ensemble de tâches d'évaluation organisées en sous-ensembles de tâches ou en blocs. Chacun des blocs contient un ensemble d'items d'évaluation d'un des domaines de compétences et couvre un large éventail de niveaux de difficulté, de facile à difficile. Les blocs d'items sont à leur tour rassemblés en livrets de tâches selon un plan d'évaluation BIE. Les répondants n'étaient pas requis de répondre à l'ensemble des tâches d'un domaine spécifique. À la place, chacun des répondants s'est vu administré aléatoirement un des livrets de tâches.

Évaluation de l'ELCA

L'évaluation psychométrique de l'ELCA comprenait les domaines de compétences suivants : compréhension de textes suivis, compréhension de textes schématiques, numératie et résolution de problèmes. L'outil d'évaluation consistait en quatre blocks d'une durée de 30 minutes (textes suivis et schématiques), deux blocks d'une durée de 30 minutes de numératie et deux blocs d'une durée de 30 minutes de tâches de résolution de problèmes.

Un plan d'évaluation de l'ELCA à quatre domaines a été mis en œuvre aux Bermudes, au Canada, en Hongrie, en Italie, en Norvège, en Nouvelle-Zélande, dans les Pays-Bas et dans les populations francophones et germanophones de la Suisse. Les États-Unis et la population italophone de la Suisse ont mis en œuvre un plan d'évaluation de l'ELCA à trois domaines excluant le domaine de la résolution de problèmes. En plus de l'évaluation de ces domaines de compétences, les pays participants ont aussi évalué l'utilisation des technologies de l'information et des communications en utilisant un ensemble de questions incorporées au questionnaire de référence de l'ELCA.

Les blocs d'items d'évaluation ont été organisés en 28 livrets de tâches pour le plan d'évaluation à quatre domaines et en 18 livrets de tâches pour le plan d'évaluation à trois domaines. Les blocs d'évaluation ont été distribués dans les livrets suivant un plan BIE par lequel chacun des livrets contenait deux blocs d'items. Ces livrets de tâches ont été aléatoirement distribués à un échantillon sélectionné de répondants. De plus, les activités de collecte de données ont été soigneusement orchestrées de manière à obtenir approximativement le même nombre de réponses complètes pour chacun des livrets de tâches, à l'exception de deux livrets de tâches à l'intérieur du plan d'évaluation à trois domaines contenant seulement des tâches de numératie où un nombre plus grand de réponses complètes était requis.

Évaluation de l'EIAA

L'État du Nuevo León, au Mexique, a mis en œuvre le plan d'évaluation de l'EIAA. Le plan d'évaluation de l'EIAA consistait en trois domaines de littératie : textes suivis, textes schématiques et textes au contenu quantitatif. Cet État a toutefois utilisé le questionnaire de référence de l'ELCA. L'utilisation des technologies de l'information et des communications a donc pu être évaluée en utilisant un ensemble de questions incorporées au questionnaire de référence de l'ELCA.

Le plan d'évaluation de l'EIAA consistait en sept livrets de tâches comprenant chacun trois blocs d'items. Les livrets de tâches étaient aléatoirement distribués à un échantillon choisi. De plus les activités de collecte étaient soigneusement orchestrées de manière à obtenir un nombre équivalent de réponses complètes pour chacun des livrets de tâches.

5.2 Population cible et base de sondage

Chaque pays participant a constitué un échantillon censé être représentatif de sa *population civile, hors établissement, âgée de 16 à 65 ans (inclusivement)*.

Les pays étaient également libres d'échantillonner des adultes de plus de 65 ans, à condition de satisfaire à une exigence minimale quant à la taille de l'échantillon de personnes de 16 à 65 ans. Le Canada a choisi d'inclure dans sa population cible des adultes de plus de 65 ans. Tous les autres pays ont restreint la population cible au groupe d'âge de 16 à 65 ans.

En raison de contraintes opérationnelles, certaines exclusions par rapport à la population cible étaient acceptables, à condition que la population observée d'un pays ne diffère pas de plus de 5 % par rapport à la population cible, c'est-à-dire que le nombre total d'exclusions par rapport à la population cible en raison d'un sous-dénombrement ne devait pas excéder 5 % de la population cible. Tous les pays affirment avoir respecté cette exigence.

Chaque pays a choisi ou constitué une base de sondage correspondant à la population cible. Le tableau ci-dessous montre la base de sondage et les exclusions par rapport à la population cible de chaque pays.

TABLEAU 5.1 Base de sondage et exclusions par rapport à la population cible

Pays	Base de sondage	Exclusions
Bermudes	Liste d'évaluation foncière □ liste à jour de tous les logements situés aux Bermudes.	Résidents en établissements, visiteurs aux Bermudes (personnes séjournant moins de six mois).
Canada	Base de données du Recensement de la population et du logement en date du 15 mai 2001 □ ménages dénombrés au moyen du questionnaire complet du recensement (échantillon de 20 %)	Résidents à long terme en établissements, membres des forces armées, personnes vivant dans des réserves indiennes, résidents de régions peu peuplées.
États-Unis	Base aréolaire – 1 883 unités primaires d'échantillonnage couvrant tous les comtés des 50 États des États-Unis plus Washington, D.C.	Personnel militaire à temps plein, résidents en établissements.
Hongrie	Base de données du Recensement de la population et du logement	Sans-abri et prisonniers,
Italie	Liste électorale – liste de personnes de 18 ans et plus qui sont des résidents de l'Italie et qui ont des droits civils	Aucune
Norvège	Registre norvégien de l'éducation (version de 2002)	Résidents permanents en établissements, personnes pour qui le niveau de scolarité atteint est inconnu.
Nouvelle-Zélande	Îlots de base du recensement (définis par Statistics New Zealand pour le recensement de la Nouvelle-Zélande)	Résidents de logements non privés (prisons, maisons de retraite, hôpitaux, résidences universitaires, etc.), résidents des régions rurales éloignées et des îles côtières (à l'exception de l'île Waiheke).
Nuevo León, Mexique	Base de données du Recensement de la population et du logement, année de référence 2000	Résidents permanents en établissements, membres de la marine mexicaine.
Pays-Bas	Données administratives municipales de base (GBA) recueillies chaque mois par le Bureau national de la statistique (CBS).	Résidents en établissements, personnes entrées illégalement au pays.
Suisse	Registre des numéros de téléphone privés (septembre 2002)	Résidents en établissements, personnes vivant dans des régions très isolées, personnes sans numéro de téléphone privé.

5.3 Plan d'échantillonnage

On a demandé à chaque pays participant de constituer un échantillon probabiliste représentatif de la population nationale âgée de 16 à 65 ans. Comme les bases de sondage et les ressources disponibles variaient évidemment d'un pays à l'autre, le plan d'échantillonnage probabiliste à utiliser a été laissé à la discrétion de chaque pays. Statistique Canada a examiné tous les plans d'échantillonnage proposés pour s'assurer que les pays respectaient les normes et les lignes directrices en matière d'échantillonnage.

Le plan d'échantillonnage de chaque pays est résumé ci-dessous. La taille de l'échantillon et le taux de réponse de chaque pays sont abordés dans la section suivante.

Bermudes

On a employé un échantillon probabiliste stratifié à deux degrés. Au premier degré, on a stratifié la liste d'évaluation foncière des logements des Bermudes par paroisse, c'est-à-dire par région géographique. À l'intérieur de chaque paroisse, on a prélevé un échantillon aléatoire de logements avec probabilité proportionnelle au nombre de logements de la paroisse. Au deuxième degré, on a sélectionné un seul répondant admissible à l'aide d'une grille de sélection inspirée de celle de Kish.

Canada

On a employé un plan d'échantillonnage probabiliste à plusieurs degrés pour prélever l'échantillon à partir de la base du recensement. Le plan a servi à constituer des échantillons distincts pour les deux langues officielles du pays, l'anglais et le français. De plus, le Canada a accru la taille de l'échantillon afin de produire des estimations concernant un certain nombre de sous-groupes de population. Des ministères provinciaux et d'autres organismes ont commandité des échantillons supplémentaires pour accroître la base ou pour cibler certaines sous-populations, dont les jeunes (de 16 à 24 ans au Québec, de 16 à 29 en Colombie-Britannique), les adultes de 25 à 64 ans au Québec, les minorités linguistiques (les anglophones au Québec, les francophones ailleurs), les nouveaux immigrants et les immigrants établis, les autochtones urbains et les résidents des territoires du Nord.

Dans chacune des dix provinces canadiennes, on a en outre divisé la base du recensement en deux strates, l'une urbaine, l'autre rurale. On a restreint la strate urbaine aux centres urbains d'une certaine taille, déterminée d'après le recensement antérieur. Le Système automatisé de regroupement des territoires (SARTE) de Statistique Canada a réparti le reste de la base de sondage en unités primaires d'échantillonnage (UPE). On a créé les UPE pour contenir une population suffisante en fonction du nombre de logements à l'intérieur d'un secteur limité et raisonnablement compact. De plus, à l'intérieur de chaque région géographique, on a ordonné la base du recensement selon le niveau de scolarité le plus élevé avant le prélèvement de l'échantillon pour s'assurer que tout l'éventail des niveaux de scolarité soit représenté.

À l'intérieur de la strate urbaine, on a utilisé deux degrés d'échantillonnage. Au premier degré, on a sélectionné systématiquement des ménages avec probabilité proportionnelle à la taille. Au deuxième degré, l'application IPAO a utilisé un algorithme d'échantillon aléatoire simple pour sélectionner une personne parmi les adultes des ménages admissibles. Dans la strate rurale, on a utilisé trois degrés pour prélever l'échantillon. Au premier degré, on a sélectionné des unités primaires d'échantillonnage avec probabilité proportionnelle à la taille de la population. Aux deuxième et troisième degrés, on a repris la même méthode de sélection à deux degrés que dans le cas de la strate urbaine.

Hongrie

On a employé un plan d'échantillonnage stratifié à deux degrés pour constituer un échantillon de personnes sélectionnées proportionnel à la taille de la population (PPT).

On a stratifié la population en sept régions et vingt comtés en tenant compte des caractéristiques démographiques et autres des régions et des comtés (taux de population active et inactive, taux de chômage, etc.), qui variaient d'une région à l'autre. Dans chaque comté, on a en outre stratifié la population en trois types d'agglomération : grande ville, petite ville et village. Puis, on a prélevé l'échantillon à deux degrés :

Degré 1 : un échantillon d'agglomérations avec PPT;

Degré 2 : des adresses choisies au hasard dans les agglomérations sélectionnées au degré 1. Pour chaque agglomération sélectionnée, on a établi la liste d'adresses à partir des fichiers du recensement de 2001 du ministère de l'Intérieur, qui contient les données les plus à jour et les plus précises sur la population de la Hongrie au moment de l'échantillonnage. C'est dans ces fichiers qu'on a sélectionné les personnes à interviewer.

Italie

On a employé un échantillon probabiliste stratifié à trois degrés pour prélever un échantillon à l'aide de listes électorales municipales. On a stratifié le pays en 22 régions géographiques. En général, on a réparti l'échantillon proportionnellement aux 22 régions. Toutefois, on a suréchantillonné les régions du Piémont, de la Vénétie, de la Toscane, de la Campanie et de Trente en vue de produire des estimations distinctes dans ces cinq régions.

Au premier degré, les unités primaires d'échantillonnage étaient les municipalités. À l'intérieur de chaque région géographique, on a stratifié les municipalités, selon la taille de leur population, en unités autoreprésentatives et non autoreprésentatives. On a sélectionné avec certitude les unités autoreprésentatives, c'est-à-dire les grandes municipalités et les municipalités métropolitaines, dans l'échantillon. Dans la strate non autoreprésentative de chaque région, on a sélectionné deux municipalités avec probabilité proportionnelle à la taille de la population cible. Au total, on a sélectionné 256 municipalités à partir des strates autoreprésentatives et non autoreprésentatives.

Au deuxième degré du plan d'échantillonnage, on a défini des « sous-listes par sexe » comme unité secondaire d'échantillonnage. La liste électorale de chaque municipalité sélectionnée comprenait un certain nombre de sous-listes qu'on a stratifiées par sexe. La liste électorale comprenait également l'adresse des résidents italiens âgés de 18 à 65 ans. On a sélectionné systématiquement le même nombre de sous-listes pour chaque sexe. Au total, on a sélectionné 1 326 sous-listes par sexe (663 dans la strate des hommes et 663 dans celle des femmes).

Au troisième degré du plan d'échantillonnage, on a prélevé systématiquement un échantillon de personnes de 18 à 65 ans à partir des unités secondaires d'échantillonnage. Puis, à l'étape de la prise de contact avec les ménages, on a inclus dans l'échantillon toutes les personnes de 16 et 17 ans vivant dans le ménage d'une personne sélectionnée de 18 à 65 ans.

Norvège

À partir de la version de 2002 du Registre norvégien de l'éducation, on a prélevé l'échantillon au moyen d'un plan d'échantillonnage probabiliste à deux degrés.

Le plan a permis de créer 363 unités primaires d'échantillonnage (UPE) à partir des 435 municipalités de la Norvège. On a regroupé ces UPE en 109 strates géographiques. Trente-huit strates étaient constituées d'une seule UPE qui était une municipalité comptant une population d'au moins 25 000 personnes. Au premier degré de prélèvement de l'échantillon, on a inclus avec certitude chacune de ces 38 UPE dans l'échantillon. On a réparti les autres municipalités en 79 strates. Les variables qui ont servi à stratifier ces municipalités étaient la structure industrielle, le nombre d'habitants, la centralité, les structures de communication, les habitudes quotidiennes de déplacement, les secteurs commerciaux et la couverture médiatique (locale). À partir de chacune de ces 79 strates, on a sélectionné une seule UPE avec probabilité proportionnelle à la taille.

Au deuxième degré du plan d'échantillonnage, on a prélevé un échantillon de personnes à partir de chaque UPE échantillonnée. On a stratifié chaque UPE sélectionnée en trois niveaux de scolarité définis par le Registre de l'éducation. On a déterminé la taille de l'échantillon de chaque UPE sélectionnée en répartissant la taille globale de l'échantillon pour chaque UPE sélectionnée avec probabilité proportionnelle à la taille de la population cible. On a ensuite réparti l'échantillon d'UPE comme suit : 30 % provenant du groupe au niveau de scolarité faible, 40 %, du groupe au niveau de scolarité moyen et 30 %, du groupe au niveau de scolarité élevé. On a exclu de l'échantillon les personnes dont le niveau de scolarité n'apparaissait pas sur le Registre de l'éducation (84 318 personnes).

Nouvelle-Zélande

On a employé un plan d'échantillonnage probabiliste stratifié à trois degrés d'échantillonnage : la réplique, le logement et le membre du ménage. On a catégorisé la population en trois strates : la strate principale (toutes les personnes de 16 à 65 ans étaient admissibles), la strate des Maoris et des îliens du Pacifique (seuls ces deux groupes ethniques étaient admissibles) et la strate des îliens du Pacifique (seuls les îliens du Pacifique étaient admissibles).

a) Degré 1 : La réplique

Sur les 38 000 îlots de base visés par le recensement de la population et des logements de la Nouvelle-Zélande en 2001, on a éliminé ceux qui comptaient neuf logements ou moins; il est resté 32 115 îlots de base comptant dix logements et plus. La couverture des logements privés permanents était de 98,6 %. Dans chaque îlot de base, la probabilité de sélection était proportionnelle au nombre de logements compris. On a sélectionné au total 896 îlots de base, qu'on a ensuite répartis en 32 répliques composées de 28 îlots de base par réplique. Chaque réplique comprenait des îlots de base répartis à peu près de la même manière du nord au sud et constituait donc un mini-échantillon probabiliste national.

b) Degré 2 : Le logement

Pour la strate principale, on a sélectionné les logements comme suit : on a calculé l'intervalle d'échantillonnage de chaque îlot de base en divisant par 15 le nombre de logements compris dans l'îlot de base. L'intervalle d'échantillonnage différait donc selon

la taille de l'îlot de base. À partir d'un point de départ aléatoire, les enquêteurs ont sélectionné les logements en fonction de l'intervalle d'échantillonnage de l'îlot de base.

En plus des logements de la strate principale, on a échantillonné jusqu'à 21 logements supplémentaires par îlot de base pour la strate des Maoris et des îliens du Pacifique et pour celle des îliens du Pacifique. Dans quatre de ces logements, les résidents des deux groupes ethniques étaient admissibles à la sélection. Dans les 17 logements restants, seuls les résidents originaires des îles du Pacifique étaient admissibles. Abstraction faite des logements de la strate principale, l'intervalle d'échantillonnage de ces logements était de 1.

c) Degré 3 : Le répondant

Pour la strate principale, on a sélectionné à l'aide de la grille de Kish une seule personne par ménage parmi tous les membres admissibles du ménage. Pour les deux strates ethniques, on a en outre tenu compte de l'ethnicité des membres du ménage (Maoris ou îliens du Pacifique pour la strate 2, îliens du Pacifique pour la strate 3) avant de sélectionner un répondant à l'aide de la grille de Kish.

Nuevo León, Mexique

On a utilisé un plan d'échantillonnage probabiliste stratifié à deux degrés à l'intérieur de chacune des strates.

On a regroupé géographiquement les 51 municipalités du Nuevo León en trois différentes strates : Strate 1 – Région métropolitaine de recensement de Monterrey, comprenant neuf municipalités; Strate 2 – les municipalités de Linares et de Sabinas Hidalgo; Strate 3 – le restant des 40 municipalités du Nuevo León. L'échantillon initial a été réparti aux trois strates proportionnellement au nombre de résidences à l'intérieur de chaque strate.

Au premier degré de prélèvement de l'échantillon, à l'intérieur de chacune des strates, on a prélevé un nombre aléatoire de ménages. Au deuxième degré, on a sélectionné dans chaque ménage une seule personne appartenant à la population cible à l'aide d'une grille de sélection inspirée de celle de Kish.

Pays-Bas

Le plan d'échantillonnage des Pays-Bas était un plan systématique stratifié en grappes à plusieurs degrés.

Au premier degré, on a stratifié le pays en quatre régions : le Nord, l'Est, l'Ouest et le Sud. À l'intérieur de chaque strate, on a sélectionné un échantillon de municipalités avec probabilité proportionnelle à la taille de leur population. Pour ce faire, on a classé les municipalités de chaque strate selon la taille de la population et on a prélevé systématiquement l'échantillon de municipalités en utilisant un point de départ aléatoire et un intervalle d'échantillonnage fixe. Les données sur la population étaient fondées sur les données administratives municipales (Gemeentelijke Basis Administratie ou GBA) recueillies par le Bureau national de la statistique (Centraal Bureau voor de Statistiek ou CBS).

Au deuxième degré, à l'intérieur de chaque municipalité sélectionnée, on a prélevé un échantillon systématique de régions définies selon le code postal. La société Experian a fourni des données au sujet des cotes de crédit (pourcentage de ménages ayant des dettes à

l'intérieur d'une région définie selon le code postal) et du pouvoir d'achat dans les régions définies selon le code postal à six caractères. On a ordonné ces régions selon la cote de crédit, puis selon le pouvoir d'achat. À partir d'un point de départ aléatoire et avec un intervalle d'échantillonnage fixe (pour ce qui est des ménages), on a ensuite sélectionné les ménages.

Au troisième degré, à l'intérieur de chaque région sélectionnée définie selon le code postal, on a choisi au hasard un seul ménage. Les données provenaient de la base de données d'Experian sur les ménages (basée sur Cendris, propriétaire actuel de la base de données centrale du bureau de poste). Cette base de données est mise à jour chaque mois.

Au quatrième degré, on a choisi au hasard une seule personne admissible à l'intérieur du ménage sélectionné.

Suisse

On a utilisé un plan d'échantillonnage probabiliste stratifié à deux degrés. On avait besoin d'estimations distinctes pour les trois régions linguistiques de la Suisse (allemande, française, italienne). Les trois régions linguistiques constituent donc les strates primaires. À l'intérieur des régions linguistiques, on a en outre stratifié la population en régions métropolitaines représentées par les cantons de Genève et de Zurich et le reste des régions linguistiques. Au premier degré d'échantillonnage, on a tiré de chaque strate un échantillon systématique de ménages à partir d'une liste de numéros de téléphone privés. Au deuxième degré, on a sélectionné dans chaque ménage une seule personne appartenant à la population cible à l'aide d'une grille de sélection inspirée de celle de Kish.

États-Unis

Aux États-Unis, on a employé un plan d'échantillonnage probabiliste stratifié à plusieurs degrés.

Au premier degré d'échantillonnage, on a prélevé un échantillon de 60 unités primaires d'échantillonnage (UPE) à partir d'un total de 1 883 UPE constituées d'un seul comté ou d'un groupe de comtés contigus, selon la taille de la population et la superficie du ou des comtés. On a stratifié les UPE en fonction des caractéristiques économiques et sociales de la population déclarées lors du recensement de 2000, soit les caractéristiques suivantes : la région du pays, le statut de région statistique métropolitaine (Metropolitan Statistical Area ou MSA), la taille de la population, les pourcentages de résidents afro-américains et de résidents hispaniques, ainsi que le revenu par habitant. On a inclus dans l'échantillon, avec certitude, les plus grandes UPE en fonction d'un seuil de taille de la population. Quant aux autres UPE, on a sélectionné une seule UPE par strate avec probabilité proportionnelle à la taille de la population.

Au deuxième degré d'échantillonnage, on a sélectionné systématiquement, à partir des UPE échantillonnées, un total de 505 segments géographiques avec probabilité proportionnelle à la taille de la population. Les segments consistent en îlots aréolaires (définis par le recensement de 2000) ou en combinaisons d'au moins deux îlots voisins. On les a constitués pour respecter des critères fondés sur la taille de la population et la proximité géographique.

Au troisième degré d'échantillonnage, on a dressé la liste des logements situés dans les segments sélectionnés, puis on a prélevé un échantillon aléatoire de logements. On a sélectionné un nombre égal de logements à partir de chaque segment échantillonné.

Au quatrième et dernier degré d'échantillonnage, on a sélectionné au hasard une seule personne admissible dans les ménages comptant moins de quatre adultes admissibles. Dans

les ménages comptant quatre personnes admissibles et plus, on a sélectionné deux adultes au hasard.

5.4 Taille de l'échantillon

On a recommandé que l'échantillon compte 5 400 unités complètes dans chaque langue officielle pour chaque pays qui mettait en œuvre l'évaluation psychométrique complète de l'ELCA (c'est-à-dire comprenant les domaines suivants : compréhension de textes suivis, de textes schématiques, numératie et résolution de problèmes). On a recommandé que l'échantillon compte 3 420 unités complètes dans chaque langue officielle si l'évaluation ne comprenait pas le domaine de la résolution de problèmes.

On a recommandé que l'échantillon compte 3 000 unités complètes dans le cas du Nuevo León au Mexique qui a évalué les compétences en littératie en utilisant les livrets de tâches psychométriques de l'Enquête internationale sur l'alphabétisation des adultes (EIAA).

Le tableau 5.2 montre le nombre final de répondants (unités complètes pour la ou les langues d'évaluation de chaque pays participant).

TABEAU 5.2 Taille de l'échantillon selon la langue d'évaluation

Pays	Langue d'évaluation	Domaines d'évaluation ¹	Nombre de répondants ²
Bermudes	Anglais	S, Sc, N, RP	2 696
Canada	Anglais	S, Sc, N, RP	15 694
	Français	S, Sc, N, RP	4 365
États-Unis	Anglais	S, Sc, N	3 420
Hongrie	Hongrois	S, Sc, N, RP	5 635
Italie	Italien	S, Sc, N, RP	6 853
Norvège	Bokmal	S, Sc, N, RP	5 411
Nouvelle-Zélande	Anglais	S, Sc, N, RP	7 131
Nuevo León, Mexique	Espagnol	S, Sc, Q	4 786
Pays-Bas	Néerlandais	S, Sc, N, RP	5 617
Suisse	Français	S, Sc, N, RP	1 765
	Allemand	S, Sc, N, RP	1 892
	Italien	S, Sc, N	1 463

1. S – Textes suivis, Sc – Textes schématiques, N – Numératie, RP – Résolution de problèmes, Q – Textes au contenu quantitatif.

2. Pour les besoins de l'échelonnement des données d'évaluation psychométrique d'un pays, les données d'un répondant sont considérées comme complètes à condition de comprendre au moins les variables du questionnaire de référence en ce qui concerne l'âge, le sexe et le niveau de scolarité.

5.5 Collecte de données

Le plan d'enquête de l'ELCA combinait les techniques de l'évaluation pédagogique à celles d'une enquête-ménage afin d'évaluer la littératie et d'obtenir les renseignements nécessaires

pour rendre ces mesures significatives. On a d'abord posé aux répondants une série de questions pour obtenir des renseignements contextuels et démographiques sur leur scolarité, leurs activités de lecture à la maison et au travail, leur situation par rapport au marché du travail, leur utilisation des technologies de l'information et de la communication, leur participation à la formation continue et une auto-évaluation de leur niveau de littératie.

Une fois le questionnaire de référence rempli, l'enquêteur a remis au répondant un livret comprenant six tâches de lecture simples (tâches préliminaires). Le répondant qui réussissait à ce premier test recevait un autre livret renfermant un nombre beaucoup plus important de tâches, tirées d'une banque d'items groupés par blocs. Chaque livret renfermait deux blocs qui représentaient environ 45 items. L'évaluation n'était pas minutée et on a exhorté le répondant à essayer chaque exercice. On lui a donc accordé le maximum de chances pour faire preuve de ses compétences, même si ces dernières s'avéraient faibles.

La collecte de données dans le cadre de l'ELCA a eu lieu entre l'automne de 2003 et le début du printemps de 2008, selon le pays. Le tableau 5.3 présente les périodes de collecte de données de chaque pays participant.

TABEAU 5.3 Période de collecte de données

Pays	Période de collecte
Bermudes	mars à août 2003
Canada	mars à septembre 2003
États-Unis	janvier à juin 2003
Hongrie	juillet 2007 à février 2008
Italie	mai 2003 à janvier 2004
Norvège	janvier à novembre 2003
Nouvelle-Zélande	août 2005 à avril 2007
Nuevo León, Mexique	octobre 2002 à mars 2003
Pays-Bas	juillet 2007 à janvier 2008
Suisse	janvier à novembre 2003

Pour assurer un haut degré de qualité des données, les lignes directrices relatives à l'exécution de l'ELCA stipulaient que chaque pays devait travailler avec un organisme ou une entreprise de collecte de données réputée posséder, de préférence, ses propres enquêteurs professionnels et expérimentés. La méthode de rémunération de ces enquêteurs devait favoriser une réponse maximale. Les entrevues ont été menées à domicile, de manière impartiale et sans aucune pression. Les pays devaient assurer la formation et la supervision des enquêteurs en mettant l'accent sur le choix d'une seule personne par ménage (dans la mesure du possible), le choix de l'un des 28 livrets de tâches principales (dans la mesure du possible), la notation du livret de tâches préliminaires et l'attribution des codes d'état. Enfin, on devait superviser le travail des enquêteurs en exerçant de fréquents contrôles de la qualité au début de la collecte de données,

puis moins de contrôles par la suite, et en offrant aux enquêteurs de l'aide durant la période de collecte de données.

Les responsables de l'ELCA ont pris quelques précautions contre les erreurs dues aux non-réponses, comme le stipulaient les lignes directrices relatives à l'exécution de l'enquête. Les enquêteurs avaient pour instruction formelle de retourner à plusieurs reprises chez les ménages non répondants afin d'obtenir le plus de réponses possible. De plus, on a demandé à tous les pays de s'assurer que les adresses fournies aux enquêteurs étaient aussi complètes que possible, afin de réduire les problèmes éventuels d'identification des ménages.

On a demandé aux pays de remplir, après la partie principale de l'enquête, un questionnaire d'évaluation afin de montrer qu'on avait bien suivi les lignes directrices, ainsi que pour cerner les problèmes de collecte qui avaient pu survenir. Le tableau 5.4 présente des renseignements sur les entrevues tirés de ce questionnaire.

TABLEAU 5.4 Renseignements sur les enquêteurs

Pays	Nombre de langues	Nombre de d'enquêteurs	Taille moyenne de l'affectation	Expérience des enquêteurs
Bermudes	1	105	40	Indéterminée
Canada	2	317	62	Enquêteurs professionnels possédant au moins deux ans d'expérience.
États-Unis	1	106	64	Enquêteurs professionnels dont environ le quart ne possédaient aucune expérience antérieure en matière d'enquête.
Hongrie	1	175	32	Enquêteurs professionnels possédant au moins deux ans d'expérience.
Italie	1	150	45	Enquêteurs professionnels dont la plupart possédaient au moins deux ans d'expérience.
Norvège	1	320	30	Seulement le tiers des enquêteurs possédaient au moins deux ans d'expérience; les autres ont été formés spécialement pour cette enquête.
Nouvelle-Zélande	1	160	45	Enquêteurs professionnels au nombre d'années d'expérience inconnu.
Nuevo León, Mexique	1	209	29	Environ 70 % des enquêteurs possédaient deux ans d'expérience.
Pays-Bas	1	277	35	Enquêteurs professionnels dont environ un sur cinq sans expérience antérieure des enquêtes.
Suisse	3	110	60	Indéterminée

5.6 Traitement de données

Comme condition de leur participation à l'ELCA, on a demandé aux pays de saisir et de traiter leurs fichiers en utilisant des méthodes qui assuraient une cohérence logique et des niveaux acceptables d'erreur dans la saisie des données. Plus précisément, on leur a demandé d'effectuer une vérification complète des notes saisies (en entrant chaque enregistrement deux fois) afin de réduire les taux d'erreur au minimum. L'exactitude de la notation des tests étant essentielle pour assurer un haut degré de qualité des données, il fallait exercer un contrôle intégral par double saisie.

Chaque pays devait également coder l'industrie, la profession et le niveau de scolarité à l'aide des codes types suivants : la Classification internationale type, par industrie (CITI), la Classification internationale type des professions (CITP) et la Classification internationale type de l'éducation (CITE). Statistique Canada a fourni des codes pour toutes les questions ouvertes, et les pays ont reçu des directives précises relativement au codage de ces questions.

Pour améliorer la comparabilité de l'analyse des données, on a demandé à chaque pays participant à l'ELCA de schématiser son ensemble de données national sous forme de disposition d'enregistrement très structurée et uniformisée. Outre qu'elle précisait la position, le format et la longueur de chaque champ, la disposition d'enregistrement internationale comprenait une description de chaque variable et indiquait les catégories et les codes à fournir pour cette variable. Dès la réception du fichier d'un pays, Statistique Canada a effectué une série de contrôles d'intervalle pour garantir la conformité avec le format prescrit, ainsi que des contrôles de cohérence et de flux de données des fichiers reçus. En cas d'anomalies, on signalait les problèmes aux pays et on leur demandait de présenter des fichiers épurés.

5.7 Notation des tests

Dans chaque pays, les personnes chargées de la notation ont reçu une formation intensive relativement à la notation des réponses aux questions ouvertes à l'aide du manuel de notation de l'ELCA. On leur a également fourni un outil pour saisir les questions fermées. Pour favoriser l'exactitude de la notation et la comparabilité entre les pays, les responsables de l'ELCA ont installé un panneau d'affichage électronique permettant aux pays de poser leurs questions en matière de notation aux spécialistes des domaines, et à ces derniers, de communiquer leurs décisions. Tous les pays avaient accès à ces renseignements et pouvaient rajuster leur notation en conséquence.

Pour assurer davantage l'exactitude, on a contrôlé de deux façons la qualité de la notation effectuée par les pays.

Premièrement, à l'intérieur d'un pays, au moins 20 % des tests ont dû faire l'objet d'une deuxième notation. Les lignes directrices relatives à la deuxième notation à l'intérieur d'un pays stipulaient une deuxième notation de toutes les grandes parties des livrets au début du processus afin de cerner et de corriger le plus grand nombre possible de problèmes de notation. Dans un deuxième temps, les correcteurs devaient sélectionner une plus petite partie d'un autre tiers des livrets de notation. La dernière étape, considérée comme une mesure de contrôle de la qualité, consistait à noter à nouveau et régulièrement une petite partie des livrets jusqu'à la fin du processus de deuxième notation. Les deux ensembles de notes devaient concorder dans une proportion d'au moins 95 % pour qu'on puisse passer à l'étape suivante du traitement. En fait, dans la plupart des cas, la fiabilité de la notation à l'intérieur d'un pays était supérieure à 95 %. Lorsqu'il y avait des erreurs, on a demandé au pays de reprendre les livrets

et de noter à nouveau toutes les questions présentant des problèmes et tous les tests notés par un correcteur problématique.

Deuxièmement, on a effectué une deuxième notation internationale. Dans une proportion de 10 %, l'échantillon de chaque pays a fait l'objet d'une deuxième notation par les correcteurs d'un autre pays. Par exemple, un échantillon des livrets de tâches des États-Unis a été soumis à une deuxième notation par les personnes qui avaient noté les livrets en anglais du Canada, et inversement. La deuxième notation visait principalement à vérifier qu'aucun pays n'avait effectué une notation uniformément différente de celle des autres pays. Statistique Canada a calculé la fiabilité de la notation entre les pays et le Educational Testing Service, de Princeton, a évalué les résultats. Encore une fois, on a exigé une exactitude rigoureuse, soit une concordance de 90 %, pour que les notes soient jugées acceptables. Chaque fois qu'on a décelé un problème, l'échantillon a été soumis à une deuxième notation. Le tableau 5.5 montre le haut niveau de concordance des notes d'un pays à l'autre.

TABLEAU 5.5 Notation – Fiabilité en pourcentage par domaine

Appariement des pays (pays chargé de la deuxième notation – pays d'origine)	Domaines psychométriques			Total (%)
	Textes suivis et textes schématiques (%)	Numératie (%)	Résolution de problèmes (%)	
Canada anglophone – Canada francophone	95	95	92	95
Canada francophone – Canada anglophone	95	97	94	95
Norvège – Canada	91	93	91	92
Canada – États-Unis	94	97	...	95
États-Unis – Canada	95	97	...	95
États-Unis – Bermudes	91	94	...	90
Bermudes – États-Unis	93	95	...	93
Canada francophone – Suisse	95	98	97	96
Suisse – Canada francophone	94	96	94	95
Suisse – Italie	96	98	96	96
Italie – Suisse	93	97	93	94
Canada – Bermudes	83	83
Canada – Nuevo León	91	95 ¹	...	92
Hongrie	94	96	93	94
Pays-Bas	91	93	93	92
Nouvelle-Zélande	96	97	94	96

... N'ayant pas lieu de figurer

1. Textes au contenu quantitatif.

TABLEAU 5.6 Résumé des opérations de notation

Pays	Début de la notation ¹	Nombre de correcteurs	Temps moyen de notation par livret
Bermudes	Au milieu	5	20 min.
Canada	Au milieu	18 ²	13 min.
États-Unis	Au début	7	12 min.
Hongrie	Au milieu	9	20 min.
Italie	Au début	9	15 min.
Norvège	Au milieu	17	8 min.
Nouvelle-Zélande	Au début	12	20 min.
Nuevo León, Mexique	Au milieu	12	n.d.
Pays-Bas	Au milieu	7	12 min.
Suisse	Au début	11	22 min.

1. Indique que la notation a commencé au début, au milieu ou à la fin de la collecte de données.

2. Dont 15 correcteurs, deux personnes pour saisir les questions fermées en résolution de problèmes et une personne pour saisir les feuilles de notation.

5.8 Réponse à l'enquête et pondération

Chaque pays participant à l'ELCA a utilisé un plan d'échantillonnage probabiliste à plusieurs degrés avec stratification et probabilités inégales de sélection des répondants. En outre, il est nécessaire de compenser les cas de non-réponse survenus à divers niveaux. L'estimation des paramètres de population et des erreurs-types correspondantes dépend donc des coefficients de pondération, ou poids.

Tous les pays participants ont employé la même méthode générale pour calculer les coefficients de pondération, mais chaque pays a élaboré ces coefficients en fonction de son propre plan d'échantillonnage probabiliste.

En général, chaque pays a calculé deux types de coefficients de pondération : des poids de population, qui servent à produire des estimations de population, et des poids de rééchantillonnage jackknife, qui servent à calculer les erreurs-types correspondantes.

Poids de population

Pour chaque enregistrement de répondant, on a établi le poids de population en calculant d'abord le poids d'échantillonnage théorique. Puis, on a calculé un poids d'échantillonnage de base en rajustant mathématiquement le poids théorique en fonction de la non-réponse. Le poids de base est le poids fondamental qui peut servir à produire des estimations de population. Toutefois, afin d'assurer la cohérence entre les poids d'échantillonnage et les chiffres de population connus d'un pays (c'est-à-dire les totaux de référence) pour les caractéristiques essentielles, on a ajusté les poids d'échantillonnage de base en fonction des totaux de référence.

Le tableau 5.7 présente les variables de référence de chaque pays et la source des chiffres de population de référence.

Poids jackknife

On a recommandé d'établir de dix à trente poids de rééchantillonnage jackknife afin de déterminer les erreurs-types des estimations d'enquête.

La Suisse a produit quinze poids de rééchantillonnage jackknife. Les autres pays en ont produit trente.

TABLEAU 5.7 **Variables de référence par pays**

Pays	Source des chiffres de population de référence	Variables de référence
Bermudes	Recensement de 2000	Âge, sexe, niveau de scolarité
Canada	Chiffres de population du recensement, juin 2003	Province, aire géographique de recensement (RMR/AR), âge, sexe
États-Unis	Current Population Survey (enquête sur l'état de la population), 2003 supplément de mars	Région de recensement, statut de région statistique métropolitaine (MSA), âge, sexe, race/ethnicité, statut d'immigrant
Hongrie	Données démographiques de 2005-2006 du Bureau central de la statistique (KSH) de la Hongrie	Âge, sexe, niveau de scolarité, région géographique
Italie	Enquête polyvalente ISTAT de 2002	Région, âge, sexe, niveau de scolarité, situation face à l'emploi
Norvège	Registre norvégien de l'éducation (version de 2002)	Âge, sexe, niveau de scolarité
Nouvelle-Zélande	Census of Populations and Dwellings (2006)	Âge, sexe, niveau de scolarité
Nuevo León, Mexique	Recensement de la population et des logements (2000)	Âge, sexe, niveau de scolarité
Pays-Bas	Données administratives municipales de base (GBA) recueillies par le Bureau national de la statistique (CBS) et base de données d'Experian.	Âge, niveau de scolarité, pouvoir d'achat mode d'occupation du logement.
Suisse	Enquête suisse sur la population active (ESPA)	Région linguistique, âge, sexe, niveau de scolarité, statut d'immigrant

Le tableau 5,8 résume la taille des échantillons et les taux de réponse de chaque pays participant.

TABLEAU 5.8 Taille de l'échantillon et taux de réponse – Résumé

Pays	Population âgée de 16 à 65 ans	Taille de l'échantillon initial (personnes de 16 à 65 ans)	Unités hors champ ¹	Nombre de répondants ² (personnes de 16 à 65 ans)	Taux de réponse ³ (personnes de 16 à 65 ans)
					%
Bermudes	43 274	4 049	745	2 696	82
Canada	21 960 683	35 270	4 721	20 059	66
États-Unis	184 260 910	7 045	1 846	3 420	66
Hongrie	6 760 050	9 178	18 356	5 635	63
Italie	38 765 513	16 727	971	6 853	44
Norvège	2 945 838	9 719	16	5 411	56
Nouvelle-Zélande	2 634 442	28 702	17 565 ⁴	7 131	64
Nuevo León, Mexique	2 382 454	6 000	36	4 786	80
Pays-Bas	10 974 940	12 734	719	5 617	47
Suisse	1 161 735	18 282	5 310	5 120	40

1. Les unités hors du champ de l'enquête sont celles qui ont été codées comme suit : résidents non admissibles, logement introuvable, logement en construction, logement inoccupé ou saisonnier, ou unités doubles.

2. Pour les besoins de l'échelonnement des données d'évaluation psychométrique d'un pays, les données d'un répondant sont considérées comme complètes à condition de comprendre au moins les variables du questionnaire de référence en ce qui concerne l'âge, le sexe et le niveau de scolarité.

3. On calcule le taux de réponse en divisant le nombre de répondants par la taille de l'échantillon initial, déduction faite des unités hors champ. 4. Si le nombre d'unités hors champ en Nouvelle-Zélande est relativement élevé, c'est qu'on a employé une méthode de présélection pour « suréchantillonner » la population maorie et celle des îliens du Pacifique. Dans les parties présélectionnées de l'échantillon, seuls les Maoris et les îliens du Pacifique étaient considérés comme admissibles.

4. Si le nombre d'unités hors champ en Nouvelle-Zélande est relativement élevé, c'est qu'on a employé une méthode de présélection pour « suréchantillonner » la population maorie et celle des îliens du Pacifique. Dans les parties présélectionnées de l'échantillon, seuls les Maoris et les îliens du Pacifique étaient considérés comme admissibles.

6.0 Méthodes d'enquête et traitement des données

6.1 Introduction

Les méthodes de l'ELCA sont inspirées des lignes directrices internationales concernant l'administration de l'ELCA. Le caractère normalisé des instruments, de l'échantillonnage, de la collecte et des méthodes de traitement (y compris les codes uniformisés pour la profession, l'industrie et l'éducation) contribue énormément à faire de l'ELCA une étude comparative internationale. Dans la présente section, nous décrivons ces procédés ainsi que les écarts du volet canadien de l'enquête par rapport au protocole. Nous abordons également certains aspects du traitement postcollecte des données qui a mené à la création du fichier de microdonnées à grande diffusion.

L'ELCA a permis de recueillir des renseignements descriptifs et des données sur les capacités des répondants échantillonnés au moyen d'un questionnaire de base et d'une série de livrets d'évaluation contenant des tâches de compréhension de textes suivis et de textes schématiques, de numératie et de résolution de problèmes. Les répondants ont passé environ 30 minutes à répondre à un ensemble commun de questions de base concernant leurs caractéristiques démographiques, leur scolarité, leur expérience sur le marché du travail et leurs activités en matière de littératie. Les réponses à ces questions de base ont permis de résumer les résultats de l'enquête en utilisant un éventail de variables descriptives et d'accroître l'exactitude des estimations des capacités de diverses sous-populations. Les renseignements de base ont été recueillis par des intervieweurs dûment formés.

Après avoir répondu aux questions de base, les répondants ont passé le reste du temps à remplir un livret de tâches de littératie visant à mesurer leurs capacités en compréhension de textes suivis et de textes schématiques, en numératie et en résolution de problèmes. La plupart de ces tâches comportaient des questions ouvertes, auxquelles les répondants devaient fournir une réponse écrite.

Pour obtenir une couverture suffisante de chacun des quatre aspects de la littératie, il fallait évaluer un nombre assez important de tâches. Par contre, il fallait aussi maintenir le fardeau de chaque répondant dans des limites de temps acceptables. Pour respecter ces deux exigences contradictoires, c'est-à-dire réduire le fardeau des répondants sans sacrifier une représentation suffisante des quatre aspects de la littératie, on a demandé à chaque répondant de n'exécuter qu'une partie de l'ensemble de tâches, en utilisant une variante de l'échantillonnage matriciel.

6.2 Modèles de méthodes, de manuels et d'instruments

Chaque pays participant à l'ELCA a reçu un ensemble de manuels d'administration et d'instruments d'enquête à utiliser comme modèle. Les pays étaient libres d'adapter ces

modèles à leur système national de collecte de données, mais en conservant un certain nombre de caractéristiques essentielles. Premièrement, les répondants devaient remplir les livrets de tâches de base et de tâches principales seuls, chez eux, sans l'aide d'une autre personne ou d'une calculatrice. Deuxièmement, les répondants ne devaient pas être rémunérés pour leur participation. Troisièmement, malgré l'interdiction d'offrir des incitatifs pécuniaires, les intervieweurs disposaient de méthodes visant à maximiser le nombre de questionnaires de base remplis et devaient utiliser les mêmes normes de codage pour traiter les cas de non-réponse. Cette dernière exigence est essentielle. Le fait de ne pas remplir les livrets de tâches de base et de tâche principales étant corrélé avec la compétence, on a besoin de renseignements de base sur les non-répondants afin d'imputer les données cognitives concernant ces personnes.

6.2.1 Questions de base

Le modèle de questionnaire de base remis à tous les pays participants comportait deux ensembles de questions : les questions obligatoires, que tous les pays devaient poser, et les questions facultatives, qui étaient recommandées mais non exigées. Les pays n'étaient pas tenus de traduire littéralement les questions obligatoires, mais on leur a demandé de respecter l'objectif conceptuel de chaque question en l'adaptant à leur enquête. Les pays étaient libres d'ajouter des questions à leur questionnaire de base si le fardeau supplémentaire des répondants ne risquait pas de réduire les taux de réponse.

Lorsque les réponses à ces questions ne compromettent pas la confidentialité des répondants, le FMGD de l'ELCA comprend le plus grand nombre possible de données recueillies. Dans le chapitre 8, nous aborderons en détail la question de la confidentialité.

6.2.2 Items de tâche

À l'instar de l'ELCA qui l'a précédée, l'IALS repose sur le principe selon lequel la difficulté de diverses tâches de littératie est déterminée par certains facteurs, qui sont stables d'une langue à l'autre et d'une culture à l'autre. On a donc remis à tous les pays participants des fichiers graphiques contenant l'ensemble des items psychométriques et on les a invités à traduire chaque item de l'anglais à leur propre langue sans modifier la présentation graphique ni les caractéristiques des tâches. Dans bien des cas, on a traduit l'item initial en fonction de ce modèle anglais pour offrir à tous le même point de départ. Cette uniformité des documents de base a permis de réduire au minimum l'incidence d'erreurs de traduction et d'adaptation.

Le processus d'adaptation des items était soumis à certaines règles. Par exemple, certains items exigeaient du répondant qu'il accomplisse une tâche qui était facilitée par l'utilisation de mots-clés. Dans certains cas, les mots-clés étaient identiques dans la question et dans le corps de l'item; dans d'autres, le mot-clé était semblable, mais pas exactement le même; dans d'autres encore, le mot-clé était un synonyme du mot utilisé dans le corps de l'item. Dans un autre cas, on demandait au répondant de choisir, dans le corps de l'item, entre plusieurs mots-clés dont un seul était correct. Les pays devaient traduire les items en préservant ces associations conceptuelles.

Au besoin, certaines conventions utilisées dans les items – unités monétaires, écriture des dates, notation décimale, etc. – ont été adaptées pour chaque pays.

Pour s'assurer que le processus d'adaptation ne compromettait pas l'intégrité psychométrique des items, on a soigneusement examiné les livrets de chaque pays pour y déceler des erreurs d'adaptation.

6.2.3 Codage normalisé des non-réponses

Il était essentiel que les pays participant à l'enquête gèrent les cas de non-réponse de manière uniforme afin de limiter, dans les estimations, le biais de non-réponse.

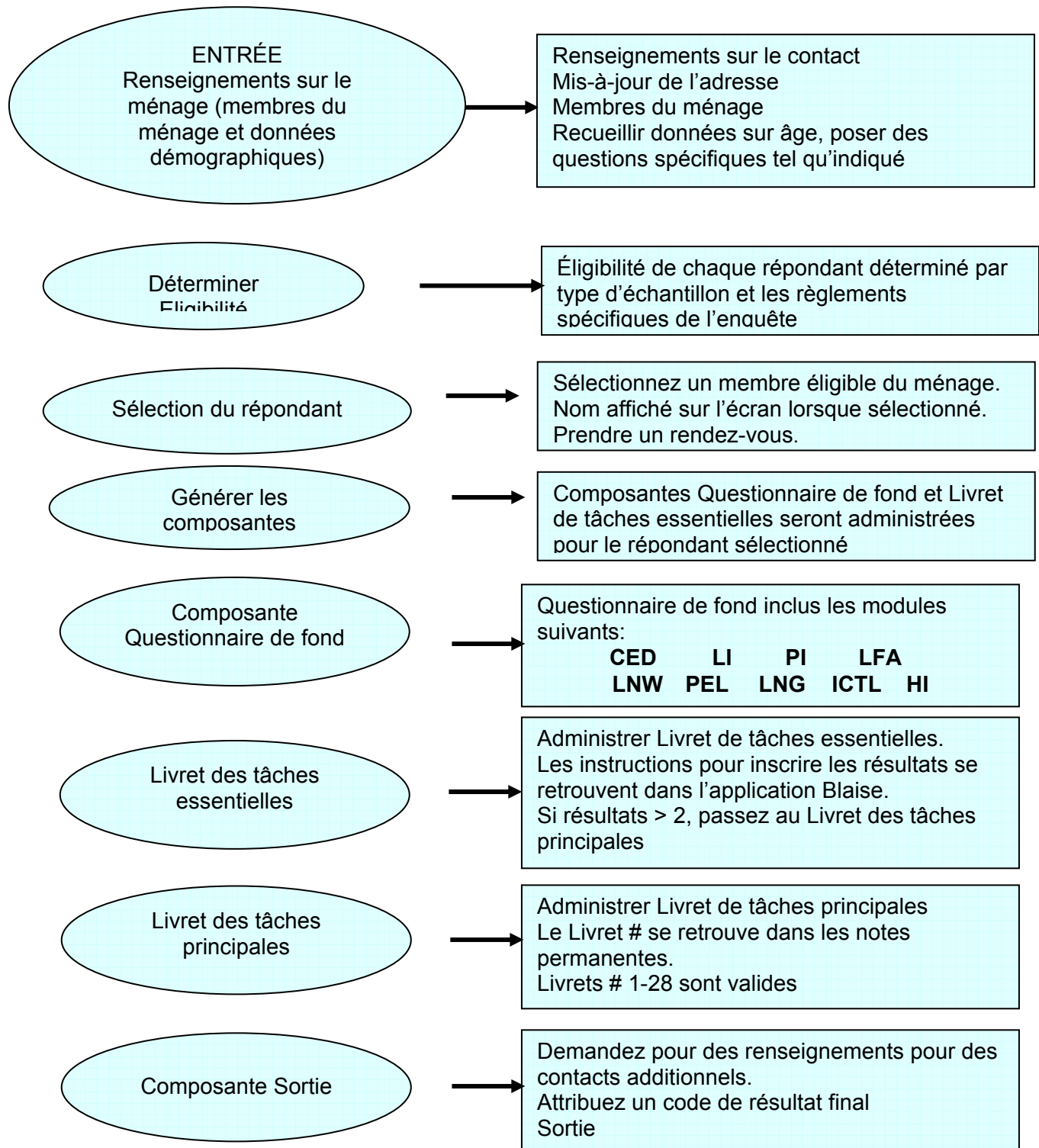
Pour les besoins de l'ELCA, le répondant devait remplir le questionnaire de base, réussir les tâches de littératie de base et tenter d'effectuer au moins cinq tâches par échelle de littératie pour que les chercheurs puissent évaluer directement ses capacités de littératie. On a imputé les données sur les capacités des personnes qui échouaient ou qui refusaient d'effectuer les tâches de base et de celles qui réussissaient les tâches de base sans tenter d'effectuer au moins cinq tâches par échelle de littératie. Comme le modèle utilisé pour imputer les estimations du niveau de littératie des non-répondants repose sur un ensemble complet de réponses aux questions de base, les pays participants ont reçu pour directive de faire remplir au moins le questionnaire de base par les personnes échantillonnées. On leur a également remis une classification détaillée des non-réponses pour les besoins de l'enquête.

Il appartenait à chaque pays de recruter ses intervieweurs. Ainsi, le nombre d'intervieweurs, leur taux de rémunération et la durée de la période d'enquête variaient d'un pays à l'autre, selon les normes et les budgets. Chaque pays a reçu un livret portant sur la formation des intervieweurs.

Au Canada, les données de l'ELCA ont été recueillies par des intervieweurs expérimentés de Statistique Canada au moyen de la technologie de l'interview personnelle assistée par ordinateur.

Le diagramme ci-dessous représente le plan de l'Enquête internationale sur l'alphabétisation et les compétences des adultes.

Enquête internationale sur l'alphabétisation et les compétences des adultes Organigramme



6.3 Notation

On a évalué les compétences en littératie des répondants d'après leur rendement à l'égard des tâches cognitives exécutées au moment de l'enquête. Contrairement aux questions à choix multiple, qu'on utilise habituellement dans les enquêtes à grande échelle et qui proposent un nombre fixe de choix de réponse, les questions ouvertes comme celles qu'on a posées dans le cadre de l'ELCA donnent lieu à un vaste éventail de réponses. Comme les données brutes sont rarement utiles à elles seules, on doit regrouper les réponses, d'une manière ou d'une autre, afin de résumer les résultats relatifs au rendement. À mesure qu'elles étaient notées, les réponses aux questions ouvertes de l'ELCA étaient classées dans trois catégories : correctes, incorrectes ou manquantes.

Les modèles employés pour évaluer les capacités des répondants et la difficulté des tâches reposent sur l'hypothèse selon laquelle les critères de notation établis pour les besoins de l'évaluation ont été appliqués uniformément à l'intérieur des pays et d'un pays à l'autre. Pour vérifier cette hypothèse, on a pris plusieurs mesures, notamment la deuxième notation à l'intérieur d'un pays et la deuxième notation entre les pays, qui sont décrites dans les sections qui suivent.

6.3.1 Deuxième notation à l'intérieur d'un pays

Pour contrôler l'exactitude de la notation, on a établi une méthode fondée sur un taux d'échantillonnage variable. Au début de la notation, presque toutes les réponses ont fait l'objet d'une deuxième notation visant à repérer les correcteurs problématiques et à déceler les réponses uniques ou difficiles qui n'étaient pas prévues dans le manuel de notation. Après avoir atteint un niveau d'exactitude satisfaisant, on a ramené le niveau de la deuxième notation à celui d'une mise à jour pour contrôler l'exactitude de tous les correcteurs. On a calculé des concordances moyennes pour tous les items. Pour s'assurer que la première et la deuxième notation étaient vraiment indépendantes, il a fallu prendre certaines précautions. Par exemple, les correcteurs devaient être des personnes différentes et le deuxième correcteur ne devait pas connaître les notes attribuées par le premier.

On s'attend à ce que les correcteurs ayant reçu une formation identique à l'intérieur d'un pays soient plus cohérents entre eux que par rapport aux correcteurs d'autres pays. Dans la plupart des cas, la fiabilité de la deuxième notation était supérieure à 97 %. Il importe de signaler que les résultats respectaient largement les tolérances statistiques établies pour les besoins de l'Enquête et qu'ils étaient nettement supérieurs à ceux atteints dans d'autres études à grande échelle utilisant des questions ouvertes.

Comme la deuxième notation à l'intérieur d'un pays visait à améliorer la qualité des données, les notes n'ont pas été mises à jour dans la base de données. En d'autres termes, les données sur la concordance présentées ici font état de la concordance minimale atteinte à l'étape de la notation. Après avoir calculé la fiabilité de la notation à l'intérieur d'un pays, on a constaté que quelques correcteurs n'étaient pas fiables. Ces correcteurs ont reçu une formation supplémentaire ou ont été remerciés. Advenant un écart entre la première notation et la deuxième, on a remplacé la première note par la bonne note si l'inexactitude était attribuable à une erreur systématique de la part du premier correcteur. Dans certains cas, le manuel de notation s'est avéré ambigu. On a alors révisé le manuel de notation et modifié la première note en fonction de la révision, mais sans modifier la deuxième note. On n'a jamais remplacé une

deuxième note, même si elle se révélait erronée par la suite. En somme, la première note tient compte des modifications et des corrections résultant des leçons tirées de l'analyse de la deuxième notation à l'intérieur d'un pays. La première note est donc plus exacte et cohérente que la deuxième, qui peut rester erronée et qui, par conséquent, sous-estime quelque peu la fiabilité de la deuxième notation. Cette sous-estimation doit toutefois être très faible puisque, dans la plupart des cas, la fiabilité est supérieure à 97 %. Ces valeurs indiquent que tous les pays participants ont atteint un très haut degré d'uniformité dans la notation.

6.3.2 Deuxième notation entre les pays

Même après avoir vérifié que tous les correcteurs pratiquaient une notation uniforme, remédié aux ambiguïtés des manuels de notation et corrigé les erreurs de notation systématiques, il fallait encore examiner la comparabilité des notes d'un pays à l'autre. Une notation exacte et uniforme à l'intérieur d'un pays ne suppose pas nécessairement que tous les pays appliquent de la même manière les directives des manuels de notation. Un biais lié à la notation peut survenir si un pays note une certaine réponse différemment des autres pays. La deuxième notation entre les pays, décrite dans la présente section, visait à assurer la comparabilité de la notation d'un pays à l'autre.

Comme nous l'avons mentionné plus haut, chaque pays a noté séparément les réponses aux items d'évaluation de l'enquête. Pour déterminer la fiabilité de la notation entre les pays à l'égard de chaque item, deux groupes distincts ont noté les réponses d'un sous-ensemble de répondants. Habituellement, ces groupes de notation étaient de pays différents. Par exemple, un échantillon de livrets de tests a été noté par deux groupes qui avaient noté les livrets en anglais du Canada et ceux des États-Unis. Statistique Canada a calculé la fiabilité de la deuxième notation entre les pays, puis ETS a évalué les résultats. À la suite de cette évaluation, on a demandé à chaque pays d'apporter quelques modifications mineures à la méthode de notation. Dans certains cas, on a constaté que les directives ambiguës dans le manuel de notation entraînaient des erreurs d'interprétation et, par conséquent, une baisse de la fiabilité.

En s'inspirant de la fiabilité de la notation entre les pays, les chercheurs peuvent repérer les items mal conçus, les critères de notation ambigus, les erreurs de traduction et d'impression d'items ou de critères de notation, les inexactitudes attribuables aux correcteurs et, surtout, les cas où un pays pratiquait systématiquement une notation différente de celle d'un autre pays. Dans ce dernier cas, il arrive que les correcteurs d'un pays notent uniformément une certaine réponse comme étant correcte, alors que ceux d'un autre pays notent la même réponse comme étant incorrecte. Avant de procéder à l'échelonnage selon la méthode TRI, il faut éliminer cette asymétrie des notes. ETS et Statistique Canada ont repéré les items asymétriques, et le pays dans lequel le problème de notation était survenu a enquêté sur les causes plausibles de ce biais systématique dans les notes. Lorsqu'on repérait une erreur systématique dans un pays donné, on corrigeait les notes initiales attribuées à l'item en question pour la totalité de l'échantillon.

Le tableau 6.2 résume la fiabilité de la deuxième notation entre les pays pour le volet canadien de l'ELCA, avant corrections.

Tableau 6.2 Fiabilité de la deuxième notation entre les pays

Appariement des pays (pays chargé de la deuxième notation – pays d'origine)	Domaines psychométriques			Total
	Textes suivis et textes schématiques	Numératie	Résolution de problèmes	
	%	%	%	
Canada anglophone – Canada francophone	95	95	92	95
Canada francophone – Canada anglophone	95	97	94	95
Norvège – Canada	91	93	91	92
Canada – États-Unis	94	97	...	95
États-Unis – Canada	95	97	...	95
États-Unis – Bermudes	91	94	...	90
Bermudes – États-Unis	93	95	...	93
Canada francophone – Suisse	95	98	97	96
Suisse – Canada francophone	94	96	94	95
Suisse – Italie	96	98	96	96
Italie – Suisse	93	97	93	94
Canada – Bermudes	83	83
Canada – Nuevo León	91	95 ¹	...	92
Hongrie	94	96	93	94
Pays-Bas	91	93	93	92
Nouvelle-Zélande	96	97	94	96

... sans objet

1. Textes au contenu quantitatif.

6.4 Saisie, traitement et codage des données

Comme condition à leur participation à l'Enquête, on a demandé aux pays de saisir et de traiter leurs fichiers en utilisant des méthodes qui assuraient une cohérence logique et des niveaux acceptables d'erreur dans la saisie des données. Plus précisément, on leur a demandé d'effectuer une vérification complète des notes saisies (en entrant chaque enregistrement deux fois) afin de réduire les taux d'erreur au minimum. L'exactitude de la notation des tests étant essentielle pour assurer un haut degré de qualité des données, il fallait exercer un contrôle intégral par double saisie.

Chaque pays devait également coder l'industrie, la profession et le niveau de scolarité en utilisant les codes types établis à cette fin (Classification internationale type par industrie, CITI; Classification internationale type des professions, CITP; Classification internationale type de l'éducation, CITE). On a également fourni des codes pour les questions ouvertes, et les pays ont reçu des directives précises relativement au codage de ces questions de façon à maintenir les erreurs de codage à des niveaux acceptables.

Pour améliorer la comparabilité de l'analyse des données, on a demandé à chaque pays participant à l'Enquête de schématiser son ensemble de données national sous forme de disposition d'enregistrement très structurée et uniformisée. Outre qu'elle précisait la position, le format et la longueur de chaque champ, la disposition d'enregistrement internationale comprenait une description de chaque variable et indiquait les catégories et les codes à fournir pour cette variable. Dès la réception du fichier d'un pays, Statistique Canada effectuait une série de contrôles des limites pour garantir la conformité avec le format prescrit. Statistique Canada a aussi effectué des contrôles de cohérence et de flux des données des fichiers reçus. Advenant des anomalies, on signalait les problèmes aux pays et on leur demandait de présenter des fichiers épurés.

6.5 Variables dérivées

On a créé un certain nombre de variables dérivées pour favoriser la recherche sur les antécédents et les résultats des compétences. Celles qui ont été ajoutées au fichier principal sont décrites en détail dans cette section.

Variables dérivées simples

Renseignements sur le ménage et le revenu

Enfants à charge de moins de 16 ans vivant dans le ménage

Nom de la VD :	KIDSHOME
Étiquette de la VD :	Enfants à charge de moins de 16 ans vivant dans le ménage
Valeurs de l'étiquette :	0 « Pas d'enfants de moins de 16 ans vivant dans le ménage »; 1 « Enfants de moins de 16 ans vivant dans le ménage »
Question(s) posée(s) :	K1 : Y compris vous-même, combien de personnes vivent dans votre ménage? K2 : Avez-vous des enfants à charge qui vivent avec vous dans votre ménage (enfants dont vous êtes responsable sur le plan financier et/ou dont vous avez la garde exclusive ou partagée)? K3 : Quel est l'âge du plus jeune enfant de votre ménage?
Pseudo-logique :	K1, K2 et K3 déterminent s'il y a un ou des enfants à charge de moins de 16 ans vivant dans le ménage.

7.0 Totalisation et analyse des données

La présente section résume les lignes directrices à respecter par les utilisateurs qui totalisent, analysent, publient ou diffusent des données tirées des bandes de microdonnées de l'enquête. À l'aide de ces lignes directrices, les utilisateurs de microdonnées devraient être en mesure de produire les mêmes chiffres que ceux qui sont produits par Statistique Canada, ainsi que des chiffres encore inédits de manière conforme aux présentes lignes directrices.

7.1 Utilisation de coefficients de pondération dans la totalisation

Les enquêtes sur la littératie reposent sur des plans d'échantillonnage complexes avec stratification, plusieurs degrés de sélection et des probabilités inégales de sélection des répondants. L'utilisation de données tirées d'enquêtes aussi complexes présente des problèmes pour les analystes, car le plan d'enquête et les probabilités de sélection ont une incidence sur les méthodes à retenir pour calculer les estimations et les variances. Afin d'éviter tout biais dans les estimations et dans l'analyse des données, on doit utiliser les coefficients de pondération de l'enquête.

Si bon nombre de méthodes d'analyse qu'on trouve dans les programmes statistiques permettent d'utiliser des coefficients de pondération, la signification ou la définition de ces derniers diffère de celle qui convient dans le cadre d'une enquête par sondage. Par conséquent, alors que les estimations produites par le programme sont souvent exactes, les variances qui sont calculées sont peu précises. Des programmes servant à calculer les erreurs-types d'estimations simples comme des totaux, des proportions et des ratios (dans le cas de variables qualitatives) sont présentés dans la section suivante.

7.2 Définitions de l'estimation catégorique et de l'estimation quantitative

Avant d'aborder les méthodes de totalisation et d'analyse des données de l'ELCA, il convient de décrire les deux principaux types d'estimations ponctuelles des caractéristiques d'une population qu'on peut produire à partir du fichier de microdonnées de l'enquête.

Estimation catégorique

Une estimation catégorique est une estimation du nombre ou du pourcentage de la population visée par une enquête qui possède certaines caractéristiques ou qui entre dans une catégorie donnée. On peut citer comme exemple le nombre d'Albertains qui se situent au niveau 1 sur

l'échelle de compréhension de textes suivis ou la proportion de Canadiens qui se situent au niveau 4 sur l'échelle de la numératie. Dans le cas du nombre de personnes possédant une certaine caractéristique, on peut aussi parler de l'estimation d'un agrégat.

Exemples de questions catégoriques :

Q : Regardez-vous parfois la télévision ou des vidéos dans une autre langue que le français ou l'anglais?

R : **Oui / Non**

Q : Comment coteriez-vous votre capacité de lecture en anglais, dont vous avez besoin dans la vie courante?

R : **Excellente / Bonne / Passable / Médiocre**

Estimation quantitative

Une estimation quantitative est une estimation du total, de la moyenne, de la médiane ou d'une autre mesure de la tendance centrale d'une quantité, fondée sur une partie ou sur la totalité des membres de la population visée par une enquête. Elle se présente essentiellement sous la forme X/Y où X est l'estimation d'un total quantitatif de la population visée par l'enquête et Y est une estimation du nombre de personnes au sein de la population visée par l'enquête qui font partie de ce total.

On peut citer comme exemple d'une estimation quantitative le nombre moyen d'employeurs que les travailleurs canadiens ont eus au cours des 12 derniers mois. Le numérateur est une estimation du nombre total d'employeurs que les travailleurs canadiens ont eus au cours des 12 derniers mois, et le dénominateur est le nombre de Canadiens ayant déclaré qu'ils ont travaillé au cours des 12 derniers mois.

Exemples de questions quantitatives :

Q : Combien d'employeurs différents avez-vous eus au cours des 12 derniers mois?

R : **employeur(s)**

Q : Combien d'heures par semaine travailliez-vous habituellement à cet emploi?

R : **heures**

7.2.1 Totalisation des estimations catégoriques

À partir du fichier de microdonnées, on peut obtenir des estimations du nombre de personnes qui possèdent une certaine caractéristique à l'intérieur d'un pays donné en additionnant les coefficients de pondération finaux de tous les enregistrements présentant la caractéristique en question.

On obtient, pour le pays, des proportions et des ratios de forme X/Y comme suit :

- a) en additionnant les coefficients de pondération finaux des enregistrements présentant la caractéristique d'intérêt pour le numérateur (X),
- b) en additionnant les coefficients de pondération finaux des enregistrements présentant la caractéristique d'intérêt pour le dénominateur (Y), puis
- c) en divisant l'estimation du numérateur par celle du dénominateur.

7.2.2 Totalisation des estimations quantitatives

À partir du fichier de microdonnées, on peut obtenir des estimations quantitatives en multipliant la valeur de la variable d'intérêt par le coefficient de pondération final de chaque enregistrement, puis en additionnant cette quantité pour tous les enregistrements d'intérêt. Par exemple, pour obtenir une estimation du nombre total d'employeurs différents que les travailleurs à temps partiel d'un pays donné ont eus au cours des 12 derniers mois, on multiplie la valeur déclarée en réponse à la question D4 (nombre d'employeurs) par le coefficient de pondération final de l'enregistrement, puis on additionne cette valeur pour tous les enregistrements comportant la valeur D5=2 (temps partiel).

Pour obtenir une moyenne pondérée de forme X/Y , on calcule le numérateur (X) selon la méthode de l'estimation quantitative et le dénominateur (Y), selon la méthode de l'estimation catégorique. Par exemple, pour estimer le nombre moyen d'employeurs que les travailleurs à temps partiel d'un pays donné ont eus au cours des 12 derniers mois :

- a) on estime le nombre total d'employeurs selon la méthode décrite ci-dessus,
- b) on estime le nombre de personnes comprises dans cette catégorie en additionnant les coefficients de pondération finaux de tous les enregistrements comportant la valeur QD5=2, puis
- c) on divise l'estimation a) par l'estimation b).

7.3 Estimations des niveaux de compétence

Le plan d'échantillonnage de l'ELCA est une adaptation du modèle logistique à trois paramètres de la théorie de la réponse d'item. Le premier paramètre (A) est la capacité de l'item de discriminer (sensibilité aux capacités) et le deuxième (B) est sa difficulté. Un troisième paramètre (C) est le paramètre inférieur de l'asymptote, qui représente la chance, possiblement non nulle, d'une bonne réponse, indépendamment de la capacité. Toutefois, comme le test de l'ELCA n'utilisait pas de questions à choix multiple, ce paramètre (C) a été fixé à zéro partout; l'équation est donc devenue, pour ainsi dire, un modèle logistique à deux paramètres. Après avoir calculé les paramètres, on attribue à chaque item une probabilité de réponse de 80 (PR80), qui correspond au niveau de capacités nécessaire à un répondant pour exécuter la tâche avec une probabilité de réussite de 80 %.

Ainsi qu'il est mentionné plus haut, on a résumé les capacités d'un répondant sur les trois échelles en utilisant les paramètres des items et les capacités du répondant conformément aux modèles d'échelonnage selon la méthode TRI. L'application différait quelque peu de la norme puisque, lors de l'ELCA, on a soumis à chaque répondant un nombre relativement restreint d'items afin de cerner plus efficacement les niveaux de capacités de la population. Toutefois, comme les données ne visent pas à estimer des niveaux de capacités individuels, il faut procéder à des analyses plus compliquées.

On a employé la méthode des valeurs plausibles pour estimer uniformément les caractéristiques clés de la population et pour en calculer d'autres avec autant d'exactitude que selon la méthode TRI standard. Cet aspect supplémentaire exige essentiellement que l'estimation des capacités soit fondée sur une série de cinq valeurs plausibles pour chacun des trois aspects de la littératie. Ces cinq valeurs plausibles—allant de prose¹ à prose⁵ pour

l'échelle de compréhension de textes suivis, de doc1 à doc5 pour celle des textes schématiques, de num1 à num5 pour celle de la numératie et de health1 à health5 pour celle des connaissances en matière de santé—ont été recodées en fonction de niveaux plausibles ayant des valeurs de 1 à 5 qui reflètent la progression, déterminée empiriquement, des capacités et des stratégies de traitement de l'information nécessaires pour exécuter des tâches de plus en plus complexes. Le niveau 1 correspond aux notes allant de 0 à 226 (inclusivement); le niveau 2, aux notes de 226,0001 à 276; le niveau 3, aux notes de 276,0001 à 326; le niveau 4, aux notes de 326,0001 à 376 et le niveau 5, aux notes égales ou supérieures à 376,0001. Les variables sont appelées plev1 à plev5 pour l'échelle de compréhension de textes suivis, dlev1 à dlev5 pour celle des textes schématiques et nlev1 à nlev5 pour celle de la numératie.

Le cadre de travail n'étant pas tout à fait le même, on a construit un peu différemment l'échelle de la résolution de problèmes. On a d'abord créé les cinq mêmes valeurs plausibles allant de 0 à 500 (prob1 à prob5), mais la définition des niveaux était légèrement différente. Par exemple, cette échelle ne définit que quatre niveaux de compétence, du niveau 1 (le plus faible) au niveau 4 (le plus élevé). S'il est nécessaire de regrouper les niveaux 4 et 5 afin de répliquer les estimations publiées (on trouve habituellement trop peu de répondants au niveau 5 pour produire des estimations fiables) pour les échelles de textes suivis, de textes schématiques, de numératie et de santé, cette étape n'est toutefois pas nécessaire pour le domaine de la résolution de problèmes.

Pour calculer des estimations ponctuelles simples relatives à l'un des cinq aspects de la littératie, il suffit d'utiliser le coefficient de pondération de la population ainsi que l'une des cinq valeurs plausibles correspondantes (choisies au hasard).

Toutefois, on peut obtenir une estimation ponctuelle plus précise en établissant la moyenne des cinq estimations produites à partir de chacune des cinq valeurs plausibles, qu'on peut calculer comme suit :

$T. = (\sum_i T_i) / 5$, où T_i est un vecteur de cinq estimations pondérées établies d'après chacune des cinq valeurs plausibles.

Il faut retenir qu'en prenant la moyenne des cinq valeurs plausibles, on ne produit qu'une estimation ponctuelle valide, et non une estimation valide de la variance. On doit utiliser **les cinq** valeurs plausibles ainsi que les 30 poids de rééchantillonnage afin de calculer correctement des estimations de la variance fondées sur le plan d'échantillonnage. Ces dernières sont abordées plus en détail à la section 8.1.2. (Utilisation des valeurs plausibles et des poids de rééchantillonnage dans le calcul de l'erreur d'échantillonnage).

7.4 Arrondissement des estimations

Afin que les estimations calculées d'après le fichier de microdonnées et destinées à la publication ou à la diffusion correspondent à celles qui sont produites par Statistique Canada, les utilisateurs doivent respecter les lignes directrices suivantes concernant l'arrondissement de ces estimations :

- a) Les estimations figurant dans le corps principal d'un tableau statistique doivent être arrondies à la centaine la plus proche selon la méthode d'arrondissement normale. Autrement dit, si le premier ou le seul chiffre à supprimer est compris entre 0 et 4, le

dernier chiffre à conserver reste inchangé. Si le premier ou le seul chiffre à supprimer est compris entre 5 et 9, le dernier chiffre à conserver est augmenté de 1. Par exemple, selon la méthode d'arrondissement normale à la centaine la plus proche, si les deux derniers chiffres sont compris entre 00 et 49, ils sont remplacés par 00 et le chiffre précédent (celui des centaines) reste inchangé. Si les deux derniers chiffres sont compris entre 50 et 99, ils sont remplacés par 00 et le chiffre précédent est augmenté de 1.

- b) Les taux partiels marginaux et les totaux marginaux figurant dans les tableaux statistiques doivent être calculés à partir de leurs éléments correspondants non arrondis, puis être arrondis à la centaine la plus proche selon la méthode d'arrondissement normale.
- c) Les moyennes, les proportions, les taux et les pourcentages doivent être calculés à partir d'éléments non arrondis (numérateurs ou dénominateurs), puis être arrondis à une seule décimale selon la méthode d'arrondissement normale. Selon la méthode d'arrondissement normale à un seul chiffre, si le dernier ou le seul chiffre à supprimer est compris entre 0 et 4, le dernier chiffre à conserver reste inchangé. Si le premier ou le seul chiffre à supprimer est compris entre 5 et 9, le dernier chiffre à conserver est augmenté de 1.
- d) Les sommes et les différences d'agrégats (ou ratios) doivent être calculées à partir de leurs éléments correspondants non arrondis, puis être arrondies à la centaine la plus proche (ou à la décimale la plus proche) selon la méthode d'arrondissement normale.
- e) Dans les cas où, à cause de restrictions d'ordre technique ou autre, on utilise une autre méthode d'arrondissement que la méthode d'arrondissement normale, produisant ainsi des estimations qui diffèrent des estimations correspondantes publiées par Statistique Canada, les utilisateurs doivent préciser la raison de ces différences dans les documents destinés à la publication ou à la diffusion.
- f) Sous aucun prétexte les utilisateurs ne doivent publier ou diffuser des estimations non arrondies. Ces dernières laissent supposer que la précision est plus grande qu'elle ne l'est en réalité.

8.0 Qualité des données

On peut évaluer la qualité des données d'une enquête en examinant deux types d'erreur : l'erreur d'échantillonnage et l'erreur non liée à l'échantillonnage.

Les estimations tirées de la présente enquête sont fondées sur un échantillon de personnes. On aurait sans doute obtenu des chiffres quelque peu différents si on avait procédé à un recensement intégral faisant appel au questionnaire, aux intervieweurs, aux superviseurs, aux méthodes de traitement, etc., effectivement utilisés pour les besoins de l'enquête. L'écart entre l'estimation obtenue à partir de l'échantillon et les résultats d'un dénombrement intégral mené dans des conditions semblables est désigné l'erreur d'échantillonnage de l'estimation.

Par ailleurs, des erreurs qui ne sont pas dues à l'échantillonnage peuvent survenir à peu près à chaque étape de l'enquête. L'intervieweur peut mal comprendre une directive, le répondant peut se tromper au moment de répondre, les réponses peuvent être mal transcrites sur le questionnaire, et des erreurs peuvent survenir au moment du traitement et de la totalisation des données. Il s'agit là d'erreurs non liées à l'échantillonnage.

8.1 Erreurs d'échantillonnage

Comme les estimations d'une enquête par sondage sont forcément entachées d'une erreur d'échantillonnage, le chercheur qui tient à adopter une saine méthode statistique voudra donner aux utilisateurs une idée de l'ampleur de cette erreur. La présente section résume les évaluations de l'erreur d'échantillonnage dont Statistique Canada se sert couramment et qu'il incite les utilisateurs à employer eux aussi s'ils produisent des estimations à partir de ce fichier de microdonnées.

Pour mesurer l'ampleur éventuelle des erreurs d'échantillonnage, on fait appel à l'erreur-type des estimations tirées des résultats d'enquête.

Par contre, étant donné la grande diversité des estimations que peut produire une enquête, l'erreur-type d'une estimation s'exprime habituellement par rapport à l'estimation à laquelle elle se rapporte. La mesure qui en découle, désignée le coefficient de variation (c.v.) de l'estimation, s'obtient en divisant l'erreur-type de l'estimation par l'estimation elle-même; elle s'exprime en pourcentage de l'estimation.

Par exemple, supposons que, sur la foi des résultats de l'enquête, on estime que 16,6 % des Canadiens se situent au niveau 1 en compréhension de textes suivis et que l'erreur-type de cette estimation s'élève à 0,013. Le coefficient de variation de l'estimation se calcule donc comme suit :

$$\left(\frac{.013}{.166} \right) \times 100\% = 7.8\%$$

8.1.1 Lignes directrices en matière de diffusion du c.v.

Parmi les critères servant à déterminer si les estimations d'une enquête sont publiables, on retrouve le coefficient de variation (c.v.), qui correspond à l'erreur-type d'une estimation exprimée en pourcentage de l'estimation.

Avant de diffuser ou de publier une estimation de l'ELCA, l'utilisateur doit d'abord déterminer le niveau de qualité de l'estimation, qui peut être acceptable, marginale ou inacceptable. La qualité des données est compromise à la fois par l'erreur d'échantillonnage et par l'erreur non liée à l'échantillonnage. Pour les besoins de la diffusion, cependant, la qualité d'une estimation est déterminée uniquement d'après l'erreur d'échantillonnage reflétée par le coefficient de variation, tel qu'il ressort du tableau 8.1. L'utilisateur doit néanmoins lire la section 8 afin de mieux comprendre les caractéristiques qualitatives de ces données.

D'abord, il convient de déterminer le nombre de répondants dont les données entrent dans le calcul de l'estimation. S'il est inférieur à 30, la qualité de l'estimation pondérée doit être considérée comme inacceptable. Pour une estimation pondérée fondée sur un échantillon d'au moins 30 personnes, l'utilisateur doit déterminer le c.v. de l'estimation et adopter les lignes directrices énoncées ci-dessous, qui s'appliquent aux estimations pondérées arrondies. Toute estimation peut être considérée comme diffusable. Toutefois, celles dont la qualité est marginale ou inacceptable doivent s'accompagner d'une mise en garde à l'intention de l'utilisateur.

Tableau 8.1 Lignes directrices en matière de niveau de qualité

Niveau de qualité de l'estimation	Lignes directrices
1. Acceptable	<p>L'estimation se caractérise par un échantillon de 30 personnes ou plus et un faible coefficient de variation, allant de 0,0 % à 16,5 %.</p> <p>Aucune mise en garde n'est requise.</p>
2. Marginale	<p>L'estimation se caractérise par un échantillon de 30 personnes ou plus et un coefficient de variation élevé, allant de 16,6 % à 33,3 %.</p> <p>Ces estimations doivent être signalées au moyen de la lettre M (ou d'un indicateur semblable). Elles doivent s'accompagner d'une mise en garde à l'intention de l'utilisateur concernant le haut niveau d'erreur associé aux estimations.</p>
3. Inacceptable	<p>L'estimation se caractérise par un échantillon inférieur à 30 personnes ou un coefficient de variation très élevé, supérieur à 33,3 %.</p> <p>Statistique Canada recommande de ne pas diffuser les estimations de qualité inacceptable. Par contre, si l'utilisateur choisit de le faire, les estimations doivent être signalées au moyen de la lettre U (ou d'un indicateur semblable) et s'accompagner de la mise en garde suivante :</p> <p>« Nous informons l'utilisateur que... (préciser la nature des données)... ne répondent pas aux normes de qualité de Statistique Canada pour ce programme statistique. Les conclusions tirées de ces données ne sauraient être fiables et risquent d'être erronées. Ces données et les conclusions qu'on pourrait en tirer ne doivent pas être publiées. Si l'utilisateur choisit de les publier, il est alors tenu de publier également le présent avertissement. »</p>

8.1.2 Utilisation des valeurs plausibles et des poids de rééchantillonnage dans le calcul de l'erreur d'échantillonnage

La présente section reprend essentiellement le contenu de la documentation qui accompagne les programmes STATTOOL (SAS et SPSS) conçus par Statistique Canada pour aider les utilisateurs à manipuler les données de l'ELCA et de l'ELCA. Les programmes et les outils évoqués dans la présente section figurent sur le CD-ROM du fichier de microdonnées à grande diffusion de l'ELCA, dans le répertoire intitulé « STATTOOL ». Certains détails de la présente section sont plutôt axés sur les comparaisons internationales que permet d'établir le fichier de microdonnées à grande diffusion de l'ELCA, mais l'exposé qui suit renseignera également le lecteur sur les utilisations et les limites pratiques des données de l'ELCA.

Calcul des estimations ponctuelles

Dans la présente section, nous allons voir comment utiliser les poids d'échantillonnage (POPWT) pour obtenir des estimations de population sous forme de pourcentages (totaux) et de moyennes. (Le calcul des erreurs-types est présenté dans la section 6.)

Tous les exemples sont fondés sur une population fictive présentant les caractéristiques suivantes :

Simplifier l'analyse provisoire

L'ensemble de données de l'EIACA est certes difficile à manipuler. Les cinq valeurs plausibles correspondant aux cinq domaines (y compris les connaissances en matière de santé), ainsi que les trente poids de rééchantillonnage, compliquent l'évaluation exacte des erreurs-types.

Dans bien des cas, la simplification du processus, notamment à l'étape exploratoire, réduirait grandement le temps de traitement nécessaire à l'analyse de l'estimation des compétences.

Il est donc recommandé, à l'étape de la recherche préliminaire, d'utiliser une seule des valeurs plausibles (VP), au lieu des cinq. Cette méthode est beaucoup plus exacte que le calcul de la moyenne des cinq VP, car ainsi, la répartition pondérée de la population reflète exactement l'estimation ponctuelle. La moyenne des VP, elle, masquerait l'erreur de test et, moins la population étudiée est nombreuse, plus les estimations s'écarteraient de la répartition réelle de la population.

Naturellement, lorsqu'on est prêt à publier les résultats de la recherche, on doit utiliser les poids de rééchantillonnage et les cinq valeurs plausibles pour produire les estimations finales avec des erreurs-types exactes. Cette méthode est décrite en détail dans la section 8.1.2.3.

Type	Sexe	Répartition de la population	Échantillon	
			Répartition (non pondérée)	Répartition (pondérée)
Région rurale	Hommes	40 %	30 %	38 %
	Femmes	60 %	70 %	62 %
	Total	20 %	50 %	19 %
Région urbaine	Hommes	51 %	45 %	50 %
	Femmes	49 %	55 %	50 %
	Total	80 %	50 %	81 %
Total	Hommes	48,8 %	37,5 %	47,7 %
	Femmes	51,2 %	62,5 %	52,3 %
	Total	100 %	100 %	100 %

D'après ce tableau, il semble que la participation des hommes était inférieure à celle des femmes tant en région rurale qu'en région urbaine. Même si les hommes représentent près de 49 % de la population, on ne trouve que 37,5 % des hommes dans l'échantillon. On observe

cette tendance dans les deux types de région. Il semble également que la région rurale soit surreprésentée, car 50 % de l'échantillon proviennent de cette région, contre seulement 20 % de la population.

Toutefois, lorsqu'on utilise les poids d'échantillonnage, les pourcentages sont fort comparables. Comment les calcule-t-on?

Pourcentages (totaux)

On a calculé comme suit le pourcentage pondéré des hommes vivant en région rurale :

$$\text{Weighted \%} = \frac{\sum_{i=1}^{rural, male} POPWT_i}{\sum_{i=1}^{rural} POPWT_i} = 38\% \text{ où } i \text{ représente la personne } i. \text{ Le numérateur est une}$$

estimation de la population totale des hommes vivant en région rurale, et le dénominateur, une estimation de la population totale vivant en région rurale.

On a calculé comme suit le pourcentage non pondéré :

$$\text{Unweighted \%} = \frac{\sum_{i=1}^{rural, male} 1_i}{\sum_{i=1}^{rural} 1_i} = \frac{n_{rural, male}}{n_{rural}} = 30\% \text{ où } n_{rural, male} \text{ est le nombre d'hommes vivant en}$$

région rurale qu'on trouve dans l'échantillon, et n_{rural} , le nombre total de personnes vivant en région rurale qu'on trouve dans l'échantillon.

Dans la dernière équation, chaque personne échantillonnée compte pour un, alors que dans la version pondérée, chaque unité échantillonnée a reçu un poids afin de représenter convenablement et proportionnellement les sous-groupes de l'échantillon (noter que le pourcentage pondéré est un ratio des totaux pondérés estimatifs).

Moyennes

Pour cet exemple fictif, supposons que nous avons également la note moyenne fondée sur la variable PROSE1, comme l'illustre le tableau ci-dessous :

Type	Sexe	Répartition de la population	Échantillon			
			Répartition (non pondérée)	Répartition (pondérée)	Avg. Prose1 (non pondérée)	Avg. Prose1 (pondérée)
Région rurale	Hommes	40 %	30 %	38 %	260	260,1
	Femmes	60 %	70 %	62 %	290	289,8
	Total	20 %	50 %	19 %	281,0	278,5
Région urbaine	Hommes	51 %	45 %	50 %	320	319,7
	Femmes	49 %	55 %	50 %	330	330,1
	Total	80 %	50 %	81 %	325,5	324,9
Total	Hommes	48,8 %	37,5 %	47,7 %	296,0	310,7
	Femmes	51,2 %	62,5 %	52,3 %	307,6	321,0
	Total	100 %	100 %	100 %	303,3	316,1

Ici encore, nous voyons que les moyennes pondérées sont très proches des moyennes non pondérées, pourvu qu'on neutralise le type de région. Ce n'est pas le cas des trois dernières lignes du tableau. Essayons de voir pourquoi. On a calculé comme suit la moyenne pondérée des hommes vivant en région rurale :

$$\text{Weighted mean} = \frac{\sum_{i=1}^{\text{rural,male}} \text{POPWT}_i * \text{PROSE1}_i}{\sum_{i=1}^{\text{rural,male}} \text{POPWT}_i} = 260.1 \quad \text{où } i \text{ représente la personne } i. \text{ Le}$$

numérateur est une estimation de la note totale pour tous les hommes vivant en région rurale, et le dénominateur, une estimation de l'ensemble de la population masculine vivant en région rurale.

On a calculé comme suit la moyenne non pondérée :

$$\text{Unweighted mean} = \frac{\sum_{i=1}^{\text{rural,male}} \text{PROSE1}_i}{\sum_{i=1}^{\text{rural,male}} 1_i} = \frac{\sum_{i=1}^{\text{rural,male}} \text{PROSE1}_i}{n_{\text{rural,male}}} = 260$$

Les résultats non pondérés et pondérés sont semblables lorsque les valeurs de PROSE1 ne varient guère d'une personne à l'autre ou que les valeurs de POPWT présentent le même comportement. Cet énoncé ne vaut pas pour les trois dernières lignes du tableau. On obtient la moyenne pondérée pour les hommes en résolvant l'équation suivante :

$$\begin{aligned} \text{Weighted mean} &= \frac{\sum_{i=1}^{\text{male}} \text{POPWT}_i * \text{PROSE1}_i}{\sum_{i=1}^{\text{male}} \text{POPWT}_i} \\ &= \frac{(38\% * 19\% * 260.1) + (50\% * 81\% * 319.7)}{47.7\%} = 310.7 \end{aligned}$$

alors que la moyenne non pondérée est donnée par :

$$\begin{aligned} \text{Unweighted mean} &= \frac{\sum_{i=1}^{\text{male}} \text{PROSE1}_i}{\sum_{i=1}^{\text{male}} 1_i} \\ &= \frac{(30\% * 50\% * 260) + (45\% * 50\% * 320)}{37.5\%} = 296.0 \end{aligned}$$

Dans la dernière équation, chaque personne échantillonnée compte pour un, alors que dans la version pondérée, chaque unité échantillonnée a reçu un poids afin de représenter convenablement et proportionnellement les sous-groupes de l'échantillon. Par exemple, les 30 %X50 %=15 % des hommes vivant en région rurale qu'on trouve dans l'échantillon ont été rajustés par les poids pour tenir compte des 38 %X19 %=7,22 % de l'ensemble de l'échantillon,

ce qui reflète beaucoup mieux ce qu'on trouve dans l'ensemble de la population. (Il convient de noter que la vraie proportion de la population des hommes vivant en région rurale est 40 %X20 % = 8 %.)

En conclusion, il faut toujours utiliser les poids d'échantillonnage pour calculer des statistiques d'après des données échantillons.

Autres poids d'échantillonnage

Comme nous l'avons vu plus haut, la somme des poids d'échantillonnage (POPWT) à l'intérieur d'un échantillon donne une estimation de la taille de la population. Il s'agit d'un poids d'échantillonnage couramment utilisé, mais la somme donne parfois un nombre très élevé et des nombres différents d'un pays à l'autre, ce qui n'est pas toujours souhaitable. Par exemple, si l'on veut calculer une estimation pondérée du rendement moyen de la population de tous les pays (ou de sous-populations à l'intérieur d'un pays), l'utilisation de la variable POPWT comme variable de pondération fait que chaque pays contribue proportionnellement à la taille de sa population, les grands pays comptant plus que les petits pays. En général, le poids POPWT n'est pas le meilleur poids à utiliser pour les analyses portant sur plusieurs pays. Une autre conséquence de l'utilisation du poids POPWT est la tendance à gonfler les résultats de tests de signification lorsque les logiciels informatiques sont incapables de calculer correctement les données pondérées. Nous allons maintenant voir deux versions possibles de poids d'échantillonnage ayant trait à ces questions en particulier. Ces versions tirent parti du fait qu'on obtient les mêmes estimations de la population sous forme de moyennes et de proportions lorsqu'on utilise une variable de pondération proportionnelle au poids de population (POPWT).

Somme donnant le poids d'échantillonnage constant (CONSTWT)

Il est possible de modifier le poids de population POPWT de sorte que tous les pays contribuent également lorsqu'on calcule la moyenne ou une proportion d'un pays à l'autre. L'équation est la suivante :

$$CONSTWT_{g,i} = POPWT_{g,i} * \left[\frac{100}{\sum_{i=1}^g POPWT_i} \right]$$

pour chaque personne du groupe d'intérêt g. La transformation des poids est différente à l'intérieur de chaque pays mais, en fin de compte, la somme de la variable CONSTWT à l'intérieur de chaque pays est de 100. La variable CONSTWT, à l'intérieur de chaque pays, est proportionnelle à POPWT multiplié par le ratio de 100 divisé par la somme de poids pour toutes les personnes du groupe d'intérêt. On peut utiliser ces poids lorsqu'on veut calculer des estimations internationales et qu'on veut que chaque pays contribue également à l'estimation internationale, sans égard à la taille du groupe d'intérêt dans le pays (voir le tableau ci-dessous).

Groupe d'intérêt	Pays	Chiffre de population (région rurale)	Estimations de population	
			Mean PROSE1 (POPWT)	Mean PROSE1 (CONSTWT)
Région rurale	A	3 700 000	290	290
	B	37 000 000	260	260
	C	7 000 000	300	300
Ensemble			268	283

Somme donnant le poids d'échantillonnage de la taille de l'échantillon (SMPLWT)

Il est possible de modifier le poids de population POPWT lorsqu'on veut utiliser la taille réelle de l'échantillon pour exécuter des tests de signification (à l'intérieur de chaque pays). L'équation est la suivante :

$$SMPLWT_{g,i} = POPWT_{g,i} * \left[\frac{n_g}{\sum_{i=1}^g POPWT_i} \right]$$

pour chaque personne du groupe d'intérêt g où n_g est la taille réelle de l'échantillon du groupe g. La transformation des poids est différente à l'intérieur de chaque pays mais, en fin de compte, la somme de la variable CONSTWT à l'intérieur de chaque pays donne la taille de l'échantillon du groupe g. La variable SMPLWT, à l'intérieur de chaque pays, est proportionnelle à POPWT multiplié par le ratio de la taille de l'échantillon (n_g) divisé par la somme des poids pour toutes les personnes du groupe d'intérêt. Si certains logiciels de statistique permettent d'utiliser la taille de l'échantillon comme diviseur dans le calcul des erreurs-types, d'autres utilisent plutôt la somme des poids, ce qui entraîne de graves erreurs-types à la baisse pour les statistiques si l'on utilise POPWT comme variable de pondération. Lorsqu'on effectue des analyses à l'aide de ce genre de logiciel, il est recommandé d'utiliser une variable comme SMPLWT comme variable de pondération. En raison de l'effet de regroupement dans les échantillons de la plupart des pays, il convient également d'appliquer à la variable SMPLWT un facteur de correction, par exemple un effet du plan de sondage.

Utilisation des valeurs plausibles pour calculer des estimations ponctuelles

Pour atteindre son objectif d'une vaste couverture des buts et des processus de la littératie, l'ELCA reposait sur un éventail d'items répartis en livrets d'évaluation. Chaque participant à l'évaluation a rempli un seul livret, ce qui a réduit au minimum le fardeau de réponse individuel. L'ELCA a utilisé un plan d'échantillonnage matriciel pour attribuer les livrets d'évaluation aux répondants afin qu'on puisse broser un tableau complet du rendement en littératie dans chaque pays à partir des composantes remplies par chaque répondant. L'ELCA a utilisé l'échelonnage selon la théorie de la réponse d'item (TRI) pour combiner les réponses individuelles en vue d'établir des estimations exactes du rendement en littératie dans la population de chaque pays. L'échelle TRI de l'ELCA utilise également la méthodologie de l'imputation multiple ou des « valeurs plausibles » pour calculer les notes de compétence en littératie de tous les répondants, même si chacun n'a répondu qu'à une partie de l'ensemble des items d'évaluation.

La plupart des tests portant sur les compétences cognitives visent à évaluer exactement le rendement de répondants individuels à des fins de diagnostic, de sélection ou de placement. On peut améliorer l'exactitude de ces mesures en augmentant le nombre d'items remis à la personne. Pour connaître la répartition des compétences dans une vaste population, toutefois, on peut obtenir des estimations plus efficaces à l'aide d'un plan d'échantillonnage matriciel. Ce type de plan sollicite un petit nombre de réponses de chaque répondant échantillonné, tout en maintenant une vaste représentation du contenu lorsque les réponses sont agrégées pour tous les répondants. Avec cette méthode, toutefois, l'avantage d'estimer les caractéristiques d'une population est plutôt contrebalancé par l'incapacité de formuler des énoncés précis au sujet des personnes; l'agrégation des notes individuelles risque donc de biaiser gravement les estimations des caractéristiques d'une population.

On a mis au point la méthodologie des valeurs plausibles afin de régler ce problème en utilisant toutes les données disponibles pour estimer directement les caractéristiques de populations et de sous-populations, puis en produisant plusieurs notes imputées (appelées valeurs plausibles) d'après ces distributions, qu'on peut utiliser dans les analyses avec un logiciel de statistique courant. Mislevy (1991) présente un examen détaillé de la méthodologie des valeurs plausibles. Les principaux points à retenir sont les suivants :

- a) Lorsqu'on veut calculer des statistiques portant sur des notes (PROSE, DOC, NUMERACY, etc.), on ne dispose pas d'une, mais de cinq valeurs de notes attribuées à chaque personne. Chaque ensemble de valeurs plausibles convient tout aussi bien pour estimer les paramètres de population.
- b) Ces statistiques fondées sur des notes doivent toujours être calculées au niveau d'une population ou d'une sous-population. Il ne faut jamais les utiliser pour établir une inférence au niveau individuel.

Utilisation des valeurs plausibles

Exemple1 : Médiane estimative de la variable PROSE dans le pays A.

Pour chaque personne, on ne dispose pas d'une, mais de cinq notes à prendre en compte, comme le montre le tableau ci-dessous.

Pays A	PROSE1	PROSE2	PROSE3	PROSE4	PROSE5
Personne 1	222	275	300	245	254
Personne 2	289	310	212	250	265
...
Personne n	285	275	243	321	312
Médiane	285	281	283	279	289

Afin d'estimer la médiane globale pour la variable PROSE dans le pays A, on estime d'abord la médiane à partir du premier ensemble de valeurs plausibles correspondant à la variable PROSE1. On répète ensuite cette première étape en utilisant les variables PROSE2 à PROSE5 pour obtenir au total cinq estimations tout aussi valables de la médiane pour le pays en question. Comme elles sont toutes aussi valables, l'étape suivante en vue d'obtenir une seule estimation de la médiane consiste à calculer la moyenne de ces cinq estimations. On obtient ainsi :

$$\text{Médiane globale} = (285 + 281 + 283 + 279 + 289) / 5 = 283,4$$

Notez qu'il ne faut pas calculer la moyenne des notes au niveau individuel. Par exemple, $(222 + 275 + 300 + 245 + 254) / 5 = 259,2$ ne constitue pas une bonne estimation de la variable PROSE pour la personne 1. Il en va de même de chacune des variables PROSEn pour une personne donnée. Les variables fondées sur les notes doivent toujours être interprétées dans le contexte de populations ou de sous-populations.

Les variables PROSE1 à PROSE5 sont des notes correspondant aux capacités brutes. Lorsqu'on regroupe ces notes en niveaux, elles donnent cinq variables de niveau appelées, d'après le domaine de la compréhension de textes suivis, PLEV1 à PLEV5 dans le FMGD. Les variables XPROSE1 à XPROSE5 représentent un recodage des variables PLEV1 à PLEV5 qu'on trouve dans le FMGD de l'ELCA, les niveaux 4 et 5 étant regroupés afin de donner un nombre suffisant de répondants à chaque niveau pour permettre une analyse exacte. On peut

procéder de même pour les variables de niveau axées sur les textes schématiques, la numératie et la santé (DLEV1 à DLEV5, NLEV1 à NLEV5 et HLEV1 à HLEV5). Quant aux niveaux de la résolution de problèmes (PSLEV1 à PSLEV5), on les laisse tels quels.

Exemple 2 : Coefficients estimatifs des paramètres de régression logistique (niveaux 2 et 3 des variables dépendantes XPROSE1 à XPROSE5).

D'après les valeurs trouvées dans le tableau précédent, on obtiendrait :

Pays A	XPROSE1	XPROSE2	XPROSE3	XPROSE4	XPROSE5
Personne 1	2	3	3	2	2
Personne 2	3	4/5	2	2	2
...
Personne n	3	2	2	3	3

On remarque avant tout qu'une personne donnée peut se trouver à différents niveaux de compétence, selon l'ensemble de valeurs plausibles pris en compte. Cette caractéristique n'invalide pourtant pas la méthodologie utilisée. Comme nous l'avons expliqué plus haut, afin d'obtenir les coefficients des paramètres de régression logistique au niveau des pays, on calcule d'abord ces paramètres en fonction du premier ensemble de valeurs plausibles. On répète ensuite cette étape quatre fois en utilisant les variables XPROSE2 à XPROSE5 pour obtenir quatre autres ensembles de paramètres estimatifs, comme le montre le tableau ci-dessous :

Pays A	XPROSE1	XPROSE2	XPROSE3	XPROSE4	XPROSE5
Coordonnée à l'origine	0,124	0,129	0,122	0,125	0,120
Bêta 1	1,051	1,059	1,049	1,055	1,060
Bêta 2	0,584	0,591	0,545	0,499	0,645
Bêta 3	3,222	4,123	3,012	3,542	3,201

Comme ces ensembles de coefficients estimatifs des paramètres sont tous aussi valables les uns que les autres, l'étape suivante en vue d'obtenir un seul ensemble d'estimations consiste à calculer la moyenne de ces cinq résultats. On obtient ainsi :

Coordonnée globale à l'origine = $(0,124 + 0,129 + 0,122 + 0,125 + 0,120) / 5 = 0,124$

Bêta 1 global = $(1,051 + 1,059 + 1,049 + 1,055 + 1,060) / 5 = 1,055$

Bêta 2 global = $(0,548 + 0,591 + 0,545 + 0,499 + 0,645) / 5 = 0,566$

Bêta 3 global = $(3,222 + 4,123 + 3,012 + 3,542 + 3,201) / 5 = 3,420$

8.1.3 Estimation de la variance d'erreur dans le cadre de l'ELCA

La méthodologie de l'ELCA est une méthode éprouvée de quantification des compétences. Il existe donc déjà une importante documentation concernant les méthodes d'estimation de la variance relativement à ce genre d'études. Le texte de la présente section provient en grande partie du guide de l'utilisateur du PIRLS 2001 de l'IEA; il a été adapté au contexte de l'ELCA.

Aperçu

Lorsqu'on analyse des données obtenues au moyen d'un plan de sondage complexe comme celui de l'ELCA, il importe de calculer des estimations exactes de la variance d'erreur à l'égard des statistiques d'intérêt. Dans le cas de l'ELCA, cette variance d'erreur peut provenir de deux sources : le processus d'échantillonnage (toujours présent) et le processus d'imputation (lorsque les statistiques d'intérêt portent sur des notes de compétence). La présente section décrit les méthodes d'estimation de ces composantes de la variance d'erreur.

Estimation de la variance d'échantillonnage

Sur le plan analytique, lorsqu'on recueille des données dans le cadre d'une enquête par sondage complexe, il est souvent difficile de produire une estimation de la variance qui soit sans biais ou cohérente avec le plan de sondage. Une catégorie de techniques appelées méthodes de rééchantillonnage permet d'estimer la variance pour le type de plans d'échantillonnage complexes comme ceux qui sont utilisés dans le cas de l'ELCA.

Le principe de base du rééchantillonnage consiste à choisir successivement des sous-échantillons dans l'ensemble de l'échantillon, à calculer la statistique d'intérêt pour chaque sous-échantillon, puis à utiliser ces sous-échantillons ou poids statistiques de rééchantillonnage pour estimer la variance de la statistique fondée sur l'échantillon complet. Différentes façons de créer des sous-échantillons à partir de l'échantillon complet donnent différentes méthodes de rééchantillonnage. Les sous-échantillons sont appelés poids de rééchantillonnage et les statistiques calculées à partir de ces poids sont appelées estimations répétées.

L'une de ces méthodes est la technique de rééchantillonnage jackknife (JRR) (Wolter, 1985). Dans le cas de l'ELCA, à l'intérieur de chaque pays, l'échantillon complet a été scindé au hasard en 30 sous-ensembles de taille égale ou presque égale, chaque sous-ensemble ressemblant à l'échantillon complet. On forme les poids de rééchantillonnage en supprimant un sous-ensemble à la fois et en multipliant les poids des autres sous-ensembles par

$$\frac{30}{29}$$

On crée ainsi 30 poids de rééchantillonnage. Cette méthode est également appelée la méthode JK1. Les poids liés à chaque poids de rééchantillonnage correspondent aux variables REPLIC01 à REPLIC30. On utilise ces poids uniquement pour calculer la variance d'échantillonnage. Les estimations ponctuelles doivent être calculées selon la méthode décrite dans la section précédente.

Calcul de la variance d'échantillonnage selon la méthode JK1

Le principe de base consiste ici à calculer l'estimation d'intérêt d'après l'échantillon complet en utilisant la variable de pondération POPWT (ou SMPLWT ou CONSTWT) ainsi que chaque poids de rééchantillonnage (en utilisant les variables REPLIC01 à REPLIC30). On utilise ensuite la variation entre les estimations fondées sur les poids de rééchantillonnage et l'estimation fondée sur l'échantillon complet pour estimer la variance d'échantillonnage pour l'échantillon complet. Supposons que $\hat{\theta}$ est l'estimation fondée sur l'échantillon complet d'un paramètre de population θ . L'estimateur de la variance d'échantillonnage $\text{var}(\hat{\theta})$ est donné par :

$$\text{var}_{\text{smp}}(\hat{\theta}) = \frac{30 \sum_{g=1}^{30} (\hat{\theta}_{(g)} - \hat{\theta})^2}{29} \quad \text{où } \hat{\theta}_{(g)} \text{ est l'estimation de } \theta \text{ fondée sur les observations comprises dans le } g^{\text{e}} \text{ poids de rééchantillonnage.}$$

Lorsque la statistique d'intérêt porte sur des notes de compétence, on fonde habituellement la variance d'échantillonnage uniquement sur le premier ensemble de valeurs plausibles au lieu de calculer cinq fois l'expression ci-dessus et de calculer la moyenne des cinq variances d'échantillonnage estimatives.

Exemple 1 : Revenu personnel moyen dans le pays A

Le revenu personnel moyen est donné par l'expression suivante :

$$\text{Mean personal income} = \frac{\sum_{i=1}^{\text{Country A}} \text{POPWT}_i * D43_i}{\sum_{i=1}^{\text{Country A}} \text{POPWT}_i} = 27\,601$$

Afin de calculer la variance d'échantillonnage, on calcule 30 fois l'expression ci-dessous, chaque calcul étant fondé sur le poids de rééchantillonnage pertinent.

$$\text{Mean personal income}_{(1)} = \frac{\sum_{i=1}^{\text{Country A}} \text{REPLIC01}_i * D43_i}{\sum_{i=1}^{\text{Country A}} \text{REPLIC01}_i} = 26\,983$$

$$\text{Mean personal income}_{(2)} = \frac{\sum_{i=1}^{\text{Country A}} \text{REPLIC02}_i * D43_i}{\sum_{i=1}^{\text{Country A}} \text{REPLIC02}_i} = 26\,146$$

...

$$\text{Mean personal income}_{(30)} = \frac{\sum_{i=1}^{\text{Country A}} \text{REPLIC30}_i * D43_i}{\sum_{i=1}^{\text{Country A}} \text{REPLIC30}_i} = 28\,965$$

Puis, on applique simplement la formule de variance donnée plus haut. On obtient ainsi :

$$\text{var}_{\text{simpl}}(\hat{\theta}) = \frac{29 \sum_{g=1}^{30} (\hat{\theta}_{(g)} - \hat{\theta})^2}{30} = \frac{29}{30} [(26983 - 27601)^2 + (26146 - 27601)^2 + \dots + (28965 - 27601)^2]$$

Enfin, la statistique $(\hat{\theta} - \theta) / \text{var}(\hat{\theta})^{1/2}$ fait l'objet d'une distribution de t approximative à 29 degrés de liberté.

Estimation de la variance d'imputation

Lorsque les statistiques d'intérêt portent sur des notes de compétence, il est nécessaire d'estimer la variance d'imputation. Comme nous l'avons mentionné dans la section précédente, on produit cinq notes par test pour les participants à l'ELCA. Ces différentes notes sont appelées valeurs plausibles (VP). Lors de l'ELCA, on a remis à chaque personne des blocs d'exercices. L'ensemble de blocs couvre tous les concepts à évaluer, mais chaque répondant ne répondait pas aux questions de tous les blocs. Grâce à un type d'affectation équilibrée de blocs aux répondants, on a couvert l'ensemble complet de questions en agrégeant les réponses des répondants. Pour un groupe de répondants semblables, on a estimé une distribution bayésienne a posteriori des notes. Les valeurs plausibles pour chaque répondant sont réalisées à partir de cette distribution a posteriori. Ces notes ne sont pas significatives pour un répondant individuel mais, en les combinant, on peut les utiliser pour estimer des moyennes et d'autres quantités de population.

Calcul de la variance d'imputation

La méthode générale d'estimation de la variance d'imputation au moyen des valeurs plausibles est la suivante :

- a) D'abord, on estime la statistique d'intérêt θ en utilisant chaque fois un ensemble différent de valeurs plausibles (M) et la variable POPWT (ou SMPLWT ou CONSTWT). Appelons ces cinq estimations $\hat{\theta}_1$ à $\hat{\theta}_5$. La statistique d'intérêt peut être n'importe quel élément estimable d'après les données-échantillons : une moyenne, l'écart entre des moyennes, des centiles, des coefficients de paramètre de régression, etc.
- b) Puis, on arrive à l'estimation globale en calculant la moyenne de $\hat{\theta}_m$ où $m=1, 2, \dots, 5$. Appelons cette estimation $\hat{\theta}$.
- c) La variance d'imputation est calculée comme suit :

$$Var_{imp}(\hat{\theta}) = \left[1 + \frac{1}{5}\right] \times \sum_{m=1}^5 \frac{(\hat{\theta}_m - \hat{\theta})^2}{4}$$

Estimation de la variance d'erreur globale

Dans des conditions idéales et avec des ressources informatiques illimitées, on calculerait la variance d'erreur globale comme suit :

$$Var(\hat{\theta}) = \sum_{m=1}^5 \frac{Var_{smp}(\hat{\theta}_m)}{5} + Var_{imp}(\hat{\theta})$$

Comme chacune des Var_{smp} suppose qu'on calcule les statistiques d'intérêt 30 fois (en utilisant chaque fois un poids d'échantillonnage différent), on peut utiliser plutôt cette formule abrégée consistant à estimer la variance d'échantillonnage à partir du premier ensemble de valeurs plausibles seulement.

$$Var(\hat{\theta}) = Var_{smp}(\hat{\theta}_1) + Var_{imp}(\hat{\theta})$$

Lorsque les statistiques d'intérêt ne portent sur aucune note de compétence, la formule de la variance d'erreur globale devient simplement la suivante :

$$Var(\hat{\theta}) = Var_{smp}(\hat{\theta})$$

Degrés de liberté

$(\hat{\theta} - \theta) / Var(\hat{\theta})^{1/2}$ fait l'objet d'une distribution de t approximative, les degrés de liberté (Jonhson et Rust, 1993) étant donnés par :

$$\nu = \frac{1}{\frac{f_m^2}{4} + \frac{(1-f_m)^2}{29}}$$

où f_m est donné par :

$$f_m = \frac{Var_{imp}(\hat{\theta})}{Var(\hat{\theta})}$$

En pratique, le nombre de degrés de liberté est fixé à 29 (ce qui est efficace lorsque f est relativement petit, par exemple inférieur à 30 %).

Établissement de comparaisons

Nous allons maintenant voir comment calculer la bonne variance d'erreur quand on compare les estimations d'enquête entre les pays, avec les estimations internationales et à l'intérieur d'un pays. Afin de simplifier le texte, nous envisageons uniquement le rendement moyen estimatif pour la variable PROSE. On devrait pouvoir généraliser directement le contenu de la présente section à n'importe quel type d'estimation d'enquête.

Comparaisons entre les pays

Quand on compare le rendement moyen estimatif pour la variable PROSE entre les pays A et B, la variance d'erreur est donnée par :

$$Var(\hat{\theta}) = Var_{Country A}(\hat{\theta}) + Var_{Country B}(\hat{\theta})$$

Par exemple, supposons que le rendement moyen estimatif est, pour le pays A, de 290 avec une variance d'erreur de 25 et, pour le pays B, de 307 avec une variance d'erreur de 30,25. L'écart entre ces deux pays est donc de $307 - 290 = 17$. La question est la suivante : cet écart de 17 points résulte-t-il d'une erreur due au fait qu'on n'a échantillonné qu'une partie de la population et qu'on n'a administré qu'une partie des items? Pour répondre à cette question, il faut d'abord calculer la statistique suivante (appelée statistique de Wald) :

$$(\hat{\theta})/Var(\hat{\theta})^{1/2} = 17/(25 + 30.25)^{1/2} = 2.287$$

Quand on compare cette valeur à la valeur critique de 95 % d'une distribution de t à 24 degrés de liberté (2,04), on conclut que les données suffisent à affirmer que ces deux pays n'ont pas le même rendement moyen estimatif.

Notez que cette méthode est également valide quand on compare les résultats de l'ELCA à ceux de l'ELCA.

Comparaisons avec les estimations internationales

Parmi les statistiques publiées, l'une des plus importantes montre le rendement moyen de votre pays par rapport à la moyenne internationale. Ici, la variance d'erreur est donnée par :

$$Var(\hat{\theta}) = \frac{(N-1)^2 Var_{smp\ A}(\hat{\theta}_A) + \sum_{k=1, k \neq A}^N Var_{smp\ k}(\hat{\theta}_k)}{N^2} + Var_{imp}(\hat{\theta}_A - \hat{\theta}_{int})$$

où N est le nombre de pays utilisés pour calculer la moyenne internationale, $\hat{\theta}_m$ représente la moyenne estimative pour le pays m , et $\hat{\theta}_{int}$ représente la moyenne internationale estimative.

Prenons, par exemple, les tableaux suivants :

	Rendement moyen	Variance d'échantillonnage
Pays A	290	20
Pays B	300	22
Pays C	286	18
Pays D	324	22
International	300	

Rendement moyen	PROSE1	PROSE2	PROSE3	PROSE4	PROSE5
Pays A	288	292	292	288	290
International	301	300	299	302	298
$\hat{\theta}_A - \hat{\theta}_{int}$	-13	-8	-7	-14	-8

Ici, on constate que le rendement moyen estimatif du pays A est de 290 avec une variance d'échantillonnage de 20 (variance d'imputation de 5) et que le rendement moyen international estimatif, fondé sur quatre pays, est de 300. L'écart entre le résultat du pays A et la moyenne internationale est donc de $290 - 300 = -10$. La question est la suivante : cet écart de 10 points résulte-t-il d'une erreur due au fait qu'on n'a échantillonné qu'une partie de la population et qu'on n'a administré qu'une partie des items? Pour répondre à cette question, il faut d'abord calculer la statistique suivante :

$$Var(\hat{\theta}) = \frac{(3)^2 20 + (22 + 18 + 22)}{4^2} + (1 + \frac{1}{5}) \frac{[(13-10)^2 + (8-10)^2 + (7-10)^2 + (14-10)^2 + (8-10)^2]}{4}$$

$$= 15.125 + 12.6 = 27.725$$

On calcule alors la statistique de Wald :

$$(\hat{\theta}) / Var(\hat{\theta})^{1/2} = 10 / (27.725)^{1/2} = 1.899$$

Quand on compare cette valeur à la valeur critique de 95 % d'une distribution de t à 24 degrés de liberté (2,04), on conclut que les données ne suffisent pas à affirmer que le rendement moyen estimatif du pays A est différent du rendement moyen international estimatif.

Comparaisons à l'intérieur d'un pays

La plupart du temps, quand on compare des sous-groupes à l'intérieur d'un pays, il n'existe pas de formule directe pour calculer la variance d'erreur globale comme c'était le cas dans les deux sections précédentes. La principale raison en est qu'habituellement, les échantillons des différents sous-groupes ne sont pas traités comme indépendants pour les besoins de tests statistiques. Par conséquent, il faut utiliser une méthode jackknife applicable à des échantillons corrélés pour estimer la variance d'échantillonnage de l'écart entre les sous-groupes. Il s'agit de calculer l'écart entre les sous-groupes une fois pour chacun des 30 échantillons répétés, et cinq fois de plus, soit une fois par ensemble de valeurs plausibles, comme nous l'avons décrit plus haut (voir les sections sur le calcul des variances d'échantillonnage et d'imputation).

Toutefois, on peut aisément utiliser des modèles de régression linéaire pour calculer ces écarts. Voici comment calculer l'écart entre les hommes et les femmes pour la variable PROSE :

- a) Créer des variables nominales pour les sous-groupes; appelons MAN la variable qui prend la valeur 1 si le répondant est un homme et 0 dans le cas contraire, et WOMAN la variable qui prend la valeur 1 si le répondant est une femme et 0 dans le cas contraire. On crée autant de variables nominales qu'il y a de sous-groupes.
- b) En utilisant POPWT, exécuter un modèle de régression linéaire avec PROSE1 comme variable dépendante et MAN ou WOMAN comme variable indépendante. (Lorsqu'il y a plus de variables nominales k et que k est supérieur à 2, choisir $k-1$ variables nominales comme variables indépendantes.) La variable nominale laissée pour compte devient le sous-groupe de référence.
- c) À l'aide des poids de rééchantillonnage, répéter l'étape b) 30 fois.
- d) À l'aide des poids POPWT, répéter l'étape b) quatre fois, soit une fois par ensemble de valeurs plausibles.
- e) Combiner les renseignements tirés des étapes b), c) et d) pour calculer l'estimation ponctuelle globale, la variance d'erreur et la statistique de Wald.

Par exemple, supposons qu'après avoir créé les variables nominales, on obtient, à la suite l'étape b, le résultat suivant :

Mean PROSE1 = 270 + 18*WOMAN (pondéré par POPWT)

Cette expression devient simplement Mean PROSE1 = 270 lorsque les répondants sont des hommes et Mean PROSE1 = 288 lorsque les répondants sont des femmes. Cela signifie que le coefficient qui précède la variable WOMAN dans le modèle de régression est l'écart entre les femmes et les hommes, alors que la coordonnée à l'origine (270) est le rendement moyen pour le niveau de référence, en l'occurrence les hommes.

À partir de l'étape c, on obtient :

Mean PROSE1 = 271 + 17*WOMAN (pondéré par REPLIC01)

Mean PROSE1 = 269 + 19*WOMAN (pondéré par REPLIC02)

Mean PROSE1 = 273 + 14*WOMAN (pondéré par REPLIC03)

...

Mean PROSE1 = 268 + 21*WOMAN (pondéré par REPLIC30)

On peut maintenant calculer comme suit la variance d'échantillonnage de l'écart entre les femmes et les hommes :

$$\text{var}_{\text{simpl}}(\hat{\theta}) = \frac{29 \sum_{g=1}^{30} (\hat{\theta}_{(g)} - \hat{\theta})^2}{30} = \frac{29}{30} [(17-18)^2 + (19-18)^2 + (14-18)^2 + \dots + (21-18)^2] = 19.575$$

À partir de l'étape d, on obtient :

Mean PROSE2 = 271 + 20*WOMAN (pondéré par POPWT)

Mean PROSE3 = 269 + 19*WOMAN (pondéré par POPWT)

Mean PROSE4 = 273 + 17*WOMAN (pondéré par POPWT)

Mean PROSE5 = 268 + 16*WOMAN (pondéré par POPWT)

On peut maintenant calculer l'estimation ponctuelle globale de l'écart en calculant la moyenne des résultats pour PROSE1 à PROSE5. Cela donne : $(18+20+19+17+16) / 5 = 18$. La variance d'imputation est donnée par :

$$\begin{aligned} \text{Var}_{\text{imp}}(\hat{\theta}) &= \left[1 + \frac{1}{5}\right] \times \sum_{m=1}^5 \frac{(\hat{\theta}_m - \hat{\theta})^2}{4} \\ &= 1.2 \times \frac{[(18-18)^2 + (20-18)^2 + (19-18)^2 + (17-18)^2 + (16-18)^2]}{4} = 3.0 \end{aligned}$$

La statistique de Wald devient :

$$(\hat{\theta}) / \text{Var}(\hat{\theta})^{1/2} = 18 / (19.575 + 3.0)^{1/2} = 3.788$$

Quand on compare cette valeur à la valeur critique de 95 % d'une distribution de t à 24 degrés de liberté (2,04), on conclut que les données suffisent à affirmer que le rendement moyen des hommes est différent de celui des femmes à l'intérieur du pays A.

8.1.4 Analyse des données de l'ELCA sur SPSS

La présente section donne des exemples élémentaires d'analyses qu'on peut effectuer en utilisant les poids d'échantillonnage et les notes dont il a été question dans les sections précédentes. Elle fournit également des détails sur certains programmes SPSS servant à mener ce genre d'analyse et sur les résultats de ces analyses. Les analyses présentées ici sont de nature simple. Les programmes calculent le pourcentage de répondants dans certains sous-groupes, le rendement moyen de ces groupes et les erreurs-types correspondantes (racine carrée de la variance d'erreur totale) pour donner des statistiques sous forme de pourcentages et de moyennes.

Dans nos exemples, nous utilisons des macros rédigées en SPSS pouvant servir à effectuer toutes les analyses décrites dans la présente section. Il s'agit de méthodes générales qu'on peut utiliser à plusieurs fins, à condition de posséder une connaissance de base du macrolangage SPSS. Si vous possédez une certaine expérience de la programmation avec ce progiciel statistique, vous serez en mesure d'apporter les modifications nécessaires aux macros pour obtenir les résultats souhaités.

Les macros SPSS

Les quatre macros SPSS disponibles sont décrites ci-dessous :

JACKMEAN.SPS

On peut utiliser cette macro SPSS pour calculer des pourcentages pondérés de répondants à l'intérieur de groupes définis et leur moyenne à l'égard d'une valeur continue donnée. Cette macro calcule également les variances d'échantillonnage JRR pour les pourcentages et les estimations moyennes. La variable peut être n'importe quelle variable continue contenue dans le fichier.

JACKMEANPV.SPS

On peut utiliser cette macro en SPSS pour calculer des pourcentages pondérés de répondants à l'intérieur de groupes définis et leurs notes de rendement moyen sur une échelle de rendement en utilisant des valeurs plausibles. Cette macro utilise des valeurs plausibles pour calculer des notes de rendement moyen. Elle calcule également les variances d'échantillonnage selon la technique de rééchantillonnage jackknife (JRR) pour les pourcentages de répondants à l'intérieur de groupes donnés, ainsi que les variances JRR et les variances d'imputation pour les notes de rendement moyen. On utilise cette macro uniquement lorsqu'on utilise plusieurs valeurs plausibles dans les analyses.

JACKREG.SPS

On peut utiliser cette macro en SPSS pour calculer des coefficients de régression linéaire et les erreurs-types correspondantes à l'intérieur de groupes définis. On peut l'utiliser avec n'importe quelle variable de l'analyse, mais elle n'utilise pas les valeurs plausibles.

JACKREGPV.SPS

On peut utiliser cette macro SPSS pour calculer des coefficients de régression linéaire et les erreurs-types correspondantes lorsqu'on utilise des valeurs plausibles comme variables dépendantes à l'intérieur de groupes définis.

Moyennes et pourcentages en l'absence de valeurs plausibles

La présente section propose un exemple de code SPSS pouvant servir à calculer les erreurs-types pour des moyennes et des pourcentages de variables autres que des valeurs plausibles. Ce code est fourni sous forme d'un programme SPSS appelé **JACKMEAN.SPS** qui calcule les pourcentages de répondants à l'intérieur de sous-groupes définis par un ensemble de variables de classification, les erreurs-types de ces pourcentages, les moyennes pour les groupes à l'égard d'une variable choisie et les erreurs-types de ces moyennes. Les erreurs-types calculées par cette macro SPSS tiennent compte du plan d'échantillonnage de l'ELCA.

Si vous utilisez cette macro, vous devez spécifier un ensemble de variables de classification, une variable d'analyse, le nombre de poids de rééchantillonnage (si ce nombre est le même pour différents pays, vous pouvez fusionner les ensembles de données des pays; sinon effectuez l'analyse pays par pays), les poids de rééchantillonnage et le poids de population à utiliser pour l'analyse. Vous devez également spécifier le fichier de données qui contient les données à traiter.

Pour utiliser efficacement la macro, vous devez connaître la syntaxe SPSS de base. Vous devez d'abord inclure la macro dans le fichier de programme où vous l'utiliserez. Si vous travaillez en mode différé, vous devez exécuter la macro pour chaque lot. Si vous utilisez le SPSS en mode interactif, vous devez exécuter la macro une fois au début de la session; elle restera active pendant toute la session. Si vous mettez fin à la session ou si vous la relancez plus tard, vous devez exécuter la macro de nouveau. Une fois la macro comprise dans une session donnée, évitez d'utiliser le mot « JACKMEAN » dans ce programme, ce qui aurait pour effet d'exécuter la macro.

Cette macro comporte les paramètres suivants :

INFILE Nom du fichier de données contenant les variables nécessaires à l'analyse (si l'emplacement du chemin est compris dans le nom du fichier, le nom du fichier doit figurer entre guillemets). Inclure uniquement les cas d'intérêt pour l'analyse (par ex., il faut exclure les répondants pour lesquels il manque des variables avant d'exécuter la macro).

CVAR Liste des variables à utiliser pour classer les répondants dans le fichier de données. Il peut s'agir d'une seule variable ou d'une liste de variables. Il est recommandé de toujours inclure la variable correspondant au pays. En l'occurrence, il fallait spécifier au moins une variable (par ex., CNTRID).

DVAR Variable pour laquelle il faut calculer des moyennes. Une seule variable doit être inscrite ici.

NJKZ Indication du nombre de poids de rééchantillonnage établis dans le fichier de données. Si vous travaillez avec les données d'un seul pays, vous devez appliquer à l'argument NJKZ autant de poids de rééchantillonnage qu'il en faut dans le pays. (Si vous travaillez avec

Les données de plus d'un pays, assurez-vous que tous les ensembles de données comportent le même nombre de poids de rééchantillonnage.)

RPWT Poids de rééchantillonnage compris dans les fichiers de données, généralement REPLIC01 à REPLIC30. Les poids de rééchantillonnage doivent être spécifiés comme suit : « REPLIC01 TO REPLIC30 ».

WGT Poids d'échantillonnage à utiliser dans l'analyse, généralement POPWT.

La manière la plus simple d'exécuter la macro consiste à utiliser la notation SPSS courante servant à appeler des macros. Il s'agit d'inscrire le nom de la macro, suivi par la liste correspondante d'arguments nécessaires à l'analyse, chacun étant séparé par une barre oblique. Par exemple, si vous exécutez la macro en utilisant le code suivant :

```
Include "c:\jackmean.sps".
```

```
Jackmean
```

```
Infile = temp           /  
Cvar = cntrid           /  
Dvar = d43              /  
Njkz = 30               /  
Rpwt = replic01 to replic30 /  
Wgt = popwt.
```

Elle calcule la moyenne du revenu personnel (D43) et l'erreur-type correspondante, à l'intérieur de chaque pays, en prenant la variable POPWT comme poids d'échantillonnage. Les données sont lues à partir du fichier système TEMP.

Le fichier qui contient ces résultats est appelé FINAL et est conservé dans le répertoire par défaut utilisé par le SPSS. Les variables contenues dans ce fichier sont les suivantes :

Variables de classification

Chacune des variables de classification est conservée dans le fichier résultant. Il y a une seule occurrence par combinaison spécifique de catégories de variables de classification.

Variable de pondération

Cette variable contient l'estimation de la taille de population des groupes définis par combinaison spécifique de catégories de variables de classification. Dans l'exemple ci-dessus, cette variable est appelée POPWT.

MNX

Cette variable contient les moyennes de la variable DVAR pour les groupes définis par la combinaison correspondante de catégories de variables de classification.

MNX_SE

Cette variable contient les erreurs-types des valeurs MNX calculées selon la méthode jackknife.

PCT

Cette variable contient les pourcentages de personnes dans les groupes pour la dernière variable de classification inscrite, à l'intérieur de la combinaison spécifiée des catégories définies à l'origine par les groupes. Dans notre exemple, on obtiendrait les pourcentages de répondants par pays.

PCT_SE

Cette variable contient les erreurs-types de PCT calculées selon la méthode jackknife.

On peut alors imprimer le fichier obtenu à l'aide de cette macro en utilisant la méthode SPSS de son choix. En voici un exemple :

```
get file = "x:\IALSS\IALSSDATA.sav"
  / keep = cntrid gendaa2 d43 popwt replic01 to replic30 .

select if (gendaa2=1 or gendaa2=2) and not(missing(d43)).

save outfile = respondent.

include "c:\IALSS\jackmean.sps".

jackmean infile= respondent          /
      cvar = cntrid gendaa2          /
      dvar = d43                     /
      njkz = 30                      /
      rpwt = replic01 to replic30    /
      wgt = popwt.

print formats   cntrid gendaa2 n (F6.0) popwt (f10.0) mnx mnx_se pct pct_se (f8.2).

report format=list automatic / var = cntrid gendaa2 n popwt mnx mnx_se pct pct_se.
```

CNTRID	Gender of Respondent	N	POPWT	MNX	MNX_SE	PCT	PCT_SE
22	1	170	109777	70949.93	7977.68	60.39	3.03
22	2	190	72007	45423.49	6084.25	39.61	3.03

Moyennes et pourcentages en présence de valeurs plausibles

La présente section donne un exemple de code SPSS pouvant servir à calculer les erreurs-types pour les valeurs plausibles et les pourcentages moyens. Ce code est fourni sous forme d'une macro SPSS appelée **JACKMEANPV.SPS** qui calcule les pourcentages de répondants à l'intérieur de sous-groupes définis par un ensemble de variables de classification, les erreurs-types de ces pourcentages, les moyennes des groupes sur l'une des échelles de rendement utilisant les valeurs plausibles, ainsi que les erreurs-types de ces moyennes. Les erreurs-types calculées par cette macro SPSS tiennent compte du plan d'échantillonnage de l'ELCA et des composantes de la variance d'imputation.

Si vous utilisez cette macro, vous devez spécifier un ensemble de variables de classification, le nom et le nombre des valeurs plausibles, le nombre de poids de rééchantillonnage (si ce nombre est le même pour différents pays, vous pouvez fusionner les ensembles de données des pays, sinon effectuer l'analyse pays par pays), les poids de rééchantillonnage et le poids de population à utiliser pour l'analyse. Vous devez également spécifier le fichier de données qui contient les données à traiter.

Pour utiliser efficacement la macro, vous devez connaître la syntaxe SPSS de base. Vous devez d'abord inclure la macro dans le fichier de programme où vous l'utiliserez. Si vous travaillez en mode différé, vous devez exécuter la macro pour chaque lot. Si vous utilisez le SPSS en mode interactif, vous devez exécuter la macro une fois au début de la session; elle restera active pendant toute la session. Si vous mettez fin à la session ou si vous la relancez plus tard, vous devez exécuter la macro de nouveau. Une fois la macro comprise dans une session donnée, évitez d'utiliser le mot « JACKMEANPV » dans ce programme, ce qui aurait pour effet d'exécuter la macro.

Cette macro comporte les paramètres suivants :

INFILE Nom du fichier de données contenant les variables nécessaires à l'analyse (si l'emplacement du chemin est compris dans le nom du fichier, le nom du fichier doit figurer entre guillemets). Inclure uniquement les cas d'intérêt pour l'analyse (par ex., il faut exclure les répondants pour lesquels il manque des variables avant d'exécuter la macro).

CVAR Liste des variables à utiliser pour classer les répondants dans le fichier de données. Il peut s'agir d'une seule variable ou d'une liste de variables. Il est recommandé de toujours inclure la variable correspondant au pays. En l'occurrence, il fallait spécifier au moins une variable (par ex., CNTRID).

PVS Valeurs plausibles à utiliser dans l'analyse. Les valeurs plausibles doivent être spécifiés comme suit : « Valeur plausible 1 TO Valeur plausible 5 », par exemple « PROSE1 TO PROSE5 ». Dans la plupart des cas, vous voudrez utiliser les cinq valeurs plausibles, mais le programme fonctionne également lorsqu'on en spécifie moins. Utilisez toujours au moins deux valeurs plausibles.

NPV Nombre de valeurs plausibles à utiliser pour l'analyse. En général, vous voudrez utiliser les cinq valeurs plausibles mais, dans certains cas, vous pouvez en utiliser moins (voir PVS ci-dessus).

NJKZ Indication du nombre de poids de rééchantillonnage établis dans le fichier de données. Si vous travaillez avec les données d'un seul pays, vous devez appliquer à l'argument NJKZ autant de poids de rééchantillonnage qu'il en faut dans le pays. (Si vous travaillez avec les données de plus d'un pays, assurez-vous que tous les ensembles de données comportent le même nombre de poids de rééchantillonnage.)

RPWT Poids de rééchantillonnage compris dans les fichiers de données, généralement REPLIC01 à REPLIC30. Les poids de rééchantillonnage doivent être spécifiés comme suit : « REPLIC01 TO REPLIC30 ».

WGT Poids d'échantillonnage à utiliser dans l'analyse, généralement POPWT.

La manière la plus simple d'exécuter la macro consiste à utiliser la notation SPSS courante servant à appeler des macros. Il s'agit d'inscrire le nom de la macro, suivi par la liste correspondante d'arguments nécessaires à l'analyse, chacun étant séparé par une barre oblique. Par exemple, si vous exécutez la macro en utilisant le code suivant :

```
Include "c:\jackmeanpv.sps".
```

```
Jackmeanpv
```

```
      Infile = temp           /  
      Cvar  = cntrid gendaa2  /  
      PVS   = PROSE1 to PROSE5 /  
      NPV   = 5               /  
      Njkz  = 30              /  
      Rpwt  = replic01 to replic30 /  
      Wgt   = popwt.
```

Elle calcule le rendement moyen en compréhension de textes suivis et l'erreur-type correspondante pour les hommes et les femmes à l'intérieur de chaque pays, en utilisant cinq valeurs plausibles et la variable POPWT comme poids d'échantillonnage. Elle calcule également les pourcentages d'hommes et de femmes à l'intérieur du pays et les erreurs-types correspondantes. Les données sont lues à partir du fichier système TEMP.

Le fichier qui contient ces résultats est appelé FINAL et est conservé dans le répertoire par défaut utilisé par le SPSS. Les variables contenues dans ce fichier sont les suivantes :

Variables de classification

Chacune des variables de classification est conservée dans le fichier résultant. Il y a une seule occurrence par combinaison spécifique de catégories de variables de classification.

Variable de pondération

Cette variable contient l'estimation de la taille de population des groupes définis par combinaison spécifique de catégories de variables de classification. Dans l'exemple ci-dessus, cette variable est appelée POPWT.

N

Cette variable contient le nombre de cas dans les groupes définis par combinaison spécifique de catégories de variables de classification.

MNX

Cette variable contient les moyennes de la première valeur plausible pour les groupes définis par la combinaison correspondante de catégories de variables de classification.

MNX_SE

Cette variable contient les erreurs-types de la moyenne pour la première valeur plausible pour les groupes calculée selon la méthode jackknife. Elle ne comprend pas l'erreur d'imputation.

PCT

Cette variable contient les pourcentages de personnes dans les groupes pour la dernière variable de classification inscrite, à l'intérieur de la combinaison spécifiée des catégories définies à l'origine par les groupes. Dans notre exemple, il s'agit du pourcentage d'hommes et de femmes à l'intérieur de chaque pays.

PCT_SE

Cette variable contient les erreurs-types de la variable PCT calculées selon la méthode jackknife.

MNPV

Cette variable contient les moyennes des valeurs plausibles pour les groupes définis par la combinaison correspondante de catégories de variables de classification.

MNPV_SE

Cette variable contient les erreurs-types pour la moyenne des valeurs plausibles pour les groupes calculée selon la méthode jackknife. Elle comprend les composantes échantillonnage et imputation.

On peut alors imprimer le fichier obtenu à l'aide de cette macro en utilisant la méthode SPSS de son choix. En voici un exemple :

```
get file = "x:\IALSS\IALSSdata.sav"
  / keep = cntrid gendaa2 popwt replic01 to replic30 prose1 to prose5.

select if (gendaa2=1 or gendaa2=2) .

save outfile = respondent.

include "c:\IALSS\jackmeanpv.sps".

jackmeanpv      infile= respondent      /
                  cvar = cntrid gendaa2      /
                  pvs = prose1 to prose5      /
                  npv=5                      /
                  njkz = 30                  /
                  rpwt = replic01 to replic30 /
                  wgt = popwt.

print formats cntrid gendaa2 (F2.0) n (F4.0) popwt (f7.0)  mnpv  mnpv_se mnx mnx_se
pct pct_se (f6.2).

report format=list automatic margin(1,255)
  / var = cntrid gendaa2 n popwt mnpv mnpv_se mnx mnx_se pct pct_se.
```

CNTRID	Gender of Respondent	N	POPWT	MNPV	MNPV_SE	MNX	MNX_SE	PCT	PCT_SE
22	1	1605	1179970	230.41	1.11	230.39	.99	49.53	.06
22	2	3196	1202504	226.13	.99	226.64	.91	50.47	.06

Coefficients de régression en l'absence de valeurs plausibles

La présente section donne un exemple de code SPSS pouvant servir à calculer des coefficients de régression linéaire et les erreurs-types correspondantes. Ce code est fourni sous forme d'une macro SPSS appelée **JACKREG.SPS** qui calcule la corrélation multiple entre les variables dépendantes et indépendantes spécifiées, ainsi que les coefficients de régression et les erreurs-types correspondantes. Les erreurs-types calculées par cette macro SPSS tiennent compte du plan d'échantillonnage de l'ELCA.

Si vous utilisez cette macro, vous devez spécifier un ensemble de variables de classification, les variables dépendantes et indépendantes, le nombre de poids de rééchantillonnage (si ce nombre est le même pour différents pays, vous pouvez fusionner les ensembles de données des pays, sinon effectuer l'analyse pays par pays), les poids de rééchantillonnage et le poids de population à utiliser pour l'analyse. Vous devez également spécifier le fichier de données qui contient les données à traiter.

Pour utiliser efficacement la macro, vous devez connaître la syntaxe SPSS de base. Vous devez d'abord inclure la macro dans le fichier de programme où vous l'utiliserez. Si vous travaillez en mode différé, vous devez exécuter la macro pour chaque lot. Si vous utilisez le SPSS en mode interactif, vous devez exécuter la macro une fois au début de la session; elle restera active pendant toute la session. Si vous mettez fin à la session ou si vous la relancez plus tard, vous devez exécuter la macro de nouveau. Une fois la macro comprise dans une session donnée, évitez d'utiliser le mot « JACKREG » dans ce programme, ce qui aurait pour effet d'exécuter la macro.

Cette macro comporte les paramètres suivants :

INFILE Nom du fichier de données contenant les variables nécessaires à l'analyse (si l'emplacement du chemin est compris dans le nom du fichier, le nom du fichier doit figurer entre guillemets). Inclure uniquement les cas d'intérêt pour l'analyse (par ex., il faut exclure les répondants pour lesquels il manque des variables avant d'exécuter la macro).

CVAR Liste des variables à utiliser pour classer les répondants dans le fichier de données. Il peut s'agir d'une seule variable ou d'une liste de variables. Il est recommandé de toujours inclure la variable correspondant au pays. En l'occurrence, il fallait spécifier au moins une variable (par ex., CNTRID).

XVAR Liste de variables indépendantes (au moins une) qui, selon le modèle de régression linéaire, servent de prédicteurs de la variable dépendante spécifiée dans la variable DVAR. Il peut s'agir de variables continues ou catégoriques ou de tout autre type de variable codée.

DVAR Variable dépendante qui, selon le modèle de régression, est prévue par la ou les variables spécifiées par le paramètre XVAR. Une seule variable doit être inscrite ici.

NJKZ Indication du nombre de poids de rééchantillonnage établis dans le fichier de données. Si vous travaillez avec les données d'un seul pays, vous devez appliquer à l'argument NJKZ autant de poids de rééchantillonnage qu'il en faut dans le pays. (Si vous travaillez avec les données de plus d'un pays, assurez-vous que tous les ensembles de données comportent le même nombre de poids de rééchantillonnage.)

RPWT Poids de rééchantillonnage compris dans les fichiers de données, généralement REPLIC01 à REPLIC30. Les poids de rééchantillonnage doivent être spécifiés comme suit : « REPLIC01 TO REPLIC30 ».

WGT Poids d'échantillonnage à utiliser dans l'analyse, généralement POPWT.

La manière la plus simple d'exécuter la macro consiste à utiliser la notation SPSS courante servant à appeler des macros. Il s'agit d'inscrire le nom de la macro, suivi par la liste correspondante d'arguments nécessaires à l'analyse, chacun étant séparé par une barre oblique. Par exemple, si vous exécutez la macro en utilisant le code suivant :

```
Include "c:\jackreg.sps".
```

```
Jackreg
```

```
Infile = temp           /  
Cvar = cntrid           /  
Xvar = regsex           /  
Dvar = d43              /  
Njkz = 30               /  
Rpwt = replic01 to replic30 /  
Wgt = popwt.
```

Elle calcule l'équation de régression pour la variable REGSEX comme prédicteur du revenu personnel. Les données sont lues à partir du fichier système TEMP.

Le fichier qui contient ces résultats est appelé REG et est conservé dans le répertoire par défaut utilisé par le SPSS. Les variables contenues dans ce fichier sont les suivantes :

Variables de classification

Chacune des variables de classification est conservée dans le fichier résultant. Il y a une seule occurrence par combinaison spécifique de catégories de variables de classification.

Mult_RSQ

Carré du coefficient de corrélation multiple du modèle.

SS_Res, SS_Reg, SS_Total

Résidu, régression et somme totale des carrés du modèle à l'intérieur de chaque groupe défini par les variables de classification.

Coefficients de régression et erreurs-types (B## et B##.SE)

Il s'agit des coefficients de régression pour chacune des variables prédictives du modèle et des erreurs-types correspondantes selon la méthode jackknife. Le coefficient zéro (B00) est la coordonnée à l'origine du modèle. Les autres coefficients reçoivent un numéro d'ordre, en commençant par 01. Ce numéro correspond à l'ordre des variables sur la liste des variables spécifiées dans le paramètre XVAR.

On peut alors imprimer le fichier obtenu à l'aide de cette macro en utilisant la méthode SPSS de son choix. En voici un exemple :

```

get file = "x:\IALSS\IALSSdata.sav"
  / keep = cntrid gendaa2 d43 popwt replic01 to replic30.

select if (gendaa2=1 or gendaa2=2) .    compute regsex = gendaa2 - 1.

save outfile = respondent.

include "c:\IALSS\jackreg.sps".

jackreg              infile= respondent              /
                    cvar = cntrid                    /
                    xvar = regsex                    /
                    dvar = d43                      /
                    njkz = 30                       /
                    rpwt = replic01 to replic30 /
                    wgt = popwt.

print formats        cntrid (F2.0) n (F4.0) mult_RSQ (f5.3)
                    SS_Total SS_Reg SS_Res (F10.0) B00 B00.SE B01 B01.SE (f6.2).

report format=list automatic margin(1,255)
  / var = cntrid n Mult_RSQ SS_Total SS_Reg SS_Res B00 B00.SE B01 B01.SE .

```

CNTRID	N	MULT_RSQ	SS_TOTAL	SS_REG	SS_RES	B00	B00.SE	B01	B01.SE
22	360	.023	1.2E+15	2.8E+13	1.18E+15	70950	7977.7	-25526	9560.6

Dans cet exemple, on crée la variable REGSEX en soustrayant un de la variable GENDAA2. Par conséquent, les hommes reçoivent un code de 0 et les femmes, un code de 1 à l'égard de cette variable. Dans ce modèle, on utilise la variable REGSEX pour prévoir les valeurs de la variable D43 (revenu personnel). Le modèle devient :

Revenu personnel = 70950(7978) pour les hommes,
 Revenu personnel = 70950(7978) – 25526(9561) pour les femmes.

Les nombres entre parenthèses sont les erreurs-types. Cela signifie que les femmes ont, en moyenne, un revenu personnel qui est de 25 526 \$ inférieur à celui des hommes et que 9 561 \$ est l'erreur-type liée à cette estimation.

Coefficients de régression en présence de valeurs plausibles

La présente section donne un exemple de code SPSS pouvant servir à calculer des coefficients de régression linéaire en prenant des valeurs plausibles comme variable dépendante et les erreurs-types correspondantes. Ce code est fourni sous forme d'une macro SPSS appelée **JACKREGPV.SPS** qui calcule la corrélation multiple moyenne entre les valeurs plausibles et les variables indépendantes spécifiées, ainsi que les coefficients de régression et les erreurs-types correspondantes. Les erreurs-types calculées par cette macro SPSS tiennent compte du plan d'échantillonnage de l'ELCA.

Si vous utilisez cette macro, vous devez spécifier un ensemble de variables de classification, les variables dépendantes et indépendantes, le nombre de poids de rééchantillonnage (si ce nombre est le même pour différent pays, vous pouvez fusionner les ensembles de données des pays, sinon effectuer l'analyse pays par pays), les poids de rééchantillonnage et le poids de population à utiliser pour l'analyse. Vous devez également spécifier le fichier de données qui contient les données à traiter.

Pour utiliser efficacement la macro, vous devez connaître la syntaxe SPSS de base. Vous devez d'abord inclure la macro dans le fichier de programme où vous l'utiliserez. Si vous travaillez en mode différé, vous devez exécuter la macro pour chaque lot. Si vous utilisez le SPSS en mode interactif, vous devez exécuter la macro une fois au début de la session; elle restera active pendant toute la session. Si vous mettez fin à la session ou si vous la relancez plus tard, vous devez exécuter la macro de nouveau. Une fois la macro comprise dans une session donnée, évitez d'utiliser le mot « JACKREGPV » dans ce programme, ce qui aurait pour effet d'exécuter la macro.

Cette macro comporte les paramètres suivants :

INFILE Nom du fichier de données contenant les variables nécessaires à l'analyse (si l'emplacement du chemin est compris dans le nom du fichier, le nom du fichier doit figurer entre guillemets). Inclure uniquement les cas d'intérêt pour l'analyse (par ex., il faut exclure les répondants pour lesquels il manque des variables avant d'exécuter la macro).

CVAR Liste des variables à utiliser pour classer les répondants dans le fichier de données. Il peut s'agir d'une seule variable ou d'une liste de variables. Il est recommandé de toujours inclure la variable correspondant au pays. En l'occurrence, il fallait spécifier au moins une variable (par ex., CNTRID).

XVAR Liste de variables indépendantes (au moins une) qui, selon le modèle de régression linéaire, servent de prédicteurs de la variable dépendante spécifiée par les valeurs plausibles. Il peut s'agir de variables continues ou catégoriques ou de tout autre type de variable codée.

ROOTPV Préfixe servant à identifier les valeurs plausibles pour l'échelle de rendement qui nous intéresse. Par exemple, la racine des valeurs plausibles en compréhension de textes suivis est « PROSE ».

NPV Nombre de valeurs plausibles à utiliser pour l'analyse. En général, vous voudrez utiliser les cinq valeurs plausibles mais, dans certains cas, vous pouvez en utiliser moins (voir PVS ci-dessus).

NJKZ Indication du nombre de poids de rééchantillonnage établis dans le fichier de données. Si vous travaillez avec les données d'un seul pays, vous devez appliquer à l'argument NJKZ autant de poids de rééchantillonnage qu'il en faut dans le pays. (Si vous travaillez avec les données de plus d'un pays, assurez-vous que tous les ensembles de données comportent le même nombre de poids de rééchantillonnage.)

RPWT Poids de rééchantillonnage compris dans les fichiers de données, généralement REPLIC01 à REPLIC30. Les poids de rééchantillonnage doivent être spécifiés comme suit : « REPLIC01 TO REPLIC30 ».

WGT Poids d'échantillonnage à utiliser dans l'analyse, généralement POPWT.

La manière la plus simple d'exécuter la macro consiste à utiliser la notation SPSS courante servant à appeler des macros. Il s'agit d'inscrire le nom de la macro, suivi par la liste correspondante d'arguments nécessaires à l'analyse, chacun étant séparé par une barre oblique. Par exemple, si vous exécutez la macro en utilisant le code suivant :

Include "c:\jackregpv.sps".

Jackregpv

Infile	= temp	/
Cvar	= cntrid	/
Xvar	= regsex	/
Rootpv	= Prose	/
NPV	= 5	/
Njkz	= 30	/
Rpwt	= replic01 to replic30	/
Wgt	= popwt.	

Elle calcule l'équation de régression pour la variable REGSEX comme prédicteur des valeurs plausibles en compréhension de textes suivis. Les données sont lues à partir du fichier système TEMP.

Le fichier qui contient ces résultats est appelé REG et est conservé dans le répertoire par défaut utilisé par le SPSS. Les variables contenues dans ce fichier sont les suivantes :

Variables de classification

Chacune des variables de classification est conservée dans le fichier résultant. Il y a une seule occurrence par combinaison spécifique de catégories de variables de classification.

Mult_RSQ

Carré du coefficient de corrélation multiple du modèle.

SS_Res, SS_Reg, SS_Total

Résidu, régression et somme totale des carrés du modèle à l'intérieur de chaque groupe défini par les variables de classification.

Coefficients de régression et Erreurs-types (B## et B##.SE)

Il s'agit des coefficients de régression pour chacune des variables prédictives du modèle et des erreurs-types correspondantes selon la méthode jackknife (avec les composantes échantillonnage et imputation). Le coefficient zéro (B00) est la coordonnée à l'origine du modèle. Les autres coefficients reçoivent un numéro d'ordre, en commençant par 01. Ce numéro correspond à l'ordre des variables sur la liste de variables spécifiées dans le paramètre XVAR.

On peut alors imprimer le fichier obtenu à l'aide de cette macro en utilisant la méthode SPSS de son choix. En voici un exemple :

```

get file = "x:\IALSS\IALSSdata.sav"
  / keep = cntrid gendaa2 popwt replic01 to replic30 prose1 to prose5.

select if (gendaa2=1 or gendaa2=2) . compute regex = gendaa2 - 1.

save outfile = respondent.

include "c:\IALSS\jackregpv.sps".

jackregpv      infile      = respondent
                  cvar      = cntrid
                  xvar      = regex
                  rootpv    = prose
                  npv       = 5
                  njkz      = 30
                  rpwt      = replic01 to replic30
                  wgt       = popwt.

print formats      cntrid (F2.0) n (F4.0) mult_RSQ (f5.3)
                  SS_Total SS_Reg SS_Res (F12.0) B00 B00.SE B01 B01.SE (f6.2) .

report format=list automatic margin(1,255)
  / var = cntrid n Mult_RSQ SS_Total SS_Reg SS_Res B00 B00.SE B01 B01.SE .

```

CNTRID	N	MULT_RSQ	SS_TOTAL	SS_REG	SS_RES	B00	B00.SE	B01	B01.SE
22	4801	.002	4395153088	10952856	4384200232	230.41	1.11	-4.27	1.50

Dans cet exemple, on crée la variable REGSEX en soustrayant un de la variable GENDAA2. Par conséquent, les hommes reçoivent un code de 0 et les femmes, un code de 1 à l'égard de cette variable. Dans ce modèle, on utilise la variable REGSEX pour prévoir les valeurs des valeurs plausibles en compréhension de textes suivis (PROSE). Le modèle devient :

Prose = 230,41 (1,11) pour les hommes,
 Prose = 230,41 (1,11) – 4,27(1,50) pour les femmes.

Les chiffres entre parenthèses sont les erreurs-types. Cela signifie que les femmes ont, en moyenne, une note en compréhension de textes suivis qui est de 4,27 inférieure à celle des hommes et que 1,50 est l'erreur-type liée à cette estimation.

8.1.5 Analyse des données de l'ELCA sur SAS

La présente section donne des exemples élémentaires d'analyses qu'on peut effectuer en utilisant les poids d'échantillonnage et les notes dont il a été question dans les sections précédentes. Elle fournit également des détails sur un programme SAS donné servant à mener ce genre d'analyse et sur les résultats de ces analyses. Les analyses présentées ici sont de nature simple. Le programme calcule le pourcentage de répondants dans certains sous-groupes, le rendement moyen de ces groupes, les nombres pondérés de répondants dans certains groupes, les centiles estimatifs de ces groupes, les coefficients de régression et coefficients de régression logistique ainsi que les erreurs-types correspondantes (racine carrée de la variance d'erreur totale).

Dans nos exemples, nous utilisons une macro rédigée en SAS pouvant servir à effectuer toutes les analyses décrites dans la présente section. Il s'agit de méthodes générales qu'on peut utiliser à plusieurs fins, à condition de posséder une connaissance de base du macrolangage SAS. Si vous possédez une certaine expérience de la programmation avec ce progiciel statistique, vous serez en mesure d'apporter les modifications nécessaires aux macros pour obtenir les résultats souhaités.

La macro SAS

La seule macro SAS disponible est décrite comme suit : **STATTOOL.SAS**

On peut utiliser ce macroprogramme en SAS pour calculer plusieurs statistiques : moyennes, centiles, fréquences, dénombrements, écarts (différences) et régressions (régression linéaire type, régression logistique et régression multinomiale). Ces statistiques sont calculées à l'intérieur de groupes définis en tenant compte des poids d'échantillonnage. Cette macro calcule également les erreurs-types JRR avec les composantes échantillonnage et imputation.

Analyses de base : moyennes, pourcentages, dénombrements, centiles, coefficients de régression et erreurs-types correspondantes

La présente section donne un exemple de code SAS pouvant servir à calculer des moyennes, des pourcentages, des dénombrements, des centiles, des coefficients de régression et les erreurs-types correspondantes pour n'importe quel type de variable (qu'il s'agisse ou non de valeurs plausibles). Ce code est fourni sous forme d'une macro SAS appelée STATTOOL.SAS qui calcule ces statistiques pour des répondants à l'intérieur de sous-groupes définis par un ensemble de n'importe quelle variable de classification (fondée ou non sur des valeurs plausibles). Les erreurs-types calculées par cette macro SAS tiennent compte des composantes échantillonnage et imputation.

Si vous utilisez cette macro, vous devez spécifier un ensemble de variables de classification, une analyse variable, le nombre de poids de rééchantillonnage (si ce nombre est le même pour différents pays, vous pouvez fusionner les ensembles de données des pays; sinon effectuez l'analyse pays par pays), les poids de rééchantillonnage et le poids de population à utiliser pour l'analyse. Vous devez également spécifier le fichier de données qui contient les données à traiter.

Pour utiliser efficacement la macro, vous devez connaître la syntaxe SAS de base. Vous devez d'abord inclure la macro dans le fichier de programme où vous l'utiliserez. Si vous travaillez en mode différé, vous devez exécuter la macro pour chaque lot. Si vous utilisez le SAS en mode interactif, vous devez exécuter la macro une fois au début de la session; elle restera active pendant toute la session. Si vous mettez fin à la session ou si vous la relancez plus tard, vous devez exécuter la macro de nouveau. Une fois la macro comprise dans une session donnée, évitez d'utiliser la chaîne « %STATTOOL » dans ce programme, ce qui aurait pour effet d'exécuter la macro.

Cette macro comporte les paramètres suivants :

WGT Poids d'échantillonnage à utiliser dans l'analyse, généralement POPWT.

RWGT Racine des variables spécifiant les poids de rééchantillonnage compris dans les fichiers de données, généralement REPLIC01 à REPLIC30. Les poids de rééchantillonnage doivent être spécifiés comme suit : « REPLIC ».

NREP Indication du nombre de poids de rééchantillonnage établis dans le fichier de données. Si vous travaillez avec les données d'un seul pays, vous devez appliquer à l'argument NREP autant de poids de rééchantillonnage qu'il en faut dans le pays. (Si vous travaillez avec les données de plus d'un pays, assurez-vous que tous les ensembles de données comportent le même nombre de poids de rééchantillonnage.)

NPV Nombre de valeurs plausibles à utiliser pour l'analyse. En général, vous voudrez utiliser les cinq valeurs plausibles mais, dans certains cas, vous pouvez en utiliser moins.

STUDY Nom de l'étude (ELCA).

CNTRYNO Identificateur du pays.

INFILE Nom du fichier de données contenant les variables nécessaires à l'analyse. (Si l'emplacement du chemin est compris dans le nom du fichier, le nom du fichier doit figurer entre guillemets.) Inclure uniquement les cas d'intérêt pour l'analyse (par ex., il faut exclure les répondants pour lesquels il manque des variables avant d'exécuter la macro).

METHOD Indication de la statistique à produire. « METHOD = mean » calcule les moyennes de la variable d'intérêt. Vous pouvez également spécifier « crosstabs » pour les totalisations croisées, « perc » pour les centiles, « diff » pour les écarts (différences), « popest » pour les chiffres de population, « reg » pour la régression linéaire type, « logistic » pour la régression logistique et « multinomial » pour la régression logistique multinomiale.

DVAR Variable pour laquelle on veut calculer des moyennes. Une seule variable doit être inscrite ici. Inscrivez le nom de la variable ou seulement la racine si la variable d'intérêt est tirée d'un ensemble de valeurs plausibles.

DVARPV Paramètre indiquant si la variable DVAR est tirée ou non d'un ensemble de valeurs plausibles. Ce paramètre prend la valeur 1 si la variable d'intérêt est tirée d'un ensemble de valeurs plausibles et 0 dans le cas contraire.

BYVAR Liste des variables à utiliser pour classer les répondants dans le fichier de données. Il peut s'agir d'une seule variable ou d'une liste de variables. Ce paramètre définit les sous-groupes pour lesquels on demande des moyennes de la variable DVAR.

BYVARPV Paramètre indiquant si la ou les variables BYVAR sont tirées ou non d'un ensemble de valeurs plausibles. Ce paramètre prend la valeur 1 si la variable d'intérêt est tirée d'un ensemble de valeurs plausibles et 0 dans le cas contraire.

Outre ces paramètres, on peut en utiliser trois autres : CRITER1, CRITER2 et CRITER3. Chacun contient un seul énoncé de programmation SAS.

La manière la plus simple d'exécuter la macro consiste à utiliser la notation SAS courante servant à appeler des macros. Il s'agit d'inscrire le nom de la macro, suivi par la liste correspondante d'arguments nécessaires à l'analyse, chacun étant séparé par une virgule. Par exemple, si vous exécutez la macro en utilisant le code suivant :

```
%include "c:\IALSS\stattool.sas";
```

```
%stattool (wgt = popwt,  
           rwgt = replic,  
           nrep = 30,  
           npv = 5,  
           study = ALL,  
           method = mean,  
           infile = in,  
           dvar = prose,  
           dvarpv = 1,  
           byvar = gendaa2 age3,  
           byvarpv = 0 0);
```

Elle calcule le rendement moyen en compréhension de textes suivis en utilisant les cinq ensembles de valeurs plausibles et l'erreur-type correspondante, à l'intérieur de chaque groupe défini par la combinaison des catégories sexe et âge, en prenant la variable POPWT comme poids d'échantillonnage. Les données sont lues à partir du fichier système TEMP.

Le fichier qui contient ces résultats est appelé FINALB et est conservé dans le répertoire par défaut utilisé par le SAS. Il existe également un fichier HTML appelé FINALB, conservé dans le lecteur C de votre ordinateur, dans le répertoire TEMP. On peut facilement accéder à ce fichier en utilisant EXCEL à partir de MICROSOFT. Les variables contenues dans ce fichier sont les suivantes :

Variables de classification

Chacune des variables de classification est conservée dans le fichier résultant. Il y a une seule occurrence par combinaison spécifique de catégories de variables de classification.

ESTIMATE

Variable contenant les moyennes de la variable DVAR pour les groupes définis par les combinaisons correspondantes de catégories de variables de classification.

STANDARD ERROR

Variable contenant les erreurs-types des valeurs de l'estimation calculées selon la méthode jackknife, y compris les composantes échantillonnage et imputation.

PROB > |T|

Variable donnant la probabilité qu'une statistique de Student soit plus grande que la valeur absolue de l'estimation observée, à l'intérieur de la combinaison spécifiée des catégories définies à l'origine par les groupes.

En voici deux exemples :

```
libname in "C:\IALSS\data";
data in;set in.IALSSdata;run;

%stattool(wgt      =  popwt,
          rwt      =  replic,
          nrep     =  30,
          =  5,
          ALL,
          npv      =  study =
          method   =  mean,
          infile   =  in,
          dvar     =  prose,
          dvarpv   =  1,
          byvar    =  gendaa2 age3,
          byvarpv  =  0 0);
```

Étude : ELCA :,

Moyennes estimatives en compréhension de textes suivis par domaine et gendaa2 age3, en fonction de cinq ensembles de valeurs plausibles et 29 degrés de liberté

Abstraction faite du domaine

Obs	Domain	GENDAA2	AGE3	estimate	Standard Error	Prob > T
1	ELCA	1	1	236,764	1,81853	0
2	ELCA	1	2	231,754	1,87374	0
3	ELCA	1	3	197,439	4,25769	0
4	ELCA	2	1	236,945	2,05748	0
5	ELCA	2	2	225,062	1,08269	0
6	ELCA	2	3	192,060	2,77765	0

Dans cet exemple, XPROSE1 à XPROSE5 servent de variables de classification. Nous estimons le revenu personnel moyen par niveau de compréhension de textes suivis.

```
libname in "C:\IALSS\data";

data in;set in.ialssdata;run;

%stattool(wgt      =  popwt,
          rwgt     =  replic,
          nrep     =  30,
          npv      =  5,
          study    =  ALL,
          method   =  mean,
          infile   =  in,
          dvar     =  D43,
          dvarpv   =  0,
          byvar    =  XPROSE,
          byvarpv  =  1,
          criter1  =  if d43 < 99999997 );
```

Étude : ELCA : si d43 < 99999997,,

Moyennes estimatives pour d43 par domaine et xprose, en fonction de cinq ensembles de valeurs plausibles et 29 degrés de liberté

Abstraction faite du domaine

Obs	Domain	xprose	estimate	Standard Error	Prob > T
1	ELCA	1	39799,22	6167,01	0,000000463
2	ELCA	2	66658,33	9378,51	0,000000081
3	ELCA	3	103063,21	27265,80	0,000724651
4	ELCA	4	172247,60	60242,53	0,007788203

On peut modifier le contenu de la « méthode » pour produire d'autres statistiques en inscrivant :

perc = pour produire des centiles,

diff = pour produire des écarts (différences),

popest = pour produire des chiffres de population,

reg = pour produire une régression linéaire type,

logistic = pour produire une régression logistique,

multinomial = pour produire un coefficient de régression logistique multinomiale.

Tous les autres paramètres restent identiques à ceux des exemples illustrés ci-dessus, les erreurs-types liées à l'échantillonnage et l'erreur de test étant dûment calculées pour chaque mesure.

8.2 Erreurs non liées à l'échantillonnage

Lorsqu'elles sont réparties sur un grand nombre d'observations, les erreurs non liées à l'échantillonnage qui surviennent au hasard ont peu d'incidence sur les estimations tirées de l'enquête. Toutefois, les erreurs qui surviennent systématiquement entraînent des distorsions dans ces estimations. Les responsables de l'enquête ont donc consacré beaucoup de temps et d'efforts à réduire les erreurs non liées à l'échantillonnage. Pour vérifier la qualité des données, on a pris des mesures d'assurance de la qualité à chaque étape du cycle de collecte et de traitement des données. Ces mesures comprenaient le recours à des intervieweurs hautement qualifiés, la formation intensive des intervieweurs à l'égard des méthodes d'enquête et du questionnaire, l'observation des intervieweurs pour déceler des problèmes de conception du questionnaire ou des erreurs d'interprétation des directives, une procédure visant à réduire au minimum les erreurs de saisie des données, ainsi que des contrôles du codage et de la révision pour vérifier la logique de traitement.

Malgré ces efforts, aucune enquête n'est à l'abri d'une erreur non liée à l'échantillonnage. Voici un aperçu des principales sources de ce type d'erreur et de son incidence sur les données de l'ELCA.

8.2.1 Base de sondage

L'utilisation du Recensement de 2001 a permis d'assurer que la base de sondage de l'ELCA était aussi inclusive que possible et que les moindres exclusions seraient effectivement calculées dans le plan d'enquête global.

8.2.2 Non-réponse

L'incidence de la non-réponse sur les résultats d'une enquête constitue une importante source d'erreurs non liées à l'échantillonnage. Il peut s'agir d'une non-réponse partielle (défaut de répondre à une ou à quelques questions seulement) ou d'une non-réponse totale.

Il y a non-réponse totale lorsque l'intervieweur ne peut communiquer avec le répondant, qu'aucun membre du ménage ne peut fournir les renseignements demandés ou que le répondant refuse de participer à l'enquête. Le taux national de non-réponse à l'ELCA se situait autour de 34 %. Toutefois, l'analyse des caractéristiques des non-répondants à l'ELCA donne à entendre que ces derniers ne se concentrent pas dans un groupe donné (ainsi, la non-réponse semble avoir été aléatoire). On a réglé le problème de la non-réponse totale en rajustant le poids des ménages ayant répondu à l'enquête, compensant ainsi pour ceux qui n'ont pas répondu.

Dans la plupart des cas, il y a non-réponse partielle à une enquête lorsque le répondant ne comprend pas ou interprète mal une question, refuse de répondre à une question ou ne se rappelle pas les renseignements demandés. En général, la fréquence des non-réponses partielles à l'ELCA était faible.

8.2.3 Erreur dans la réponse

Il convient de mentionner d'autres sources possibles d'erreur non liée à l'échantillonnage qui sont propres à l'ELCA. D'abord, certains répondants ont pu trouver le test intimidant, ce qui peut avoir eu une incidence négative sur leur rendement. Contrairement aux études « habituelles », il existe de « bonnes » et de « mauvaises » réponses aux items de test de l'ELCA. De plus, pour de nombreux répondants, c'était la première fois qu'ils subissaient un test depuis bon nombre d'années. En outre, même si les intervieweurs n'imposaient pas de limite de temps pour répondre aux questions, le fait que quelqu'un observait et attendait la réponse a peut-être imposé une contrainte de temps involontaire. Par conséquent, même si les items ont été choisis avec soin pour correspondre étroitement à des tâches courantes, il est possible que les réponses au test ne révèlent pas pleinement les compétences en littératie des répondants en raison du contexte propre à un test. Enfin, comme il s'agissait d'un test, les répondants devaient accomplir les activités de manière complètement indépendante alors que, dans la réalité, on a souvent l'occasion de parcourir des textes imprimés avec des parents, des amis et des collègues. Il est donc possible que les compétences mesurées par l'enquête ne reflètent pas l'étendue complète des aptitudes de certains répondants dans un contexte ordinaire.

8.2.4 Notation

Dans le cadre de l'ELCA, une autre source possible d'erreur non liée à l'échantillonnage a trait à la notation des items du test, notamment ceux qui ont été notés sur une échelle (par exemple, les items qui exigeaient que les répondants rédigent une réponse). Pour réduire au minimum la fréquence des erreurs de notation, on a déployé des efforts particuliers, dont la centralisation de la notation et la vérification par échantillonnage.

9.0 Dispositions 'enregistrement et fréquences univariés

Veillez referer au document «ELCA_manueldecode_F.pdf» pour les dispositions d'enregistrement et les fréquences univariés de l'ELCA.

10.0 Principaux participants au projet

Direction et coordination de l'étude internationale

M. T. Scott Murray

Directeur de l'étude internationale ELCA, Statistique Canada, Ottawa

M. Yvan Clermont

Coordonnateur de l'étude internationale ELCA, première vague de pays, Statistique Canada, Ottawa

Mme. Sylvie Grenier

Coordonnatrice de l'étude internationale ELCA, deuxième vague de pays, Statistique Canada, Ottawa

M. Patrick Werquin

Coordonnateur de l'étude internationale ELCA, OCDE, Paris

Notation et échelonnage de l'étude internationale

M. Irwin Kirsch

Educational Testing Service, Princeton

M. Kentaro Yamamoto

Educational Testing Service, Princeton

Mme. Minh-Wei Wang

Educational Testing Service, Princeton

Mme Julie Eastland

Educational Testing Service, Princeton

Responsables nationaux de l'étude

Bermudes	M. Crispin Boney <i>Statistics Department, Government of Bermuda, Hamilton</i>
Canada	M. Jean Pignal <i>Statistique Canada, Ottawa</i>
Hongrie	Janos Wiedermann, <i>Budapest</i>
Italie	Mme. Vittoria Gallina <i>Istituto Nazionale per la Valutazione del Sistema dell'Istruzione, Frascati</i>
Norvège	M. Egil Gabrielsen <i>Centre for Reading Research, Stavanger</i>
Nouvelle-Zélande	Paul Satherley <i>Ministère de l'éducation, Wellington</i>
Nuevo León, (Mexique)	M. Edmundo Guajardo Garza <i>Ministerio de Educación, Monterrey</i>
Pays-Bas	Willem Houtkoop <i>Max Goote Expert Center for educational research, Amsterdam</i>
Suisse	M. Philippe Hertig <i>Office fédéral de la statistique, Neuchâtel</i> M. Philipp Notter <i>University of Zürich, Zürich</i>
États-Unis	Mme. Mariann Lemke <i>National Center for Education Statistics, Washington</i> M. Eugene Owen <i>National Center for Education Statistics, Washington</i>

Experts du domaine et collaborateurs

Textes suivis et textes schématiques

M. Irwin Kirsch
Educational Testing Service, Princeton

M. Kentaro Yamamoto
Educational Testing Service, Princeton

M. Julie Eastland
Educational Testing Service, Princeton

M. Stan Jones
Atlantic Heath Promotion Research Center, Yarmouth

Numératie

M. Iddo Gal
University of Haifa, Haifa

Mme. Mieke van Groenestijn
Utrecht University of Professional Education, Utrecht

Mme. Myrna Manly
El Camino College, Palos Verdes

Mme. Mary Jane Schmitt
TERC, Cambridge

M. Dave Tout
Language Australia, Melbourne

M. Yvan Clermont
Statistique Canada, Ottawa

M. Stan Jones
Atlantic Heath Promotion Research Center, Yarmouth

Experts du domaine et collaborateurs

Résolution de problèmes

M. Eckhard Klieme
German Institute for International Educational Research, Frankfurt

M. Jean-Paul Reeß
LIFE Research and Consult, Bonn

Mme. Anouk Zabai
LIFE Research and Consult, Bonn

Questionnaire de référence

Mme. Lynn Barr-Telford
Statistique Canada, Ottawa

M. Stan Jones
Atlantic Heath Promotion Research Center, Yarmouth

M. Trevor Williams
WESTAT, Rockville

Équipe d'enquête, analystes et équipe de production

Mme. Danielle Baum
Statistique Canada, Ottawa

M. Richard Desjardins
Statistique Canada, Ottawa

Mme. Sylvie Grenier
Statistique Canada, Ottawa

M. John Leung
Statistique Canada, Ottawa

Mme. Carrie Munroe
Statistique Canada, Ottawa

M. Owen Power
Statistique Canada, Ottawa

Auteurs

M. Yvan Clermont
Statistique Canada, Ottawa

M. Richard Desjardins (rédacteur en chef)
Statistique Canada, Ottawa

Mme. Urvashi Dhawan-Biswal
Ressources humaines et Développement des compétences Canada, Ottawa

Mme. Lauren Dong
Statistique Canada, Ottawa

M. Irwin Kirsch
Educational Testing Service, Princeton

Mme. Carrie Munroe
Statistique Canada, Ottawa

M. T. Scott Murray (rédacteur)
Statistique Canada, Ottawa

M. Owen Power
Statistique Canada, Ottawa

Mme. Isabelle Recotillet
Centre d'études et de recherches sur les qualifications, Marseilles

M. Kjell Rubenson
Université de Colombie-Britannique, Vancouver

M. George Sciadas
Statistique Canada, Ottawa

M. Albert Tuijnman (rédacteur)
European Investment Bank, Luxembourg

M. Ben Veenhof
Statistique Canada, Ottawa

M. Matthias Von Davier
Educational Testing Service, Princeton

M. Patrick Werquin
Organisation de coopération et de développement économiques, Paris

M. J. Douglas Willms
Université du Nouveau-Brunswick, Fredericton