Functional and Elective Illiteracy in Canada*

Frank L. Graves
ekos research associates inc.,
ottawa, ont.

Brian Kinsley
department of communications,
government of canada

Using nationally representative data from the 1978 Canadian Readership Survey, the authors model a range of reading activity. They demonstrate empirical problems in associating illliteracy with education at less than a Grade 9 level. The survey data indicate that a majority of Canadians identified as illiterate in this manner do in fact read. The paper identifies and describes the truly functionally illiterate, (i.e., those who report that they cannot read). This discussion turns to the problem of explaining the far more numerous elective illiterates (i.e., those who can read but do not). Using explanatory statistical models based on log-linear and logit-techniques, the article attempts to explain (in a statistical sense) elective illiteracy in Canada. In addition to confirming the expected bivariate associations between reading activity and age, occupation, education, language, and settlement type, this research uncovers three statistically and substantively significant higher order specification effects involving (1) reading, education and language, (2) reading, age and settlement type, and (3) reading, settlement type, and language.

Les auteurs dégagent un éventail de niveaux de lecture à partir des doneés recueillies auprès d'un échantillin représentatif de la population canadienne dans le cadre de l'Enquête Canadienne sur la Lecture de 1978. Ils mettent en évidence les problèmes pratiques qui se posent lorsqu'on considère comme analphabète toute personne dont le niveau de scolarité est inférieur à la 9e année. Les données de l'enquête indiquent en effet que la majorité des Canadiens identifiés comme analphabètes à partir de ce critère sont en fait capables de lire. L'article identifie et décrit alors les véritables analphabètes fonctionnels (i.e., ceux qui répondent qu'ils ne savent pas lire). Puis, ayant disposé du problème de l'analphabétisme fonctionnel, les auteurs poursuivent leur analyse en s'intéressant à une population beaucoup plus nombreuse: celle des analphabètes par choix

* This research was supported by the Research and Statistics Directorate, Department of Communications, Government of Canada. The paper reflects the views of the authors only and does not necessarily represent the views of the Department of Communications. The authors wish to acknowledge the contribution of our colleagues at Ekos Research and Carleton University. Professor Hugh McRoberts provided particularly valuable advice. The comments of our reviewers are also gratefully acknowledged. In particular, Professor J. Dale Burnett gave us extremely thorough comments which substantially improved our paper. Naturally the authors alone are responsible for any errors in this paper.

(i.e., ceux qui savent lire mais ne lisent pas). A l'aide de modèles statistiques explicatifs élaborés notamment à partir de techniques d'analyse linéaire par transformation logarithorique, cette étude évalue différentes façons d'expliquer et de prédire cette forme d'analphabétisme. En plus de confirmer les effets simples d'interaction prévus entre le niveau de lecture d'une part, et l'âge, l'occupation, le niveau de scolarité et le langue parlée d'autre part, cette recherche fait ressortir des effets spécifiques statistiquement significatifs et importants, impliquant (1) la lecture, le niveau de scolarité et la langue parlée, (2) la lecture, le niveau de scolarité et le type de localité de résidence, et (3) la lecture, l'âge et la localité de résidence.

INTRODUCTION: AIMS AND STRATEGY

In an increasingly technocratic society, reading and writing skills are basic survival tools. Conversely, illiteracy may be viewed as a nearly insurmountable barrier to social mobility and a serious impediment to adaptation to everyday life in post-industrial society. Whether an individual does not read because of incapacity or by choice, this individual necessarily loses much of the information vital for effective operation in our society.

Although reading and writing are companion literacy skills, the data used in this study only permit us to examine reading ability and behaviour. Given this constraint, the present study has three basic aims. First, we will demonstrate empirically the inadequacy of conventional operational definitions of illiteracy which rely on measures of educational attainment, most specifically those based on Grade 8 or less, as the criterion for functional illiteracy. Our data indicate that most individuals with an education at less than the Grade 9 level, do, in fact, engage in some reading activity. Second, we identify and briefly describe the truly functionally illiterate, i.e., those who say they cannot read and have not read as a leisure time activity. Finally, we distinguish a category of nonreaders who do not read despite the fact that they claim to be able to read. We then apply log-linear and logit modeling techniques to develop a predictive model to explain and predict these "elective" illiterates. These are individuals who indicated on the survey that they had not read anything in the last week and had not read a magazine or book in the six weeks prior to the survey date, i.e., since 1 January 1978. We believe that this latter exercise is particularly valuable since the logit models isolate substantively and statistically significant higher order effects which are neither intuitively obvious nor discernible using other techniques.

Our analysis is based on the categories which follow from the cross classification of reading ability (yes/no) and reading behaviour (yes/no). The four logical categories evident in Table 1 constitute a simple literacy typology. Following a brief descriptive profile of the "functional illiterate" we will turn our attention to the problem of explaining elective illiteracy (i.e., why do Canadians who have the ability choose not to read?).

TABLE 1
Literacy Typology Cross-Classification of Reading Ability and Reading Behaviour

	Reading ability	
Reading behaviour	*Can read*	*Can't read*
Reads	Literate	Structural zero cell
Does not read	Elective illiterate	Functional illiterate

THE DATA

The analysis is based on data drawn from the Survey of Leisure Activities and Reading Habits commissioned by the Secretary of State Department and conducted by Statistics Canada in February, 1978. The survey was undertaken in conjunction with the monthly Labour Force Survey, which permitted a link between the leisure and reading habits data and sociodemographic characteristics and labour force participation. The sample was a multistage, stratified, random sample representative of the noninstitutionalized Canadian population 15 and over, excluding the Yukon and Northwest Territories. The survey was self-completed, and it was dropped off at the respondents' homes and picked up after completion. Of the 20,000 individuals receiving questionnaires, nearly 17,000 or about 80% completed them. The distribution of the sample according to education level of respondent is set out in Table 2. It is important to note that nonreaders could have received assistance in filling out the survey from literate members of the household. In fact, they would be likely to do so given that the survey was an official government survey. It is also noteworthy that about equal proportions of respondents from each education category did not complete the questionnaire. Hence, there was no discernible bias toward the better educated.

Concerning the instrument itself, the data available for analysis included six items on labour force participation, 18 on sociodemographic characteristics, 24 on general leisure and cultural activities and 25 on reading and library use. Many questions were multi-dimensional and comprise some 240 variables in all.[1]

DO THE ILLITERATE READ: EMPIRICAL AND CONCEPTUAL PROBLEMS OF IDENTIFYING THE ILLITERATE[2]

The O.E.D. defines an illiterate as one who is "ignorant of letters" or "unable to read."

TABLE 2
Distribution of Respondents, and Nonrespondents by Level of Education, Canada 1978 (%)

	Type of respondent	
Level of education	*Respondent*	*Nonrespondent*
0–8 years	81	19
9–10 years	80	20
11–12 years	80	20
13 years	80	20
Some postsecondary	80	20
Certificate or diploma	83	17
Degree	81	19
Total	81	19

The UNESCO Expert Committee on Standardization of Educational Statistics developed this definition further stating:

A person is literate when he has acquired the essential knowledge and skills which enable him to engage in all those activities in which literacy is required for effective functioning in his group and community, and whose attainment in reading, writing and math make it possible to continue to use these skills toward his own and the community development. (Harman, 1970)

However, these definitions do not permit one to easily identify those who are illiterate. There are two approaches that have been taken. Sets of instruments have been developed which measure how well individuals understand what they are reading and whether they can follow instructions based on this material (e.g., Evans et al., 1977). This method, while undoubtedly valid, is cumbersome and impractical for large-scale studies, cross-national comparisons or comparisons with previous periods in time.

As an alternative, governments, agencies and researchers have employed an indirect measure of literacy. Literates are defined as those who have achieved a certain level of education such as Grade 5 or Grade 8. Although educators and policy makers do not treat this definition uncritically, it does appear in many important research and policy contexts. For example, Statistics Canada used Grade 5 as the cut-off level for its analysis of the 1961 Census data and still maintains this definition. The Canadian Association for Adult Education defines the literate as one who has at least nine years of formal education. It is clear, then, that the definition of literacy varies according to specific requirements. This definition (Grade 9 or more) was used in a major study by Thomas (1976) in which she compared literacy rates across Canada and concluded that the illiteracy rate had dropped significantly. Indeed, by this definition it has declined from

TABLE 3
Type of Reading Activity by Educational Level, Canada 1978 (%)*

Educational level	*Book readers*	*Reader–magazines or newspapers, but not books*	*Non-readers*	*Total*
0–8 years	32.8(1604)	45.4(1908)	21.8(1487)	100 (4459)
9–10 years	54.0 (337)	36.8(1182)	9.2(2200)	100 (3719)
11 years	66.7 (141)	27.4 (571)	5.9(1739)	100 (2451)
12–13 years	71.3 (145)	24.9 (967)	3.8(3793)	100 (4905)
Some postsecondary	76.6 (35)	20.4 (244)	3.0(1067)	100 (1346)
Certificate or diploma	76.4 (80)	21.0 (341)	2.6(1288)	100 (1679)
Degree	85.7 (8)	13.6 (135)	0.8(1016)	100 (1159)
Totals	60.0(2350)	30.8(5348)	9.2 (139)	100(16704)

*Readers were defined as those who have read at all for leisure within the last year.
NOTE: *N* in parentheses

41% in 1961 to 28% in 1976 for individuals 15 and over who are not in school. But is this an adequate definition of literacy?

If we turn to the results of the 1978 Survey of Reading Habits, we can immediately see the problem with using Grade 8 or less to identify the illiterate population. As Table 3 indicates, over 78% of those with less than Grade 9 claimed to have read as a leisure activity. This is a rather high level of reading activity for a group defined as "illiterate." Of course it may be that the "quality" of reading activity for this group was quite low. Table 3 does show that this group prefers less demanding nonbook reading but 33% read books as a leisure time activity.

THE FUNCTIONAL ILLITERATE

There are an estimated 88,000 individuals 15 and over who have indicated an inability to read any language and are nonreaders. Over 90% have less than a Grade 8 education, and many may not have any formal education at all. Clearly, while those with less than Grade 9 education are not necessarily illiterate or nonreaders, those who cannot read lack at least a Grade 9 level of education.

Where are those who cannot read located in Canada? Projecting from our representative sample, we estimate that 18,000 reside in the Atlantic provinces, 26,000 in Québec, 29,000 in Ontario, and 15,000 in Western Canada. Table 4 indicates, for example, that while 20% of Canada's functionally illiterate reside in the Atlantic region, this region contains only 9% of Canada's population. Thus the table shows that the Atlantic provinces and rural areas have a proportionately higher concentration of illiterates than other regions or urban areas of the country.

TABLE 4
Estimated Population 15 and Over who Could not Read by Region and Settlement Size, Canada, 1978 (%)

Location	*Distribution of people who cannot read*	*Population distribution*
Region		
Atlantic	20(18,000)	9
Quebec	30(26,000)	27
Ontario	32(29,000)	37
Western Canada	18(15,000)	27
Settlement size		
Over 100,000	33(29,000)	57
Under 100,000 nonrural	32(28,000)	25
Rural	35(31,000)	18
Total	100(88,000)	100

NOTE: *N* in parentheses

MODELING READERS AND ELECTIVE ILLITERATES

Introduction

We can now turn to the principal task and, using log-linear and logit modeling, conduct explanatory modeling to determine which variables and combination of variables "cause" a respondent to fall into the elective nonreader category. Since our data are nationally representative, we have the unique opportunity to apply these results to the system of factors producing elective illiteracy in modern Canadian society.

Several factors limit the types of statistical analyses we can conduct. The level of measurement of our dependent variable (as well as many of our independent variables) is ordinal at best, and hence conventional parametric techniques such as least squares regression are inappropriate. We could consider dummy-dependent regression; however, the split on our dependent variable is too extreme.[3]

Rather than resort to cruder, nonparametric techniques, we utilized powerful log-linear and logit methods which permit the quantitative analysis of qualitative data. Given the nature of most social data, we firmly believe that these newer techniques represent the most important recent development in quantitative social research methodology.†

† Editor's Note: See, for example, J. Dale Burnett, Loglinear analysis: A new tool for educational researchers. *Canadian Journal of Education*, 1983, *8*(2), 139–154.

The mathematics underlying Goodman's system are rather complex but the results of applying this approach are intuitively appealing and easily understood. To simplify, log-linear and logit models predict the logged odds that one will fall in a specified cell of a contingency table (e.g., the "illiterate" vs. "literate" cells) as a function of belonging to certain other cells of the table. Through a hierarchical fitting process one develops a statistical model which identifies the significant effects producing variation in the table (defined in terms of a likelihood ratio chi square). The objective is to develop a model to account for the bulk of the chi square in a crossclassified table with the fewest "effects" (i.e., parameters). By way of analogy to regression, one might say that in log-linear modeling we predict cell frequencies on the basis of other cell frequencies, whereas in a regression model, we predict scores on the basis of other scores.[4]

Dependent and Independent Variables

The variable which we are interested in "explaining" is reading (R) – i.e. did the respondent engage in any reading in the six weeks prior to the survey. Recall that we have removed from this analysis those who did not read because they were not able to read[5]. Hence we are only trying to predict whether or not someone elects not to read even if that individual is capable of reading. Of the 14,953 cases remaining in the analysis, 1,738 or about 11.6% fall into the elective illiterate category.

The independent variables include: education (E), age (A), settlement size (S), language (L), and occupational group (O). The variables and the response categories associated with them are defined in Table 5.

It should be noted that the variables were dichotomized in order to take full advantage of the regression-like logit models. In collapsing categories to create the dichotomous variables necessary for a logit model, one can lose information and in some cases produce misleading results (cf., Gillespie, 1977, pp. 108–109; Knoke, 1975). In order to avoid this potential problem, we conducted a complete log-linear analysis of the tables, based on more complex polytomous variables, before we proceeded to the logit analysis of the dichotomous variables. We also tested different models involving alternate cutting points for the dichotomous variables. Although there is some information loss in collapsing our data in this way, we can confidently state that this simplification of the table produces no misleading results.[6]

The Log-linear Model

The aim of log-linear modeling is to develop the most parsimonious model which best predicts the raw data.[7] The present study employed Goodman's

TABLE 5
List of Variables Used in the Analysis

Variable	*Symbol*	*1st Category*	*2nd Category*
Predictor variables			
Education	(E)	Less than Grade 9	Greater than or equal to Grade 9
Language	(L)	Anglophone	Francophone
Settlement type	(S)	Metropolitan cities over 100,000	Nonmetropolitan All other areas
Age	(A)	Less than 35 (younger)	35 or older (older)
Occupation	(O)	White-collar	Nonwhite-collar
Dependent variable			
reading activity	(R)	Did not read	Did read

ECTA[8] software. We also fitted the same model using LOGLIN.[9] Although the *p*-values and parameters were slightly different, the conclusions and fit were identical.

Typically, a log-linear analysis begins with a baseline model of mutual independence. However, since our concern in this exercise lies principally with the sole dependent variable reading activity (R) we are not particularly interested in explaining the χ^2 due to the mutual dependencies of the five independent variables. Hence the appropriate baseline model is one which takes into account the intercorrelations amongst the predictor variables (i.e., E, S, L, A, O). The $(ESLAO)$ term is included in order to take the observed associations among predictor variables for granted, in the same sense as we do in multiple regression.

After fitting numerous possible models to the original table we found that we could not reject a model based on three specification effects and one associative effect. An associative effect involves the relation of two terms, wheras a specification effect involves the interaction of three or more terms. In Davis's (1974) terms the model would be: (REL) (RAS) (RSL) (RO). By virtue of the hierarchical principle, all of the lower order effects contained in these higher order effects are logically implied. Hence (REL) implies that the (RE) (EL) (RL) (R) (E) (L) effects all must be included in the model. Recent literature suggests that nonhierarchical, or nested models are possible but in these types of models the interpretation of parameters becomes problematic. Equation 1 summarises this model.

$$FIJKLM = \theta + \lambda^{RO} + \lambda^{REL} + \lambda^{RAS} + \lambda^{RSL} \qquad (1)$$

The θ, from the Model 1, is equivalent to a "grand mean effect" and the

TABLE 6

Original Raw Cell Frequencies and Predicted Cell Frequencies

	Metropolitan (100,000 population)				*Nonmetropolitan (100,000 population)*				
	35 Years		*35 Years*		*35 Years*		*35 Years*		
	<Grade 9	*≥Grade 9*	*<Grade 9*	*≥Grade 9*	*<Grade 9*	*≥Grade 9*	*<Grade 9*	*≥Grade 9*	*Totals*
White-collar									
English									
Nonreader	8(7.1)	50(42.5)	26(24.6)	27(28.3)	28(29.8)	88(95.8)	52(57.5)	45(35.5)	324
Reader	33(32.7)	894(890.2)	111(112.0)	578(584.7)	56(69.1)	1,068(1,009.7)	216(213.2)	570(598.5)	3,526
French									
Nonreader	2(2.7)	15(15.9)	9(8.1)	5(6.4)	7(9.0)	30(28.4)	17(15.0)	4(6.4)	89
Reader	9(9.5)	150(175.2)	46(28.2)	75(69.7)	18(23.2)	215(230.1)	66(62.2)	86(82.7)	665
Nonwhite-collar									
English									
Nonreader	23(18.9)	55(60.9)	97(97.0)	55(60.7)	129(121.5)	215(210.6)	345(350.6)	116(116.7)	1,035
Reader	77(73.9)	1,093(1,084.4)	362(378.2)	1,074(1,065.8)	238(239.0)	1,817(1,885.8)	1,129(1,103.9)	1,698(1,672.8)	7,488
French									
Nonreader	7(7.2)	18(22.8)	35(32.2)	18(13.7)	34(36.5)	63(62.4)	91(91.3)	24(21.0)	290
Reader	21(21.5)	224(213.5)	84(95.2)	131(127.1)	92(80.4)	463(429.7)	306(321.8)	215(231.0)	1,536
Totals	180	2,499	770	1,963	602	3,959	2,222	2,758	4,953

NOTE: Predicted or fitted cell frequencies in parentheses

TABLE 7
Selected List of Alternative Models Tested

Model	*L.R.*^{2}	*P-Value*
(1) Mutual independence (no effects)	813.6	0
(2) Conservative baseline (controling for interindependent variable interactions) [EASLO] [R]	356.5	0
(3) Fitted Model [REL] [RAS] [RSL] [RO]	49.7	.117
(4) [REL] [RAS] [RO]	61.6	.020
(5) [REL] [RAS] [RSL]	56.7	.040
(6) [RE] [RE] [RSL] [RO]	58.2	.031
(7) [REL] [RA] [RSL] [RO]	66.8	.005

λ's represent the effect on the logged odds of being in a given category of the table as a function of the variable(s) specified in the lambda's superscripts. For example, *RE* would represent the departure from the grand mean effect as a function of being in the first categories of both the reading and education variables.

Table 6 shows the level of fit we can achieve between our actual and estimated values using this model. Although the quality of fit is fairly obvious from examining these results, additional evidence confirms this. For example, the *p*-value associated with testing this particular model (hypothesis) is .12 which means that we cannot reject the model at conventional significance levels. Furthermore, an index of dissimilarity of 1.8 indicates that less than 2% of the cases in the table would have to be rearranged in order to achieve a perfect fit. Our baseline model has a χ^2 likelihood ratio of 356.5 and under the proposed model only 49.7 χ^2 remain with 39 degrees of freedom. In other words there is about a 94% reduction in error. This figure is analogous to an R^2 of .94.

To compare the performance of various other hypothetical models and effects, Table 7 was prepared to summarize the likelihood ratio χ^2 and *p*-values associated with testing other potential models. Since we can safely decompose the likelihood ratio χ^2 with this aproach (this is not advisable with the more conventional Pearsonian least-squares chi square) we can conclude whether the effect of dropping or adding a term to the model produces a significant drop in chi square. For example, removing the reading, settlement size, and language specification effect increases the χ^2 by 12. As one can see from examining the *p*-values associated with various alternate models, the actual fitted model which we have selected is the only one which permits us to reject the hypothesis that the model does not fit at a greater than .05 significance level.

The most promising alternate model is the one which removes the

TABLE 8
Predicted Odds Being a Reader or Non-Reader for Various Combinations of Explanatory Variables

Settlement size	*Age*	*Education*	*Language*	*Occupation*	*Nonliterate odds*	*Literate odds*	*Rank*
>100K	<35	≥GR9	E	WC	.048	20.95	1
>100K	≥35	≥GR9	E	WC	.048	20.66	2
>100K	<35	≥GR9	E	NWC	.056	17.81	3
>100K	≥35	≥GR9	E	NWC	.057	17.56	4
<100K	≥35	≥GR9	E	WC	.059	16.86	5
<100K	≥35	≥GR9	E	NWC	.070	14.33	6
<100K	≥35	≥GR9	F	WC	.077	12.92	7
>100K	<35	≥GR9	F	WC	.091	11.02	8
<100K	≥35	≥GR9	F	NWC	.091	11.00	9
>100K	≥35	≥GR9	F	WC	.092	10.90	10
>100K	<35	≥GR9	E	WC	.095	10.54	11
<100K	≥35	≥GR9	F	NWC	.107	9.36	12
<100K	<35	≥GR9	F	NWC	.108	9.28	13
<100K	<35	≥GR9	E	NWC	.112	8.95	14
<100K	≥35	≥GR9	F	WC	.123	8.10	15
<100K	<35	≥GR9	F	NWC	.145	6.89	16
<100K	<35	<GR9	E	WC	.217	4.61	17
>100K	≥35	<GR9	E	WC	.220	4.55	18
<100K	≥35	<GR9	F	WC	.241	4.15	19
>100K	<35	<GR9	E	NWC	.256	3.91	20
>100K	≥35	<GR9	E	NWC	.256	3.90	21
<100K	≥35	<GR9	E	WC	.270	3.71	22
>100K	<35	<GR9	F	WC	.284	3.52	23
<100K	≥35	<GR9	F	NWC	.284	3.52	24
>100K	≥35	<GR9	F	WC	.287	3.48	25
<100K	≥35	<GR9	E	NWC	.318	3.15	26
>100K	<35	<GR9	F	NWC	.335	2.99	27
>100K	≥35	<GR9	F	NWC	.338	2.96	28
<100K	<35	<GR9	F	WC	.388	2.58	29
<100K	<35	<GR9	E	WC	.431	2.32	30
<100K	<35	<GR9	F	NWC	.454	2.20	31
<100K	<35	<GR9	E	NWC	.508	1.97	32

occupation-reading effect. This effect only produces a marginally significant improvement in fit. However, in light of the fact that the occupation effect is in addition to and net of the effects of education (that is, it is a partial effect which controls for the impact of education), and in light of the significant improvement in χ^2 explained, we have included this term in the

final model. It is important to note that as in the case of multiple regression the effects in the logit model are partial effects (independent of the effects of all other terms in the model). The substantive meaning of these effects will become clearer as we complete this discussion.

LOGIT MODEL

Since we have dichotomized our data, we can apply the more powerful regression-like logit techniques. One basic difference between the log-linear and logit models is that the logit model specifies the likelihood of belonging to a particular category of the dependent variable (i.e., Category 1) and only includes terms which involve the dependent variable (i.e., reading). Equation 2 represents the logit model in the fashion suggested by Goodman (1976).

$$\Phi\, JKLMN = +\beta^{RO} + \beta^{REL} + \beta^{RAS} + \beta^{RSL} \qquad (2)$$

This logit model is created by doubling the corresponding λ of the log-linear model. The resulting equation identifies the relative (partial) impact of these effects:

$$\Phi\, JKLMN = -.08^{RO} + .09^{REL} - .12^{RAS} - .09^{RSL} \qquad (3)$$

Using this model we can predict how individuals belonging to various combinations of categories of the independent variables would perform on the dependent variable. In other words, we can accurately predict whether an individual will choose to read or not to read on the basis of minimal dichotomous information regarding his education, age, language, occupation, and residence.

DISCUSSION

One should generally be cautious about interpreting the parameters of the log-linear and logit models. This is because there are variations in the algorithms used by different computer packages available for estimating log-linear models which produce slightly different estimates, (cf., Holt, 1979). However, as Page (1977) and Holt (1979) argue, estimates based on the odds ratios calculated from the fitted values of the table (predicted from the logit model) will not suffer from these variances. Page (1977) also argues that it is desirable to work with real odds ratios rather than logged odds ratios. The actual odds ratios are substantively more meaningful than the logged terms and are achieved by taking the antilog of the predicted odds ratio.

Table 8 gives the actual odds ratio for being in the nonreading and reading category of the table as a function of various combinations of the independent variables. Following Page and Holt these estimates are derived

from the fitted table. Because the odds ratios in the nonreading category are small fractions, which obscure interpretation, we present the odds ratio of the reading category as well.

The worst performance is produced by a nonmetropolitan, younger, less than Grade 9 anglophone who does not belong to a white-collar occupational group. The odds that this type of individual will be a reader are less than two to one. (Recall that the overall, univariate odds for being in the nonreader category are really nine to one.) At the other end of the continuum we have a younger, white-collar anglophone, with more than a Grade 9 education, who lives in a metropolitan area. Such individuals are nearly 21 times as likely to be readers as nonreaders, or over 10 times as likely to be readers as are young, nonwhite-collar anglophones with less than a Grade 9 education who live in a nonmetropolitan area.

Not suprisingly, the strongest associative effect in the model is that between reading and education (+.66). Clearly those who possess a greater than Grade 9 education are more likely to read. By taking the antilog of the reading-education effect in the logit model, we see that (net of the effects of the other terms in the model) being in the less than Grade 9 class of the education variable increases one's odds of being in the elective nonreader category by nearly a factor of two (1.9 to 1). Similarly, although in a much weaker fashion (i.e., +.11), being younger also increases one's odds of being an elective nonreader. On the other hand, living in a city of over 100,000, being anglophone, or belonging to a white-collar occupational group decreases one's odds of being an elective illiterate.

The erosion of reading behaviour among the younger, post-World War II generation is probably due to the burgeoning of competing leisure media such as television. It is interesting to note that the younger (under 35) population is less likely to read, in spite of the fact that they are much more likely to have a Grade 9 education or more. Whether it is a function of the popularity of competing electronic media, or changes in the efficacy of public school curricula, Canadian educators must note with some alarm that our massive investments in human capital have not produced a more literate society which reads more. Perhaps these linkages underline the inextricable connection of social and technological change.

Reading and settlement size also are significantly associated (−.13). The odds of being a nonreader decrease if one lives in a large metropolitan area. A recent study by deVille, Dugas, and Graves (1980) has shown that cultural participation increases in areas where the comunity's cultural infrastructure is well developed. This finding applies to reading activity and reading resources. Since larger metropolitan areas have more bookstores and libraries, this may explain this association.

The same argument regarding the supply of cultural facilities in general, and reading facilities in particular, can be used to account for the language effect we have identified (−.13). Canadian francophones may read less

and experience higher levels of elective illiteracy as a function of the paucity of library services for Québécois (Ben-Gera and Kinsley, 1980). Furthermore, French books are in shorter supply, and often more costly, even from commercial sources such as bookstores.

The final bivariate effect, between ocupation and reading (−.08), is somewhat predictable. White-collar respondents are less likely to be elective illliterates than nonwhite-collar respondents. It is important to keep in mind that this effect is independent of, and in addition to, the effect of education. Hence, we can conclude that membership in the white-collar occupations encourages reading activity.

None of these bivariate effects are particularly surprising. Yet if we were to rely only on these simple relationships to form our explanations and predictions of elective illiteracy, we would be misled. This is because there are significant higher order specification effects which confound these simple linkages. One of the advantages of log-linear and logit methods is that the analyst is more likely to discover these types of effects in the hierarchical process of fitting a model.

For example, the specification effect between reading, education, and language (+.09) suggests that being both anglophone *and* having less than a Grade 9 education further increases one's odds of being an elective illiterate. The same effect applies to francophones who have greater than a Grade 8 education. This indicates that education is more efficacious in determining reading behaviour among Canadian anglophones than it is among Canadian francophones.

These variations in the role of education and literacy for the two language groups are quite remarkable. One could speculate that education generally has a greater impact on the life-styles and consumption patterns of English Canadians than it does on these aspects of Canadian francophone society. One could further speculate that Canadian francophone culture is less rooted in the written word and less differentiated across educational boundaries.

There is also a significant specification effect for reading, age, and settlement size (−.12). This effect means that being both young and living in a large metropolitan area significantly decreases one's odds of being electively illiterate. This may be due to the fact that the superior literary cultural infrastructure associated with larger cities, and the increased reading opportunities this provides, is a more important factor for younger Canadians than it is for older Canadians. Once again this effect is in excess of and net of the simple effects between age and reading and settlement size and reading.

Finally, there is a significant specification effect due to the mutual relations among reading, settlement size, and language (−.09). This effect implies that if a respondent is both anglophone and from a metropolitan area (or francophone from a nonmetropolitan area) his/her odds of being

in the elective illiterate category are decreased. In other words, both the basic settlement size and language effects, which show that anglophone and urbanites are more likely to read, are reinforced in the case where one is both anglophone and from a large metropolitan area.

SUMMARY AND CONCLUSIONS

We initiated this discussion by discriminating between the illiterate and literate population. We then distinguished within the illiterate group those who were so by virtue of their inability to read and those who chose not to read (i.e., the elective illiterate). We then turned our attention to the task of modeling elective illiteracy – a far more ubiquitous problem in Canadian society. This analysis confirmed our expectations concerning bivariate relationships between education and elective illiteracy, settlement type/elective illiteracy, language/elective illiteracy, age/elective illiteracy, and occupation/illiteracy. The results also indicated that higher order specification effects exist which significantly altered these simple associations. We found that education operates differently across linguistic boundaries, with age having different effects, contingent upon settlement type. Not having completed elementary schooling increases the probability of illiteracy for anglophones more so than it does for francophones. It also appears that urbanites (as opposed to nonmetropolitan residents) who are younger are less likely to be in the elective illiterate category. Furthermore, living in a nonmetropolitan area is more likely to increase a francophone's odds of being a nonreader than it is for a nonmetropolitan anglophone respondent. Quite significantly, despite the fact that the younger population is much more likely to have a Grade 9 or better education than the older population, the former are significantly more likely to fall into the elective illiterate group.

It appears that some of the findings presented here are explicable in terms of traditional differences between Canada's dominant cultural groups. Francophone culture appears to be less differentiated in terms of acquired education perhaps because that culture is not as closely tied to formal, written reservoirs as is the less oral anglophone culture. The fact that a basic anglophone education is more effective in producing readers than francophone lower education is important. This finding may be due to differences in educational curricula and resources, or perhaps it is a function of cultural differences. But this would have to be tested explicitly. We also suggest that the diminished participation in reading activity evident in nonmetropolitan francophone areas is due to a less well-developed distribution systems. We have speculated that many of these effects may be understandable in light of the technological and cultural changes accompanying the emergence of postindustrial society. In this society, pragmatic or effective illiteracy will seriously impair the quality of life of any individual,

whether he or she chooses to or simply is unable to read. The distribution and reproduction of educational and cultural capital are important elements in the principles governing social inequality. Whereas researchers have increased attention devoted to the conspicuous consumers of cultural capital, little attention has been directed to this equally pressing problem of cultural poverty. The present research provides a hint as to the extent and causes of illiteracy in Canada. Equipped with this basic knowledge, educators may begin to take steps to ensure that postindustrial society is not a prelude to postliterate society.

NOTES

[1] For additional information concerning the survey and sample distribution see Watson (1980. Appendices 1 and 4).

[2] This section was based on some of the material presented by Watson (1980).

[3] See M.W. Gillespie (1977. pp 105–106) who cites the fact that Knoke (1975) and Goodman (1970) argue that dummy-dependent regression is reasonable only when the split is no more than 75:25, whereas our split on the elective illiteracy/reader variable is 1728/13215 or 11.6:88.4.

[4] Several excellent descriptions of log-linear and logit techniques are available – e.g., Bishop, Fienberg, and Holland (1975), S. Fienberg (1977) and several papers by Leo Goodman (1970; 1976).

We have also removed from the analysis those who said they were not able to read but claimed to have read. This logically contradictory group constituted less than 0.25% of the total number of respondents. This seemingly illogical group may include some of the blind who do not read any language in the usual sense of the term but do read braille.

[6] For example, if the log-linear model where age is trichotomised into younger, middle-aged and older categories, we find that being younger significantly enhances one's probability of being illiterate, being middle-aged makes no significant difference, and being older significantly lessens the probability that a respondent will be in the illiterate category of the table.

[7] For a discussion of fitting and selection procedures, see J. Davis (1974) or, S. Fienberg (1977, pp. 459–558).

[8] ECTA (Everyman's Contingency Table Analysis) can be obtained from Leo A. Goodman, Department of Statistics, University of Chicago, 1118 East 58th Street, Chicago, Ill. 60637.

[9] LOGLIN, written by Donald C. Olivier and Raymond K. Neff, can be obtained from the Program Librarian, Health Sciences Computing Facility, Harvard School of Public Health, 665 Huntingston Avenue, Boston, Mass. 02115.

REFERENCES

Ben-Gera, M. & Kinsley, Brian L. *Library and bookstore use in Canada.* Ottawa: Department of Communications, 1980.

Bishop, Y. M., Fienberg, S. E., & Holland, P.W. *Discrete multivariate analysis: Theory and practice.* Cambridge, Mass.: MIT Press, 1975.

Davis, J.A. Hierarchical models for significance tests in multivariate contingency tables. In H. L. Costner (Ed.), *Sociological Methodology.* San Francisco: Jossey-Bass, 1974, 189–231.

deVille, B., Dugas, T., & Graves, F. *Community infrastructure and cultural participation: Technical report.* Secretary of State, 1980.

Evans, Peter et al. *British Columbia reading assessment* (3 vols.). Victoria: British Columbia Ministry of Education, 1977.

Fienberg, S. E. *The analysis of cross-classified categorical data.* Cambridge, Mass.: MIT Press, 1977.

Gillespie, M. W. Log-linear techniques and the regression analysis of dummy dependent variables. Further bases comparison. *Sociological Methods and Research,* August 1977, *6,* 103–122.

Goodman, Leo. The multivariate analysis of qualitative data: interactions among multiple classifications. *Journal of the American Statistical Association,* 1970, *65,* 226–256.

Goodman, Leo. The relationship between modified and more usual multiple regression approaches to the analysis of dichotomous variables. In D. Heise (Ed.), *Sociological Methodology.* San Francisco: Jossey-Bass, 1976, 83–110.

Harman, D. Illiteracy: An overview. *Harvard Educational Review,* 1970, 226–243.

Holt, D. Log-linear models for contingency table analysis: On the interpretation of parameters. *Sociological Methods and Research,* February 1979, *7,* 330–336.

Holzner, B. & Marx, J. H. *Knowledge application.* Boston: Allyn and Bacon.

Knoke, D. A comparison of log-linear and regression models for systems of dichotomous variables. *Sociological Methods and Research,* May 1975, *3,* 416–434.

Olivier, D. C., & Neff, R. K. *LOGLIN 1 users guide.* Ottawa: Carleton University, 1976.

Page, William. Interpretation of boundaries log-linear model effect as odds ratios approach. *Sociological Methods and Research,* May 1977, *5,* 419–436.

Thomas, A. M. *Adult basic education and library activities 1975–1976.* Toronto: World Literacy of Canada, 1976.

Watson, Kenneth F. *Leisure reading habits.* Ottawa: Infoscan, 1980.

Frank L. Graves is president of Ekos Research Associates Inc., 171 Nepean St., Suite 304, Ottawa, Ont.

Brian Kinsley is a Research Analyst with the Department of Communications/Ministère des Communications, Ottawa, Ont. K1A 0C8.