Canadian Heritage | Patrimoine canadien

Research, Analysis and Industry Development

Explanatory Model For Leisure Book Reading Rates In Canada

Final report

Prepared by Créatec +

Canada

Printed edition
Catalogue No. CH44-74/2006E
ISBN 0-662-49056-8

PDF edition
Catalogue No. CH44-74/2006E-PDF
ISBN 0-662-42950-8

Cette publication est également offerte en français et peut être consultée sur le site Web du ministère du Patrimoine canadien : www.patrimoinecanadien.gc.ca.

FINAL REPORT

EXPLANATORY MODEL FOR LEISURE BOOK READING RATES IN CANADA

Submitted to
Canadian Heritage
Publishing Policy and Programs

Contract #: 45164955

November 30, 2005

LES ÉTUDES DE MARCHÉ CRÉATEC +

206 Pine Avenue East – Montreal, Quebec, H2W 1P1
Tel.: (514) 844-1127 - Fax: (514) 288-3194
Email: info@createc.ca / Web site: www.createc.ca

TABLE OF CONTENTS

1. SUMMARY OF RESULTS

- The results of a survey carried out by Créatec + on behalf of Canadian Heritage in January 2005 ("Reading and Buying Books for Pleasure") highlighted the same key demographic and social factors that influence book reading as those brought to light by most of the previous Canadian, European and US studies[1], namely gender, education level, geographical location and language.
-
  → The results showed that book reading for pleasure is mainly a female pursuit, varies greatly by activity, region and language, and decreases when the education level drops.

  → Men's and Francophones' reading habits still lag behind, and the results among Francophones outside Quebec are a cause for concern.

- However, the close interrelationships among language, region and education, in particular, mask or dilute the true relevance of these variables and, perhaps, others as well. To better understand the true importance of these variables and reassure policy makers and programmers that they are on the right track, we carried out modelling activities similar to those carried out by the National Endowment for the Arts in their report *"Reading at Risk"* by using data collected in 2002 from the Americans.

  → This explanatory and predictive analysis is the first to be carried out using Canadian data.

- With the help of the multiple logistic regression (MLR) technique, explanatory models for the general reading rate (at least one book per year) and regular reading rate (every day or nearly) were developed, for Canada as a whole and then for Francophones in Quebec and outside Quebec, where reading rates are somewhat lagging. Section 2 of the present report contains an overview of the statistical underpinning for this method as well as a guide for interpreting the results.

---

[1] *Reading at Risk: A survey of Literacy Reading in America – National Endowment for the Arts, June 2004.*

Reading in Canada. An Ekos survey carried out on behalf of Canadian Heritage, 1978 and 1991.

Enquête sur les pratiques culturelles des Québécoises et des Québécois. Ministère de la culture et des communications du Québec, 1999.

Canadian Book Buying Habits: What Influences Purchases? – Pollara, 2002.

Europeans and culture. Eurobarometre, 2001.

ALL CANADIANS

- Overall, the results of our statistical models show that although the situation in Canada is very different from that in the United States (where the reading rate is on the decline and much lower than in Canada), general reading and regular reading are chiefly influenced by the same two variables, i.e. gender and level of education, the other variables decreasing in relevance after their interrelationships are taken into account.

  → Women read far more than men. Our MLR analyses show that the odds[2] that a woman will read at least one book per year are 2.9 times greater than those of a man, and the odds that a woman will read regularly are 2.4 times greater than those of a man.

  → The more education a person has, the more likely that person is to read. The odds of reading are 4.5 times greater among those with a university education than among those with a high school education or less, and the odds of regular reading are 2.6 times greater.

- Contrary to what has been observed in the United States, however, reading in Canada is not influenced by age. Even after controlling for the effect of the other variables, age is not a relevant variable for explaining reading rates in Canada, which tends to confirm the absence of a generation gap and contradicts the belief that young people avoid reading books for pleasure or that this type of reading among young people is a major problem in Canada.

  → The higher level of education in younger generations likely helps maintain reading rates among them.

- The results of the MLR also contradict the widely-held belief that the Internet has had harmful effects on book reading rates. Thus, chatting over the Internet has a positive and significant correlation with the general leisure book reading rate. In a broader sense, being connected to the Internet has no bearing on reading rates. It should be recalled that neither the general reading rate nor the average time devoted to reading has declined since 1991, which is when the Internet arrived on the scene.

  → The odds of reading books are 1.7 times higher among respondents who chat over the Internet than among those who do not.

- Although Francophones have the lowest reading and regular reading rates in the country, the overall influence of membership in a language community is significantly lower than the influence of gender or level of education, when the effect of the other variables is controlled for.

[2] *Read Section 2 to learn the specific statistical meaning of the term "odds", compared to "probability".*

→ A close examinations of the data suggests that a lower level of education among Francophones, especially those outside Quebec but inside Quebec as well, goes a long way toward explaining why both these language communities are behind in their reading in relation to other Canadians. After the effects of education and gender are taken into account:

– Anglophones living in an environment where their official language is the majority (hereafter "majority Anglophones") have 1.8 times higher odds of reading regularly than Francophones outside Quebec.

– Francophones in Quebec have 1.1 times higher odds of reading regularly than Francophones outside Quebec. This finding suggests that after taking into account the effects of educational gaps, the odds of reading regularly are almost the same for Francophones in Quebec and their counterparts outside Quebec.

- Engaging in a hobby, crafts, music or arts activities has a positive influence on general and regular reading rates (similar observation in the NEA analyses).

→ This factor ranks third among influences on the general reading rate, behind gender and level of education.

– The odds of reading books for pleasure are twice as high among respondents who engage in a hobby, crafts, music or arts activities than among respondents who do not.

→ This factor ranks fourth among influences on the regular reading rate, trailing gender, level of education and membership in majority or minority official language communities.

– Respondents who engage in a hobby, crafts, music or arts activities have 1.7 times higher odds of reading regularly than those who do not.

- Last, not skipping over unfamiliar words (which can reflect good reading skills or a positive self-assessment of one's ability to read) has a positive correlation with the general reading rate and reading regularity. When looked at more closely, however, this apparent facility with reading cannot be considered an indicator of better education. It could just as easily result from a greater interest in language in general and reading in particular, rather than a good understanding of the written word.

→ This variable ranks sixth among influences on the general reading rate, right behind chatting on-line.

→ It ranks fifth among influences on regular reading.

FRANCOPHONES IN QUEBEC

- Among Francophones in Quebec, gender and level of education have a significant influence on the general reading rate, but this influence is less pronounced than among Canadians as a whole. Knowing that there is a public library near one's home and engaging in volunteer activities are more useful predictors of the general reading rate than gender and education level.

- This observation suggests that the first step in promoting reading to Francophones in Quebec should be to make them more aware of the places where books can be found. The results also suggest that in Quebec, the more one is involved in one's community, the more one tends to read for pleasure (an observation similar to one made concerning Americans in the NEA study).

- But it is still between low and high level of education that the difference in the odds ratio for reading remains greatest, which suggests that level of education remains a crucial variable in any national policy aimed at promoting book reading for pleasure.

  → The odds of reading are 6.4 times higher among respondents with a university education than among those with no more than a high school education.

  → The odds of reading are 2.9 times higher among respondents who report doing volunteer work than among those who do not.

  → The odds of reading are 2.5 times higher among respondents who say they live close to a library than among those who do not.

    – Unfortunately, it is impossible to determine whether the actual proximity of a public library is what this boils down to or whether the respondents' perception is selective, reflecting instead their level of interest in reading; only factual data on the existence of such establishments would shed light on the true importance of this factor, which is closely linked to the general reading rate.

  → The odds of reading are 2.3 times higher among women than among men.

- As for regular reading, the main influencing factor is gender, followed by known proximity of a public library, family income and attendance at shows. Level of education does not have a significant bearing on the regularity of reading books for pleasure, after the preceding variables have been taken into account.

- These observations on the apparent shift in predictive value of the regularity of reading among Francophones in Quebec from level of education to income could be explained by two factors that combine to make access to books more difficult:

  – higher prices for books in French in Quebec than elsewhere in Canada;

  – the number of public libraries in Quebec is lower than the national average.

  → If this assumption is plausible, we should be cautious before concluding that level of education does not play a role in regular reading in Quebec, because income and education are linked. These findings suggest that difficulty accessing leisure books impedes regular reading more than level of education does, which illustrates just how important this impediment is in Quebec.

  → The odds of reading regularly are 2.9 times higher among women than among men.

  → The odds of reading regularly are 2.1 times higher among respondents who consider that there is a library near where they live than among those who do not.

  → The odds of reading regularly are 1.8 times higher among respondents with a family income of over $20,000 than among those with an income of under $20,000.

  → The odds of reading regularly are 1.6 times higher among respondents who say they attend shows than among those who say they do not.

FRANCOPHONES OUTSIDE QUEBEC

- Among Francophones outside Quebec (whose reading rates are reported to be the lowest in the country), gender influences the general reading rate greatly, but clearly less than income. Not skipping over unfamiliar words also influences the reading rate (it should be recalled that Francophones outside Quebec, in far greater numbers than the other respondents, report that they have difficulty reading and that they read to improve their language).

  → The odds of reading are 5.5 times higher among respondents with a family income of over $40,000 than among those with an income of under $20,000.

  → The odds of reading are 2.4 times higher among women than among men.

  → The odds of reading are 2.4 times higher among respondents who say they do not skip over unfamiliar words than among those who say they do.

- Regular reading in this language community is influenced primarily by level of education, followed by gender and watching movies at the theatre.

  → The odds of reading regularly are 7.2 times higher among respondents with a university education than among those with a high school education or lower.

→ The odds of reading regularly are 2.1 times higher among women than among men.

→ The odds of reading regularly are 1.9 times higher among respondents who watch movies at the theatre than among those who do not.

- So the income factor is present with this other Francophone group, but this time it influences the general reading rate, whereas level of education influences reading regularity.

- All in all, the results as a whole show that gender and level of education are the best predictors of book reading for pleasure and should therefore be the building blocks of a policy aimed at narrowing the reading gaps observed between the various social environments, over any other type of segmentation of the population, be it in terms of age categories or even language.

  → Narrowing the gaps between men and women, and between levels of education, should have spin-offs in other environments that are also lagging behind, which this type of statistical analysis demonstrates.

  → It should be pointed out, however, that among respondents who consider French to be their main official language, income also plays a large role, likely on account of a higher price structure for books in French.

  → Furthermore, in Quebec it seems that lower availability or more limited awareness of public libraries is accentuating the difficulty of having access to books and is impacting on reading rates.

- The following tables sum up the key findings from the prediction models prepared using the Logit analysis.

  → These findings can also be considered as indicators that can be followed in future to pinpoint changes in reading rates rather than relying solely on an observational approach, which would complement the follow-up table for leisure book reading habits.

  → The following chapter (Methodology) contains a guide to help interpret the statistical results presented in these summary tables.

SUMMARY OF PARTIAL CORRELATIONS (R)

READING			REGULAR READING		
Independent variables		R	Independent variables		R
TOTAL CANADA					
1.	Gender	0.187	1.	Gender	0.177
2.	Level of education	0.177	2.	Level of education	0.135
3.	Arts activities	0.115	3.	Language community	0.088
4.	Language community	0.078	4.	Arts activities	0.084
5.	Internet chatting	0.070	5.	Not skipping over words	0.037
6.	Not skipping over words	0.064			
FRANCOPHONES IN QUEBEC					
1.	Library nearby	0.150	1.	Gender	0.196
2.	Does volunteer work	0.133	2.	Library nearby	0.120
3.	Gender	0.129	3.	Family income	0.081
4.	Level of education	0.098	4.	Attends shows	0.071
FRANCOPHONES OUTSIDE QUEBEC					
1.	Family income	0.234	1.	Level of education	0.217
2.	Gender	0.137	2.	Gender	0.118
3.	Not skipping over words	0.132	3.	Movies at theatre	0.106

ODDS RATIO SUMMARY

GENERAL READING RATE			REGULAR READING		
Independent variables categories		Odds ratio*	Independent variables categories		Odds ratio
TOTAL CANADA					
1.	University education	4.5	1.	University education	2.6
2.	Woman	2.9	2.	Woman	2.4
3.	College education	2.2	3.	Majority Anglos	1.8
4.	Arts activities	2.0	4.	Arts activities	1.7
5.	Majority Anglos	1.8	5.	Minority Anglos	1.6
6.	On-line chatting	1.7	6.	College education	1.4
7.	Not skipping over words	1.5	7.	Not skipping over words	1.2
FRANCOPHONES IN QUEBEC					
1.	University education	6.4	1.	Woman	2.9
2.	Volunteer work	2.9	2.	Library nearby	2.1
3.	Library nearby	2.5	3.	Income of $20 000 +	1.8
4.	Woman	2.3	4.	Attends shows	1.6
FRANCOPHONES OUTSIDE QUEBEC					
1.	Income of $40 000 +	5.5	1.	University education	7.2
2.	Not skipping over words	2.4	2.	Woman	2.1
3.	Woman	2.4	3.	Movies at theatre	1.9
4.	Income between $20 and 40 000	2.3			

* *An odds ratio is obtained by dividing the odds of reading (regularly) for one category by the odds of reading (regularly) for its reference category.*

2. METHODOLOGY

2.1 CONTEXT

- On behalf of Canadian Heritage, Créatec + conducted a national telephone survey in January 2005 ("Reading and Buying Books for Pleasure"), the purpose of which was to provide a detailed statistical picture of the habits of Canadians with respect to buying and reading books for pleasure, as well as to update the findings of *Reading in Canada 1991,* undertaken by Ekos on behalf of Canadian Heritage.

- The data collected this year indicate that reading for pleasure is a solidly established and widespread habit, showing little if any change over the past 15 years and suffered no effects from the popularity of the Internet.

  → Eighty-seven per cent of the Canadians who took part in the survey read, if only part of a book per year, and one half (54 per cent) read virtually every day.

  → Canadians appear to be distinctly different from their American counterparts, almost half of whom read an average of less than one book per year and whose reading rate has substantially decreased over the past 20 years, particularly among those in the 18-24 age group.

  → Canadians read in far greater numbers than residents of the member countries of the European Union, according to the average reading rates for these countries (60 per cent).

  → Among industrialized countries for which comparable results are available, Nordic countries lead the way when it comes to reading for pleasure. Canada ranks number one (87 per cent), followed by Sweden (80 per cent) and Finland (75 per cent).

- Like other prior studies[3], this survey highlights the close link between certain demographic and social factors and attitudes to reading books, for example, gender, education level, geographical location and language. Moreover, the same disparities in reading rates that had been observed in 1991 were observed this year, in particular those having to do with language and gender.

  → For example, the mainly female leisure readership varies greatly by activity and decreases when the education level drops.

  → The reading habits of Francophones in Quebec are still deficient, and the 2005 survey shows that the reading situation of Francophones outside Quebec lags farthest behind.

    – In Quebec, the regular reader rate remains the lowest in the country (46 per cent) while it exceeds 50 per cent everywhere else, peaking in the Western Provinces at 59 per cent (B.C.) and 60 per cent (Prairies).

    – Only 31 per cent of Francophones outside Quebec sampled in this survey reported that they were regular book readers compared to 55 per cent of other respondents.

- However, the knowledge amassed to date on Canadians' reading habits is descriptive. While very useful in a background role as part of an overall analysis of book reading for pleasure, this type of information does have its limits. Indeed, the close interconnections among language, region and level of education, in particular, mask the true relevance of these (and perhaps other) variables. This makes it difficult for policy makers and programmers to ensure that their actions are appropriate, since their knowledge of the true importance (determinance) of the variables with which they deal is, at best, imperfect.

---

[3] *Reading at Risk: A survey of Literacy Reading in America – National Endowment for the Arts, June 2004.*

Reading in Canada. An Ekos survey carried out on behalf of Canadian Heritage, 1978 and 1991.

"Enquête sur les pratiques culturelles des Québécoises et des Québécois." Ministère de la culture et des communications du Québec, 1999.

Canadian Book Buying Habits: What Influences Purchases? – Pollara, 2002.

"Europeans and Culture." Eurobarometre, 2001.

- This issue was addressed by the National Endowment for the Arts in the United States, which conducted explanatory analyses on 2002 data.[4] The purpose of their analyses was to predict the rate and frequency of literary reading, using a statistical technique that assigns each of the main socio-demographic variables, which are statistically linked to reading habits, its true influence, i.e. its individual influence after removing from consideration the influence of all the other variables. To do so, they used the multiple logistic regression statistical model, also known as Logit Analysis.

  → For example, their results showed that, all things being otherwise equal, the odds that an American adult having completed post-secondary education will read literary works[5] are at least twice (2.3) as high as those for an adult who did not pursue higher learning. To put it differently, failure to pursue higher learning translates into only 42 per cent of the odds of being a reader, compared to those who did pursue a higher education.

- The modelling contained in this report is completely similar.

2.2 PURPOSE OF THE MODELLING

- Using the data from this year's survey[6], the aim of the Logit regression analyses was to generate a reading habits prediction model with the help of socio-demographic variables (age, gender, membership in a linguistic group, region, level of education, etc.) and behavioural variables (does volunteer work, visits museums, parents read to respondents at early age, etc.).

- When the predictive value of the model is good, this multivariate statistical analysis technique guides us toward the socio-demographic and behavioural variables most likely to predict reading habits and gives us an accurate measure of their actual influence:

  1. Their net contribution (controlling for the effect of the other variables) to reading habits;

  2. Their place, rank, hierarchy among all the explanatory variables.

---

4 *Reading at Risk: A Survey of Literacy Reading in America – National Endowment for the Arts. Research Division Report #46, June 2004. In particular, see Appendix C.*

5 *Novels, poetry collections.*

6 *Reading and Buying Books for Pleasure – January 2005. Créatec survey commissioned by Canadian Heritage.*

- This technique was used to predict two reading rates among three populations. Thus, we developed the following six prediction models:

  1. Total Canada

     - Reading rate (at least one book per year)
     - Regular reading rate (at least every day)

  2. Francophones in Quebec

     - Reading rate (at least one book per year)
     - Regular reading rate (at least every day)

  3. Francophones outside Quebec

     - Reading rate (at least one book per year)
     - Regular reading rate (at least every day)

2.3 OVERVIEW OF LOGISTIC REGRESSION

1. Multiple logistic regression

- Also known as *Logit analysis*, logistic regression is particularly useful in seeking to understand or predict the effect of one or more categorical variables on a dichotomous (binary) variable, i.e. which can take only two values (e.g. reader, non-reader). This is a robust form of discriminant analysis that is not subject to the constraints of the normal distribution.

- This statistical method can be used to model the likelihood that an event or state will happen, given the values or states of a series of quantitative and/or qualitative "explanatory" variables (e.g. gender, level of education, age, having a hobby, etc.).

- The results of logistic regression can be used to explain a situation but also to characterize a specific population for which the state of the explained variable is unknown. The model is simply used with the values of the descriptive variables known for this population or specific group.

- Contrary to conventional multiple regression or discriminant analysis, this technique does not require a normal distribution of predictors, nor homogeneity of variances.[7]

[7] *Julie Desjardins. L'analyse de régression logistique. Tutorial in Quantitative Methods for Psychology 2005, Vo. 1 (1), p. 35-4.*

- Logistic regression can be carried out in a variety of ways. The one that was used is the direct method, where no specific hypothesis is formulated concerning the order or importance of the predictor variables selected in the model. In fact, this type takes into account only superior predictors, because each is assessed as though it was the latest to enter into the equation. Consequently, if a predictor is strongly correlated with the dependent variable, it is possible that it shows only a weak prediction in the presence of the other predictors. That is the method we used.

  → One of the main advantages of *Logit* regression is that it translates the relationships between the explained variable (e.g. reading rate) and the explanatory variables (e.g. level of education) in terms of the odds of reading in relation to reference categories (odds ratio).

  → Gamblers are very familiar with the notion of odds. It is the relationship between the probability that an event will take place and the probability that it will not. For example, if the odds ratio is 5 to 1, one would expect the event to take place five times more often than it does not.

  → An equivalent way of characterizing the frequency of an event is its probability.

  → To illustrate this, the following table matches the odds ratio of occurrence with its equivalent probability.

Odds	Probability
0.11	0.1
0.25	0.2
0.43	0.3
0.67	0.4
1.00	0.5
1.50	0.6
2.33	0.7
4.00	0.8
9.00	0.9

Odds*: 0 = p/1-p = probability of event taking place / probability of event not taking place.*
***Probability**: P = 0/1+0*

- Although odds and probability are equivalent ways of characterizing an event, the results of the *Logit* analysis do not allow the probabilities of an event to be calculated, because it provides an odds ratio $(p_1/1\text{-}p_1)/(p_2/1\text{-}p_2)$. Thus, if the odds of reading among women are 2.3 times those of men, this must not be interpreted as meaning that their probability of reading is 2.3 times higher.

→ For example, this result would be the same if:

– 70 per cent of women and 50 per cent of men read (odds ratio = 2.3 / 1.0 = 2.3 and probability ratio = 70/50 = 1.4).

– 82 per cent of women and 67 per cent of men read (odds ratio = 4.6 / 2.0 = 2.3 and probability ratio = 82/67 = 1.2).

→ The odds ratio is therefore an indicator of the extent of the difference in probabilities that an event will occur for two categories of the same explanatory variable, and thus makes it possible to target the largest gaps.

2. Typical equation

- In the Logit prediction models, we find:

→ An explained variable (Y), e.g. reading or not reading;

→ Explanatory variables (x), e.g. gender, region, age, etc.;

→ Coefficients (β - bêta) which measure the influence of each of the explanatory variables (x) on the probability that the explained variable (Y) will occur.

- The logistic regression formula is as follows:

$$\log (p_i / 1-p_i) = \alpha + \beta_1 x i_1 + \beta_2 x i_2 + ... + \beta_k x i_k$$

where *k* is the number of explanatory variables, *i* = 1,...., *n* the individuals in the sample and p_i the probability that Y=1 (reading). The term 'log (p_i /1-p_i)' is what is called the logit or log-odds. Most researchers use natural logarithms, but base-10 logarithms can also be used.[8]

3. Multicollinearity

- When two variables are intercorrelated (collinearity phenomenon) and one of them has little to do with the explained variable, this makes the model needlessly cumbersome, and it becomes difficult to determine the unique or net influence of each variable when taking into account the influence of the others.

→ Many socio-demographic variables are intercorrelated. For example, income is correlated with education, or mother tongue with region.

[8] *Logit provides us with the odds ratio, which enables us to visualize where the main gaps are.*

→ For this reason, care must be taken in choosing the explanatory variables that will enter into the model while eliminating the variables that have high collinearity with others and that are less relevant in terms of explaining the reading rates.

4. Modelling steps

- For each reading rate to be predicted, the modelling involved three steps:

  1. A close examination of the frequency distributions for each of the explanatory variables in order to identify whether it would be appropriate to group some categories.

     – For example, following this examination, we grouped together certain age categories.

  2. Conduct analyses of the collinearity of the explanatory variables and determine which ones to include in the model;

  3. Multiple logistic regression analysis between the selected explanatory variables and reading behaviour.

     – The variable category where the lowest reading rate is observed was used as the reference category for calculating and interpreting the odds ratios.

- We used SPSS's LOGLINEAR procedure to develop all the models presented in this report, rather than SAS PROC LOGISTIC, since social science researchers are generally more familiar with SPSS.

- For further details on Logit analysis, visit the site operated by Paul-Marie Bernard of Université Laval's Département de médecine sociale et préventive www.uquebec.ca/reglog.

2.4 GUIDE FOR INTERPRETING THE RESULTS

- Logistic regression produces various results (presented in the tables associated with each model) that we will briefly describe here:

  1. *β* (bêta) is the regression coefficient of each category of explanatory variables. It measures the (unique) effect of this specific variable on the odds of reading when the effect of all the other variables is held constant.

  2. *Exp (β)* or the exponential of β is the odds ratio. This ratio is obtained by dividing the odds (p1/1-p) for a sub-group, where p is the probability of an event occurring (reading books for pleasure, for example), by the odds for the reference sub-group.

For example, in the model for Canada as a whole, the regression coefficient of Francophones in Quebec for reading books for pleasure is 0.38 (logarithm for the odds of reading in this language community in relation to Francophones outside Quebec) when the effect of all the other variables is held constant.

By taking the inverse of the natural logarithm of the regression coefficient ($e^{0.38}$), an odds ratio of 1.474 is obtained. It is this indicator that is used to interpret the results. This ratio indicates that the odds of Francophones in Quebec reading are 1.5 times higher than those of Francophones outside Quebec (the reference for the language community variable). In other words, the reading odds of a Francophone living in an environment where his or her official language is the minority (a "minority Francophone") are 68 per cent of those of a "majority Francophone" (and 55 per cent of those of a "majority Anglophone").

3. *R* is the partial correlation coefficient of a variable with the explained variable. This coefficient reflects the link between the independent variable and the dependent variable, while at the same time controlling for the effect of the other variables (collinearity).

   For example, in the model to predict reading books for pleasure (at least one per year) for Canada as a whole, the language community has a regression coefficient of 0.078, which, once the effect of all the other variables is taken into account, is well below the one for gender (R = 0.187) or level of education (R = 0.177). In other words, once the effect of education and gender has been taken into account, the effect of the official language community is significant but not nearly as important as that of the most influential variables.

   This type of results suggests to policy makers that it is much more useful to promote reading first and foremost to those with a lower level of education and men, throughout the country, rather than to other communities or regions.

4. *Classification table*: this table compares reality (reader or not) with the model's prediction (reader predicted or not). It is an indication of the quality of the logistic regression model.

5. *Nagelkerke pseudo* R^2: an estimate of variance of the predicted variable explained by the model, another indicator of the predictive value of the model.

2.5 VARIABLES CONSIDERED FOR INCLUSION

- For each of the logistic regression analyses presented in the report, 24 variables were considered for inclusion in the model, i.e. 14 socio-demographic variables and 10 behavioural variables:

  → Preferred official language;
  → Age group;
  → Type of household;

- → Level of education;
- → Employment activity;
- → Membership in a visible minority;
- → Access to the Internet;
- → Family income;
- → Gender;
- → Area of residence (urban or rural);
- → Region;
- → Language community (minority or majority / Francophone or Anglophone);
- → Access to a bookstore nearby[9];
- → Access to a public library nearby[10];
- → Never skips over unfamiliar words;
- → Their parents read to them at an early age;
- → Watches television shows or videos;
- → Listens to music;
- → Engages in a hobby, crafts, music, arts activities;
- → Does volunteer work;
- → Chats on-line;
- → Watches movies at a theatre;
- → Visits museums or art galleries;
- → Attends artistic performances.

- According to the predicted variable and the population considered, certain variables were eliminated, either because they were not enough correlated with the predicted variable or because they were strongly correlated with other variables whose correlation with the predicted variable was higher.

- Ultimately, each of the six logistic regression analyses was conducted using between three and six predictor variables. The variables selected by the modelling are presented in section 3 (results of prediction models).

[9] *"That offers books in French" for analyses specific to Francophones outside Quebec.*

[10] *"That offers books in French" for analyses specific to Francophones outside Quebec.*

3. RESULTS OF PREDICTION MODELS

3.1 TOTAL CANADA – GENERAL READER RATE

1. VARIABLES IN THE MODEL

- An examination of the data for the Canadian sample as a whole allowed an initial identification of the variables with little or no relation to the general reading rate (at least one book per year).

  → We compared the frequency distributions of independent variables for readers and non-readers. The frequency distributions showed that seven of them were not sufficiently associated with the reading rate, and they were excluded right from the start. Thus, the series of independent variables likely to influence reading was first reduced to 19 variables (see the **Appendix** for the list of this sub-series).

- Next, we proceeded with a multiple correlation analysis between the sub-series of 18 independent variables and the dependent variable.

  → At this stage, 13 of the 19 variables were eliminated from the model, either because they were not correlated enough with the dependent variable or because they were strongly correlated (collinearity) with another explanatory variable.

- Thus, the logistic regression analysis to predict the reading of at least one book per year among Canadian respondents as a whole used six independent variables:

  1. Level of education;
  2. Gender;
  3. Language community;
  4. Does not skip over unfamiliar words;
  5. Chats on-line;
  6. Engages in arts activities or hobbies.

- It was observed that age, region and income did not have enough of a specific relation with the general reading rate or of a unique effect to be included in the model. These observations suggest that for policies aimed at promoting reading in Canada, age is not a priority consideration – a finding, in fact, that emerged from the descriptive data. They also suggest that an official language community's majority or minority situation is a more relevant policy-making consideration than region or simply the preferred reading language, which could not be deduced solely from the survey's descriptive data.

2. CODING OF VARIABLES

- The dependent variable was coded "0" for non-readers and "1" for readers.

- The categories of independent variables were grouped together and recoded in ascending order of their reader rates, and the resultant categories are presented in **Table 1a** (below). The category whose value is "0" represents the reference category for each variable in the odds ratio interpretation (all the other categories are compared to the reference category in the logistic regression analysis).

TABLE 1A VARIABLES SELECTED TO PREDICT THE GENERAL READING RATE IN CANADA

Independent variables	Categories
▪ *Level of education*	• **[High school or less]** • College • University
▪ *Gender*	• **[Man]** • Woman
▪ *Language community*	• **[Minority]** • Majority French • Majority English
▪ *Does not skip over unfamiliar words*	• **[No]** • Yes
▪ *On-line chatting*	• **[No]** • Yes
▪ *Engages in arts activities and hobbies*	• **[No]** • Yes

*N.B. The reference category is in **[bold]**.*

3. INTERPRETATION OF RESULTS

- **Table 1b** presents the partial correlation coefficients of each of the variables and categories of variables. It measures the correlation between the independent variable and the dependent variable when the effect of the other variables included in the model is taken into account. Table 1b also ranks each variable in the predictive model.

- Thus, gender (R=0.187) is the variable most associated with the general reading rate in Canada, followed by:

  → level of education (R=0.177)
  → engages in arts activities and hobbies (0.115)
  → language community (0.078)
  → chatting on-line (0.070)
  → and not skipping over unfamiliar words (0.064).

- Since the official language community contributes relatively little (R=0.078) to the model, this suggests that the low general reading rate observed among Francophones outside Quebec – and to a lesser extent among Francophones in Quebec – in relation to the other language groups is primarily attributable to other variables, in particular level of education.

TABLE 1B PARTIAL CORRELATIONS OF PREDICTOR VARIABLES FOR LEISURE BOOK READING – TOTAL CANADA (N=1 963) –

Independent variables	Variables		Categories	
	R*	Rank	R*	Rank
Level of education	0.177	2		
▪ College			0.125	3
▪ University			0.150	2
Gender	0.187	1		
▪ Woman			0.187	1
Language community	0.078	4		
▪ Majority French			---	---
▪ Majority English			0.086	5
Not skipping over unfamiliar words	0.064	6		
▪ Agree			0.064	7
Internet chatting	0.070	5		
▪ No			0.070	6
Engages in arts activities and hobbies	0.115	3		
▪ No			0.115	4

* *Partial correlation coefficient (R) which measures the level of correlation between the independent variable and the dependent variable when the effect of the other variables included in the model is taken into account. All the partial correlation coefficients indicated are statistically significant to P < 0.01.*

- In the bottom part of **Table 1c**, we see that the model correctly classifies a high proportion of respondents, namely 66.8% of the cases observed. More specifically, the model manages to produce 66.1% true positives (classified by the model as readers when they are in actual fact readers) and 71.3% true negatives (classified by the model as non-readers when they are in actual fact non-readers).

- **Table 1c** presents the regression coefficients (β) and the odds ratios of each of the categories of predictor variables for general reading in Canada. The regression coefficient reflects the unique effect (while controlling for the effect of the other variables) of each predictor variable on the logarithm for the odds ratio that a person reads for pleasure in Canada.

- The odds ratio[11] reflects the increase in the odds of reading associated with membership in a particular category rather than in the reference category. Thus, the odds ratio of reading among respondents with a university education (β = 1.523) are four and a half times[12] (4.586) higher than among respondents with a high school education or less. In other words, the odds of reading among those who have not completed post-secondary studies are equivalent to 21 per cent of the odds among those who have.

- The highest odd ratios were observed for the following categories:

  → University education (4.5 times higher than those among respondents with high school or less);

  → women (2.9 times higher than among men);

  → college education (2.2 times higher than those among respondents with high school or less);

  → arts activities (2.0 times higher than those among respondents who do not engage in arts activities);

  → Anglophones in the majority (1.8 times higher than those among respondents living in a minority language environment).

- Looking at the odds ratios, one notes the following:

  1. The largest gap is between categories of education. The odds of reading among respondents who did not complete post-secondary studies are only 21 per cent of those among respondents who reached the university level. There is also a substantial gap between respondents who did not go on to university after college and those who completed university.

  2. The gap is also very large between men and women, well above that between majority and minority language situations. Thus, the odds of reading among men are only 34 per cent of those among women.

- These results speak for themselves: the <u>first</u> priority of policies aimed at promoting book reading for pleasure in Canada should be to target those with the lowest levels of education and men rather than other groups. In other words, of all the socio-demographic variables taken into account, it is narrowing the reading gaps between levels of education and between men and women that is most likely to improve general reading rates in Canada.

---

[11] *Obtained by calculating the inverse of the natural logarithm of the coefficient β.*

[12] *The odds of reading and the probability of reading are not synonymous. The odds of reading (p/1-p) reflect the ratio of the probability of reading divided by the probability of not reading.*

- These observations are very similar to those reported in the United States[13]:

  → In the US, respondents with a university education have 2.3 times higher odds of reading than respondents who did not complete high school (in the United States, the reference category was different and excluded lower levels of education), but the results are comparable, i.e. twice as high in Canada, if these two levels alone are considered.

  → American women have 2.3 times higher odds of reading than men, which is on a par with the Canadian experience. Although comparisons between the two countries might not be totally accurate, they suggest that in Canada, the gender gap is even larger than in the US.

[13] *Reading at Risk: A Survey of Literacy Reading in America – National Endowment for the Arts. Research Division Report #46, June 2004. In particular, see Appendix C.*

TABLE 1C **SUMMARY OF RESULTS FROM THE LOGISTIC REGRESSION TO PREDICT BOOK READING FOR PLEASURE – TOTAL CANADA (N=1 963) –**

Independent variables		Regression coefficient β	Standard error	p*
	▪ Constant	-0.383	0.191	0.045
Level of education	▪ College	0.820	0.159	0.000
	▪ University	1.523	0.250	0.000
Gender	▪ Woman	1.080	0.143	0.000
Language community	▪ Majority French	0.388	0.201	0.053
	▪ Majority English	0.598	0.162	0.000
Not skipping over unfamiliar words	▪ Agree	0.410	0.141	0.003
Internet chatting	▪ Yes	0.557	0.178	0.001
Arts activities, hobbies	▪ Yes	0.713	0.149	0.000
Correct classification	▪ Overall	66.8%		
	▪ True positive	66.1%		
	▪ True negative	71.3%		
Nagelkerke R^2		9.5%		

* *Significant if $p < 0.05$.*

Odds ratio

Independent variables		Reference category	Odds ratio for reading	Inverse odds ratio[14]
Level of education	▪ College	**High school or less**	2.271	0.44
	▪ University		4.586	0.21
Gender	▪ Woman	**Man**	2.944	0.34
Language community	▪ Majority French	**Minority**	1.474	0.67
	▪ Majority English		1.818	0.55
Not skipping over unfamiliar words	▪ Yes	**No**	1.508	0.66
Internet chatting	▪ Yes	**No**	1.745	0.57
Arts activities, hobbies	▪ Yes	**No**	2.041	0.48

[14] *Odds of reading in the reference category divided by the odds of reading in the other categories of the same variable.*

3.2 TOTAL CANADA – REGULAR READER RATE

1. VARIABLES IN THE MODEL

- A review of the non-weighted data from the Canadian sample series made possible an initial identification of the variables with little or no relation to the regular reading rate.

  → We compared the frequency distribution of the independent variables for regular readers and those who do not read on a regular basis. The series of independent variables likely to influence the regular reading rate was thus reduced to 14 variables (see **Appendix** for the list of this sub-series).

- Next, we proceeded with a multiple correlation analysis between the sub-series of 14 independent variables and the dependent variable, and nine other variables were eliminated from the model at this stage because of collinearity.

- Thus, the logistic regression analysis for predicting regular reading among Canadian respondents as a whole was carried out using the five following independent variables:

  1) Level of education
  2) Gender
  3) Language community
  4) Not skipping over unfamiliar words
  5) Engages in arts activities or hobbies

- As in the case with general reading of books for pleasure, age, income and region did not have enough of a specific link with regular reading to be included in the model.

- Contrary to what was observed for the general reading rate, Internet chatting does not have enough of a link with regular reading, and so this variable was not selected.

2. CODING OF VARIABLES

- The dependent variable was coded “0” for those who do not read on a regular basis and “1” for regular readers.

- The categories of independent variables were regrouped and recoded according to their regular reader rates and the resultant categories are presented in **Table 2a** (below). The category whose value is “0” represents the reference category for each variable in the odds ratio analysis (all the other categories are compared to the reference category through the logistic regression analysis).

TABLE 2A **VARIABLES SELECTED TO PREDICT THE REGULAR READING RATE IN CANADA**

Independent variables	Categories
▪ *Level of education*	• **[High school or less]** • College • University +
▪ *Gender*	• **[Man]** • Woman
▪ *Language community*	• **[Minority French]** • Majority French • Minority English • Majority English
▪ *Not skipping over unfamiliar words*	• **[Disagree]** • Agree
▪ *Engages in arts activities or hobbies*	• **[No]** • Yes

*N.B. The reference category is in **[bold]**.*

3. INTERPRETATION OF RESULTS

- **Table 2b** presents the partial correlation coefficients of each of the variables and the categories of each variable. It measures the correlation between the independent variable and the dependent variable when the effect of the other variables included in the model is taken into account. Table 2b also ranks each variable in the predictive model.

  → Thus, gender (R=0.177) has the most influence on the regular reading rate in Canada, followed by:

  – Education level (R=0.135);
  – Language community (R=0.088);
  – Engages in arts activities or hobbies (R=0.084);
  – Not skipping over unfamiliar words (R=0.037).

- Again, the official language majority or minority situation contributes relatively little (R=0.088) to the model, which suggests that the low regular reading rate among Francophones outside Quebec, in relation to the other language communities, can be explained to a great extent by other variables, such as education, for example.

TABLE 2B **PARTIAL CORRELATIONS OF PREDICTOR VARIABLES FOR REGULAR READING OF BOOKS FOR PLEASURE – TOTAL CANADA (N=1 963) –**

Independent variables	Variables		Categories	
	R*	Rank	R*	Rank
Level of education	0.135	2		
▪ College			0.065	5
▪ University			0.136	2
Gender	0.177	1		
▪ Woman			0.177	1
Language community	0.088	3		
▪ Majority Francophones			---	---
▪ Minority Anglophones			0.037	6
▪ Minority Anglophones			0.076	4
Never skips over unfamiliar words	0.037	5		
▪ Agree			0.037	6
Engages in arts activities or hobbies	0.084	4		
▪ Yes			0.084	3

* *Partial correlation coefficient (R) which measures the level of correlation between the independent variable and the dependent variable when the effect of the other variables included in the model is taken into account. All the partial correlation coefficients indicated are statistically significant at P < 0.01, except for those in the categories "Minority Anglophones" and "Never skips over unfamiliar words", which are statistically significant at P < 0.05.*

- In the bottom part of **Table 2c**, we see that the model correctly classifies a high proportion of respondents, namely 64.1% of the cases observed. More specifically, the model manages to produce 62.0% true positives (classified by the model as regular readers when they are in actual fact regular readers) and 66.7% true negatives (classified by the model as not regular readers when they are in actual fact not regular readers).

- **Table 2c** also presents the regression coefficients (β) and the odds ratios of each of the categories of predictor variables for regular reading in Canada. The regression coefficient reflects the net effect (while controlling for the effect of the other variables) of each predictor variable on the logarithm for the odds ratio that a person regularly reads for pleasure in Canada.

- The odds ratio[15] reflects the increase in the odds of regularly reading associated with membership in a particular category rather than in the reference category. Thus, the odds ratio of reading regularly among respondents with a university education (β = 0.966) are two and a half times[16] higher than among respondents with a high school education or less.

[15] *Obtained by calculating the inverse of the natural logarithm of the coefficient β.*

[16] *Remember: the odds (P/1-P) do not directly reflect the probability of reading.*

- The highest odds ratios were observed for the following categories:

  → University education (2.6 times higher than among respondents with high school or less);

  → Women (2.4 times higher than among men);

  → Majority Anglophones (1.8 times higher than among minority Francophones);

  → Engages in arts activities or hobbies (1.7 times higher than among respondents who do not);

  → Minority Anglophones (1.6 times higher than among minority Francophones).

- As is the case with the reading rate, the two variables that most influence the regular readers rate are the same in Canada and in the US, namely level of education and gender.

  → Respondents with a higher level of education read more often than those with a lower level of education.

    – In Canada, those with a college education have 1.4 times higher odds of reading regularly than those with a high school education or less, and those with a university education have 2.6 times higher odds.

    – In the United States, respondents having completed at least a few years of college have 1.2 times higher odds of reading regularly than those who never made it past high school.

  → More women read regularly than men.

    – In Canada, their odds of reading regularly are 2.4 times higher than those of men.

    – In the US, this ratio is 1.6 times higher.

- The other explanatory variables in the model differed from one country to the other, reflecting their different geo-demographic realities but also because the questions were not always comparable.

  → Canada:

    – Majority or minority official language situation;
    – Not skipping over unfamiliar words;
    – Engages in arts activities or hobbies.

→ United States:

- Race and ethnicity;
- Number of hours watching television;
- Age;
- Employment activity;
- Mother's level of education;
- Volunteer work;
- Attends artistic performances.

- All the same, this confirms that, contrary to the United States (where young people read less than older people), age does not influence Canadians' regular reading rate.

- An analysis of the odds ratio shows us that:

1. The largest gap is between levels of education. Those who have not completed post-secondary studies have odds of reading regularly equivalent to only 38 per cent of those of respondents who have completed their university studies.

2. There is also a large gap between men and women. The odds of reading regularly among men are 40 per cent of those among women. This gap is larger than the one between minority Francophones and majority Anglophones (the odds of reading regularly among the former are 54 per cent of those among the latter).

- These results are very similar to those for the general reading rate and underscore the importance of seeking first and foremost to improve the pleasure book reading rate among men and those with the lowest education levels before targeting other groups. They also validate the idea that the Internet does not have a negative effect on reading.

TABLE 2C **SUMMARY OF LOGISTIC REGRESSION FOR PREDICTING REGULAR READING OF BOOKS FOR PLEASURE – TOTAL CANADA (N=1 963) –**

Independent variables		Regression coefficient β	Standard error	p*
	▪ Constant	-1.654	0.179	0.000
Level of education	▪ College	0.398	0.107	0.000
	▪ University	0.966	0.133	0.000
Gender	▪ Woman	0.912	0.097	0.000
Language community	▪ Majority Francophones	0.115	0.171	0.502
	▪ Minority Anglophones	0.482	0.198	0.015
	▪ Majority Anglophones	0.615	0.146	0.000
Not skipping over unfamiliar words	▪ Agree	0.231	0.096	0.017
Engages in arts activities or hobbies	▪ Yes	0.532	0.115	0.000
Correct classification	▪ Overall	64.1%		
	▪ True positive	62.0%		
	▪ True negative	66.7%		
Nagelkerke R^2		10.7%		

* *Significant if p < 0.05*

Odds ratio

Independent variables		Reference category	Odds ratio for reading	Inverse odds ratio[17]
Level of education	▪ College	**High school or less**	1.489	0.67
	▪ University		2.628	0.38
Gender	▪ Woman	**Man**	2.489	0.40
Language community	▪ Majority Francophones	**Minority Francophones**	1.122	0.89
	▪ Minority Anglophones		1.620	0.61
	▪ Majority Anglophones		1.850	0.54
Not skipping over unfamiliar words	▪ Agree	**Disagree**	1.260	0.79
Engages in arts activities or hobbies	▪ Yes	**No**	1.703	0.58

[17] *Odds of reading regularly in the reference category divided by the odds of reading in the other categories of the same variable.*

3.3 FRANCOPHONES IN QUEBEC – GENERAL READER RATE

1. VARIABLES IN THE MODEL

- Following a review of the data from the sample of Francophones in Quebec and an analysis of frequency distributions and correlations, 15 variables were not considered in the Logit analysis (see **Appendix** for the list of excluded and included variables) because they did not have any influence on the general reading rate or due to their collinearity with other more influential variables.

  → In addition, three variables did not apply to this sub-group (official language, majority or minority linguistic status, and region).

- Six variables were therefore retained for the logistic regression analysis:

  1) Education level
  2) Gender
  3) Resides near a public library
  4) Does volunteer work
  5) Employment activity
  6) Does not skip over unfamiliar words

- Note that two of those six variables (employment activity and does not skip over unfamiliar words) were not significant in the six-variable Logit model and were therefore eliminated. We are therefore presenting the results of a four-variable model.

2. CODING OF VARIABLES

- The dependent variable was coded “0” for non-readers and “1” for readers.

- Independent variable categories were grouped together and recoded according to their reader rates and the resulting categories are presented in **Table 3a**. The category with a value of “0” represents the reference category for each variable in the odds ratio analysis (all the other categories are compared to the reference category through logistic regression analysis).

TABLE 3A **VARIABLES USED TO PREDICT GENERAL READING RATES AMONG FRANCOPHONES IN QUEBEC**

Independent Variables	Categories
▪ *Education level*	• **[High school or less]** • College • University
▪ *Gender*	• **[Man]** • Woman
▪ *Resides near a public library*	• **[Disagree]** • Agree
▪ *Employment activity*	• **[Unemployed]** • Employed • Student
▪ *Does not skip over unfamiliar words*	• **[Disagree]** • Agree
▪ *Does volunteer work*	• **[No]** • Yes

*N.B. The reference category is in **[bold]**.*

3. INTERPRETATION OF RESULTS

- **Table 3b** presents the partial correlation coefficients for each variable and for the categories of each variable. It represents the correlation between the independent variable and the dependent variable when the effect of the other variables in the model is taken into consideration. Table 3b also shows where each variable ranks in the prediction model.

- Three variables had a very comparable influence on the general reading rate among Francophones in Quebec:

  → access to a public library near their place of residence (R=0.150);
  → involvement in volunteer work (R=0.133); and
  → gender (R=0.129).
  → Education level (R=0.098) ranked fourth in terms of influence.

TABLE 3B **PARTIAL CORRELATIONS FOR PREDICTOR VARIABLES FOR READING BOOKS FOR PLEASURE – FRANCOPHONES IN QUEBEC (N=360) –**

Independent Variables	Variables		Categories	
	R*	Rank	R	Rank
Education level	0.098	4		
▪ College			---	---
▪ University			0.115	4
Gender	0.129	3		
▪ Woman			0.129	3
Resides near a public library	0.150	1		
▪ Agree			0.150	1
Does volunteer work	0.133	2		
▪ Yes			0.133	2

* *The partial correlation coefficient (R) measures the level of correlation between the independent variable and the dependent variable when the effect of the other variables in the model is taken into consideration. All the partial correlation coefficients listed are statistically significant at P < 0.01 except for the "education level" category, which is statistically significant at P < 0 .05.*

- The bottom of **Table 3c** shows that the model correctly classifies a high proportion of respondents–67.5%. More specifically, the model produces a true positive rate of 67.1% (classified by the model as readers when they are, in actual fact, readers) and a true negative rate of 69.8% (classified by the model as non-readers when they are, in actual fact, non-readers).

- **Table 3c** also presents the regression coefficients (β) and the odds ratios for each of the categories of predictor variables for the general reading rate among Francophones in Quebec. The regression coefficient reflects the unique effect (by controlling for the effect of the other variables) of each predictor variable on the logarithm of the odds ratio that a Francophone in Quebec reads for pleasure.

- The highest odds ratios were observed for:

  → University education (6.4 times higher than among respondents with a high school education or less);

  → Does volunteer work (2.9 times higher than among those who do not do volunteer work);

  → Resides near a public library (2.5 times higher than among respondents who do not);

  → Women (2.3 times higher than among men).

- The evaluation of the odds ratios shows us that:

    1. the greatest difference is between education categories. Among Francophones in Quebec, the odds of reading for those who have completed a university education are 6.4 times higher than for those who have not completed a post-secondary education. In other words, the odds of reading among those who did not complete post-secondary studies are equal to 15% of the odds among those who have completed a university education.

    2. the difference between men and women, while significant, is much smaller than the difference between levels of education. Thus, the odds of men reading are only 42% of the odds of women.

TABLE 3C **SUMMARY OF LOGISTIC REGRESSION TO PREDICT THE READING OF BOOKS FOR PLEASURE – FRANCOPHONES IN QUEBEC (N=360) –**

Independent Variables		Regression Coefficient β	Standard Error	p*
	▪ Constant	0.155	0.310	0.616
Education level	▪ College	0.452	0.342	0.186
	▪ University	1.860	0.757	0.014
Gender	▪ Woman	0.846	0.318	0.007
Resides near a public library	▪ Agree	0.939	0.317	0.003
Does volunteer work	▪ Yes	1.077	0.396	0.006
Correct classification	▪ Overall	67.5%		
	▪ True positive	67.1%		
	▪ True negative	69.8%		
Nagelkerke R^2		17.8%		

* *Significant if p < 0.05.*

Odds Ratios

Independent Variables		Reference Category	Odds Ratio for Reading[18]	Inverse Odds Ratio[19]
Education level	▪ College	**High school or**	1.572	0.63
	▪ University	**less**	6.429	0.15
Gender	▪ Woman	**Man**	2.332	0.42
Resides near a public library	▪ Agree	**Disagree**	2.559	0.39
Does volunteer work	▪ Yes	**No**	2.938	0.34

[18] *The odds ratio, which is obtained by calculating the inverse of the natural logarithm of the coefficient β, reflects an increase in the odds associated with membership in a particular category rather than in the reference category. Francophones in Quebec who have a university education (β = 1.860 and odds ratio of 6.429) have an odds ratio for reading that is six and a half times (6.429) higher than among their counterparts who have a high school education or less. Remember that the odds (P/1-P) are not a direct reflection of the probability of reading, but rather the number of readers divided by the number of non-readers.*

[19] *Odds (P/1-P) of reading for the reference category divided by odds of reading for the other categories of the same variable.*

3.4 FRANCOPHONES IN QUEBEC– REGULAR READER RATE

1. VARIABLES IN THE MODEL

- Following a review of the data from the sample of Francophones in Quebec and an analysis of frequency distributions and correlations, 18 variables were not considered in the Logit analysis (see **Appendix** for the list of excluded and included variables) because they did not have any influence on regular reader rates or due to their collinearity with other more influential variables.

  → Three variables did not apply to this sub-group (official language, majority or minority language status, and region).

- Ultimately, 6 variables were retained for the logistic regression analysis:

  1) Gender
  2) Visible minority
  3) Family income
  4) Resides near a public library
  5) Does not skip over unfamiliar words
  6) Attends arts performances

2. CODING OF VARIABLES

- The dependent variable was coded “0” for those who do not read regularly and “1” for those who do read regularly.

- Independent variable categories were grouped together and recoded according to their regular reader rates and the resulting categories are presented in **Table 4a**. The category with a value of “0” represents the reference category for each variable in the odds ratio analysis (all the other categories are compared to the reference category through logistic regression analysis).

TABLE 4A **VARIABLES USED TO PREDICT REGULAR READING RATES AMONG FRANCOPHONES IN QUEBEC**

Independent Variables	Categories
▪ *Gender*	• **[Man]** • Woman
▪ *Visible minority*	• **[Yes]** • No
▪ *Family income*	• **[$20K -]** • $20K +
▪ *Resides near a public library*	• **[Disagree]** • Agree
▪ *Does not skip over unfamiliar words*	• **[Disagree]** • Agree
▪ *Attends arts performances*	• **[No]** • Yes

*N.B. The reference category is in **[bold]**.*

3. INTERPRETATION OF RESULTS

- **Table 4b** presents the partial correlation coefficients for each variable and for the categories of each variable. It represents the correlation between the independent variable and the dependent variable when the effect of the other variables in the model is taken into consideration. Table 4b also shows where each variable ranks in the prediction model.

- Gender (R= 0.196) is the variable that has the most influence on the regular reading rate among Francophones in Quebec, followed by access to a public library near their place of residence (R=0.120). Income (R=0.081) and attending arts performances (R=0.071) also have a certain influence.

TABLE 4B **PARTIAL CORRELATIONS FOR PREDICTOR VARIABLES FOR REGULAR BOOK READING FOR PLEASURE – FRANCOPHONES IN QUEBEC (N=360) –**

Independent Variables	Variables		Categories	
	R*	Rank	R*	Rank
Gender ▪ Woman	0.196	1	0.196	1
Visible minority ▪ No	---	---	---	---
Income ▪ $20K +	0.081	3	0.081	3
Resides near a public library ▪ Agree	0.120	2	0.120	2
Never skips over unfamiliar words ▪ Agree	---	---	---	---
Attends arts performances ▪ Yes	0.071	4	0.071	4

* *The partial correlation coefficient (R) measures the level of correlation between the independent variable and the dependent variable when the effect of the other variables in the model is taken into consideration. All the partial correlation coefficients listed are statistically significant at P < 0.01 except for the categories "income $20K +" and "attends arts performances," which are significant at P < 0.05.*

- The bottom of **Table 4c** shows that the model correctly classifies a remarkable proportion of respondents – 64.7%. More specifically, the model produces a true positive rate of 65.2% (classified by the model as regular readers when they are, in actual fact, regular readers) and a true negative rate of 64.3% (classified by the model as regular non-readers when they are, in actual fact, regular non-readers).

- **Table 4c** also presents the regression coefficients (β) and the odds ratios for each of the categories of predictor variables for regular reading among Francophones in Quebec.

- The highest odds ratios were observed for the following categories:

  → Women (3 times higher than among men);

  → Not a visible minority (2.8 times higher than among visible minorities);

  → Resides near a public library (2.1 times higher than among others);

  → Family income greater than $20 000 (1.8 times higher than among those whose family income is less than $20 000).

TABLE 4c **SUMMARY OF LOGISTIC REGRESSION TO PREDICT REGULAR BOOK READING FOR PLEASURE – FRANCOPHONES IN QUEBEC (N=360) –**

Independent Variables		Coefficient of Regression β	Standard Error	p*
	▪ Constant	-3.245	0.734	0.000
Gender	▪ Woman	1.092	0.238	0.000
Visible minority	▪ No	1.056	0.678	0.029
Family income	▪ $20 000 and +	0.593	0.256	0.021
Near a public library	▪ Agree	0.751	0.247	0.002
Never skips over unfamiliar words	▪ Agree	0.291	0.234	0.213
Attends arts performances	▪ Yes	0.511	0.240	0.033
Correct classification	▪ Overall ▪ True positive ▪ True negative	64.7% 65.2% 64.3%		
Nagelkerke R^2		18.4%		

* *Significant if p < 0.05.*

Odds Ratios

Independent Variables		Reference Category	Odds Ratio for Reading Regularly[20]	Inverse Odds Ratio[21]
Gender	▪ Woman	**Man**	2.982	0.33
Visible minority	▪ No	**Yes**	2.874	0.34
Family income	▪ $20K and +	**$20K -**	1.809	0.55
Near a public library	▪ Agree	**Disagree**	2.121	0.47
Never skips over unfamiliar words	▪ Agree	**Disagree**	1.338	0.74
Attends arts performances	▪ Yes	**No**	1.667	0.59

[20] *The odds ratio, which is obtained by calculating the inverse of the natural logarithm of coefficient β, reflects an increase in the odds associated with membership in a particular category rather than in the reference category. Francophone women in Quebec (β = 1.092 and odds ratio of 2.982) have an odds ratio for reading regularly that is three times higher than among men. Remember that the odds (P/1-P) are not a direct reflection of the probability of reading, but rather the number of regular readers divided by the number of regular non-readers.*

[21] *Odds (P/1-P) of reading for the reference category divided by odds of reading for the other categories of the same variable.*

3.5 FRANCOPHONES OUTSIDE QUEBEC – GENERAL READER RATE

1. VARIABLES IN THE MODEL

- Following a review of the data from the sample of Francophones outside Quebec and an analysis of frequency distributions and correlations, 19 of 24 variables were not considered in the Logit analysis (see **Appendix** for the list of excluded and included variables) because they did not have any influence on reader rates or due to their collinearity with other more influential variables.

  → Ultimately, two variables could not be applied to this sub-group (official language and majority or minority linguistic status).

- Therefore, there were only three variables left to be included in the logistic regression analysis:

  1) Family income
  2) Gender
  3) Never skips over unfamiliar words.

2. CODING OF VARIABLES

- The dependent variable was coded “0” for non-readers and “1” for readers.

- Independent variable categories were grouped together and recoded according to their reader rates and the resulting categories are presented in **Table 5a**. The category with a value of “0” represents the reference category for each variable in the odds ratio analysis (all the other categories are compared to the reference category through logistic regression analysis).

TABLE 5A VARIABLES USED TO PREDICT GENERAL READER RATES AMONG FRANCOPHONES OUTSIDE QUEBEC

Independent Variables	Categories
▪ *Family income*	• **[$20 000 -]** • $20 000 - $40 000 • $40 000 +
▪ *Gender*	• **[Man]** • Woman
▪ *Never skips over unfamiliar words*	• **[Disagree]** • Agree

*N.B. The reference category is in **[bold]**.*

3. INTERPRETATION OF RESULTS

- **Table 5b** presents the partial correlation coefficients for each variable and for the categories of each variable.

- Family income (R=0.234) has the most influence on reading rates among Francophones outside Quebec, followed by gender (R=0.137) and the tendency not to skip over unfamiliar words (R=0.133).

TABLE 5B PARTIAL CORRELATIONS FOR PREDICTOR VARIABLES FOR BOOK READING FOR PLEASURE – FRANCOPHONES OUTSIDE QUEBEC (N=265) –

Independent Variables	Variables		Categories	
	R*	Rank	R*	Rank
Family income	0.234	1		
▪ $20 000 – $40 000			0.091	4
▪ $40 000 and +			0.247	1
Gender	0.137	2		
▪ Woman			0.137	2
Never skips over unfamiliar words	0.132	3		
▪ Agree			0.132	3

* *The partial correlation coefficient (R) measures the level of correlation between the independent variable and the dependent variable when the effect of the other variables in the model is taken into consideration. All the partial correlation coefficients listed are statistically significant at P < 0.01 except for the income " $20 000 - $40 000" category, which is statistically significant at P < 0.05.*

- The bottom of **Table 5c** shows that the model correctly classifies a very high proportion of respondents – 77%. More specifically, the model produces a true positive rate of 81.3% (classified by the model as readers when they are, in actual fact, readers) and a true negative rate of 60.7% (classified by the model as non-readers when they are, in actual fact, non-readers).

- **Table 5c** also presents the regression coefficients (β) and the odds ratios for each of the categories of predictor variables for general reading among Francophones outside Quebec.

- The highest odds ratios were observed for the following categories:

  → Family income of $40 000 or more (5.5 times higher than among respondents with a family income of less than $20 000);

  → Does not skip over unfamiliar words (2.4 times higher than among those who do);

  → Women (2.4 times higher than among men);

  → Family income between $20 000 and $40 000 (2.3 times higher than among respondents with a family income of less than $20 000).

TABLE 5C **SUMMARY OF LOGISTIC REGRESSION TO PREDICT BOOK READING FOR PLEASURE – FRANCOPHONES OUTSIDE QUEBEC (N=265) –**

Independent Variables		Regression Coefficient β	Standard Error	p*
	▪ Constant	-0.334	0.328	0.308
Family income	▪ $20 000 -$40 000 ▪ $40 000 and +	0.840 1.716	0.406 0.396	0.038 0.000
Gender	▪ Woman	0.894	0.333	0.007
Does not skip over unfamiliar words	▪ Agree	0.897	0.343	0.009
Correct classification	▪ Overall ▪ True positives ▪ True negatives	77.0% 81.3% 60.7%		
Nagelkerke R^2		19.9%		

* *Significant if p < 0.05.*

Odds Ratios

Independent Variables		Reference Category	Odds Ratio for Reading[22]	Inverse Odds Ratio[23]
Family income	▪ $20 000 - $40 000 ▪ $40 000 and +	**$20K -**	2.317 5.563	0.43 0.17
Gender	▪ Woman	**Man**	2.445	0.40
Never skips over unfamiliar words	▪ Agree	**Disagree**	2.452	0.40

[22] *The odds ratio, which is obtained by calculating the inverse of the natural logarithm of coefficient β, reflects an increase in odds associated with membership in a particular category rather than in the reference category. Francophone respondents outside Quebec with a family income of $40 00 or more (β = 1.716 and odds ratio of 5.563) have an odds ratio for reading that is five times higher than among respondents with a family income of less than $20 000. Remember that the odds (P/1-P) are not a direct reflection of the probability of reading, but rather the number of readers divided by the number of non-readers.*

[23] *Odds (P/1-P) of reading for the reference category divided by odds of reading for the other categories of the same variable.*

3.6 FRANCOPHONES OUTSIDE QUEBEC – REGULAR READER RATE

1. VARIABLES IN THE MODEL

- Following a review of the data from the sample of Francophones in Quebec and an analysis of frequency distributions and correlations, 18 variables were not considered in the Logit analysis (see **Appendix** for the list of excluded and included variables) because they did not have any influence on regular reader rates or due to their collinearity with other more influential variables.

  → Also, two variables could not be applied to this restricted sub-group (official language and majority or minority linguistic status).

- Ultimately, four variables were retained for the logistic regression analysis:

  1) Education level
  2) Gender
  3) Never skips over unfamiliar words
  4) Goes to the movies

2. CODING OF VARIABLES

- The dependent variable was coded “0” for those who do not read regularly and “1” for those who do.

- Independent variable categories were grouped together and recoded according to their regular reader rates and the resulting categories are presented in **Table 6a**. The category with a value of “0” represents the reference category for each variable in the odds ratio analysis (all the other categories are compared to the reference category through logistic regression analysis).

TABLE 6A VARIABLES USED TO PREDICT REGULAR READING RATES AMONG FRANCOPHONES OUTSIDE QUEBEC

Independent Variables	Categories
▪ *Education level*	• **[High school or less]** • Some university • University +
▪ *Gender*	• **[Man]** • Woman
▪ *Never skips over unfamiliar words*	• **[Disagree]** • Agree
▪ *Goes to the movies*	• **[No]** • Yes

*N.B. The reference category is in **[bold]**.*

3. INTERPRETATION OF RESULTS

- **Table 6b** presents the partial correlation coefficients for each variable and for the categories of each variable and the rank of each variable in the prediction model.

- Education level (R=0.217) has the most influence on regular reading rates among Francophones outside Quebec, followed by gender (R=0.118) and watching movies at the theatre (R=0.106).

TABLE 6B PARTIAL CORRELATIONS FOR PREDICTOR VARIABLES FOR REGULAR BOOK READING FOR PLEASURE – FRANCOPHONES OUTSIDE QUEBEC (N=265) –

Independent Variables	Variables		Categories	
	R*	Rank	R*	Rank
Education level	0.217	1		
▪ Some university			---	
▪ University			0.229	1
Gender	0.118	2		
▪ Woman			0.118	2
Never skips over unfamiliar words	---	---		
▪ Agree			---	---
Goes to the movies	0.106	3		
▪ Yes			0.106	3

* *The partial correlation coefficient (R) measures the level of correlation between the independent variable and the dependent variable when the effect of the other variables in the model is taken into consideration. All the partial correlation coefficients listed are statistically significant at P < 0.01 except for the "goes to the movies" category, which is statistically significant at P < 0.05.*

- The bottom of **Table 6c** shows that the model correctly classifies a high proportion of respondents – 67.6%. More specifically, the model produces a fairly low percentage of true positives, 56.3% (classified by the model as regular readers when they are, in actual fact, regular readers), but a high percentage of true negatives, 75.8% (classified by the model as non-readers when they are, in actual fact, non-readers).

- **Table 6c** also presents the regression coefficients (β) and the odds ratios for each of the categories of predictor variables for regular reading among Francophones outside Quebec.

- The highest odds ratios were observed for the following categories:

  → University education (7.2 times higher than among respondents with a high school education);

  → Women (2.1 times higher than among men);

  → Watching movies at the theatre (1.9 times higher than among respondents who do not go to the movies).

TABLE 6C **SUMMARY OF LOGISTIC REGRESSION TO PREDICT REGULAR BOOK READING FOR PLEASURE – FRANCOPHONES OUTSIDE QUEBEC (N=265) –**

Independent Variables		Regression Coefficient β	Standard Error	p*
	▪ Constant	-1.789	0.330	0.000
Education level	▪ Some university	0.381	0.310	0.218
	▪ University	1.980	0.432	0.000
Gender	▪ Woman	0.769	0.289	0.007
Never skips over unfamiliar words	▪ Agree	0.461	0.275	0.093
Goes to the movies	▪ Yes	0.682	0.277	0.013
Correct classification	▪ Overall	67.6%		
	▪ True positives	56.3%		
	▪ True negatives	75.8%		
Nagelkerke R^2		22.5%		

* *Significant if p < 0.05.*

Odds Ratios

Independent Variables		Reference Category	Odds Ratio for Reading[24]	Inverse Odds Ratio[25]
Education level	▪ Some university	**High school or less**	1.464	0.68
	▪ University		7.247	0.13
Gender	▪ Woman	**Man**	2.158	0.46
Never skips over unfamiliar words	▪ Agree	**Disagree**	1.586	0.63
Goes to the movies	▪ Yes	**No**	1.977	0.50

[24] *The odds ratio, which is obtained by calculating the inverse of the natural logarithm of coefficient β, reflects an increase in the odds associated with membership in a particular category rather than in the reference category. Francophones outside Quebec with a university education (β = 1.980 and odds ratio of 7.247) have an odds ratio for reading regularly that is seven times higher than among respondents with a high school education or less. Remember that the odds (P/1-P) are not a direct reflection of the probability of reading regularly, but rather the number of readers divided by the number of non-readers.*

[25] *Odds (P/1-P) of reading for the reference category divided by odds of reading for the other categories under the same variable.*

APPENDIX

INCLUDED/EXCLUDED VARIABLES

TABLE 1 **EXCLUDED/INCLUDED VARIABLES FOR THE LOGIT MODEL**
TOTAL CANADA – GENERAL READER RATE

TOTAL CANADA READERS	Reference Category	LOGIT Model Included Variables
▪ Official language	French	
▪ Age	35 years +	
▪ Type of household	---	
▪ Education level	High school	Included
▪ Employment activity	---	
▪ Visible minority	---	
▪ Access to the Internet	No	
▪ Family income	$20K -	
▪ Gender	Man	Included
▪ CMA	No	
▪ Region	QC/Atl.	
▪ Language community	Minority	Included
▪ Resides near a bookstore (Q4.1.14)	Disagree	
▪ Resides near a public library (Q4.1.15)	Disagree	
▪ Resides near a bookstore with French-language books (Q4.1.17)	---	
▪ Resides near a public library with French-language books (Q4.1.18)	---	
▪ Never skips over unfamiliar words (Q5.1)	Disagree	Included
▪ Parents read (Q6.1)	Never / almost never	
▪ TV shows or videos (Q2.3)	---	
▪ Listens to music	---	
▪ Arts activities or hobbies	No	Included
▪ Does volunteer work	No	
▪ Chats on-line	No	Included
▪ Goes to the movies	No	
▪ Museums and art galleries	No	
▪ Arts performances	No	

--- : *No reference category because no trends were observed, even after response categories were grouped together.*

TABLE 2 **EXCLUDED/INCLUDED VARIABLES FOR THE LOGIT MODEL**
TOTAL CANADA – REGULAR READERS

TOTAL CANADA REGULAR READERS	Reference Category	LOGIT Model Included Variables
▪ Official language	---	
▪ Age	---	
▪ Type of household	---	
▪ Education level	High school -	Included
▪ Employment activity	---	
▪ Visible minority	---	
▪ Access to the Internet	No	
▪ Family income	$20K -	
▪ Gender	Man	Included
▪ CMA	---	
▪ Region	---	
▪ Language community	Fr. min.	Included
▪ Resides near a bookstore (Q4.1.14)	Disagree	
▪ Resides near public library (Q4.1.15)	Disagree	
▪ Resides near a bookstore with French-language books (Q4.1.17)	---	
▪ Resides near a public library with French-language books (Q4.1.18)	---	
▪ Never skips over unfamiliar words (Q5.1)	Disagree	Included
▪ Parents read (Q6.1)	Never / almost never	
▪ TV shows or videos (Q2.3)	---	
▪ Listens to music	---	
▪ Arts activities or hobbies	No	Included
▪ Does volunteer work	No	
▪ Chats on-line	---	
▪ Goes to the movies	No	
▪ Museum and art galleries	No	
▪ Arts performances	No	

--- : *No reference category because no trends were observed, even after response categories were grouped together.*

TABLE 3 **EXCLUDED/INCLUDED VARIABLES FOR THE LOGIT MODEL**
FRANCOPHONES IN QUEBEC –GENERAL READER RATE

FRANCOPHONES IN QUEBEC READERS	Reference Category	LOGIT Model Included Variables
▪ Official language	N/A	
▪ Age	16-24 years of age	
▪ Type of household	Couple without children	
▪ Education level	High school -	Included
▪ Employment activity	Unemployed	
▪ Visible minority	---	
▪ Access to the Internet	---	
▪ Family income	$20K -	
▪ Gender	Man	Included
▪ CMA	No	
▪ Region	N/A	
▪ Language community	N/A	
▪ Resides near a bookstore (Q4.1.14)	Disagree	
▪ Resides near a public library (Q4.1.15)	Disagree	Included
▪ Resides near a bookstore with French-language books (Q4.1.17)	N/A	
▪ Resides near a public library with French-language books (Q4.1.18)	N/A	
▪ Never skips over unfamiliar words (Q5.1)	Disagree	
▪ Parents read (Q6.1)	Never	
▪ TV shows or videos (Q2.3)	---	
▪ Listens to music	---	
▪ Arts activities or hobbies	No	
▪ Does volunteer work	No	Included
▪ Chats on-line	---	
▪ Goes to the movies	---	
▪ Museums and art galleries	No	
▪ Arts performances	No	

N/A: Variable that does not apply to this sub-group.

--- : No reference category because no trends were observed, even after response categories were grouped together.

TABLE 4 **EXCLUDED/INCLUDED VARIABLES FOR THE LOGIT MODEL**
FRANCOPHONES IN QUEBEC – REGULAR READERS

FRANCOPHONES IN QUEBEC REGULAR READERS	Reference Category	LOGIT Model Included Variables
▪ Official language	N/A	
▪ Age	16-34 years	
▪ Type of household	---	
▪ Education level	High school -	
▪ Employment activity	Employed	
▪ Visible minority	Yes	Included
▪ Access to the Internet	---	
▪ Family income	$20K -	Included
▪ Gender	Man	Included
▪ CMA	No	
▪ Region	N/A	
▪ Language community	N/A	
▪ Resides near a bookstore (Q4.1.14)	Disagree	
▪ Resides near a public library (Q4.1.15)	Disagree	Included
▪ Resides near a bookstore with French-language books (Q4.1.17)	N/A	
▪ Resides near a public library with French-language books (Q4.1.18)	N/A	
▪ Never skips over unfamiliar words (Q5.1)	Disagree	Included
▪ Parents read (Q6.1)	---	
▪ TV shows or videos (Q2.3)	---	
▪ Listens to music	---	
▪ Arts activities or hobbies	---	
▪ Does volunteer work	---	
▪ Chats on-line	No	
▪ Goes to the movies	---	
▪ Museums and art galleries	No	
▪ Arts performances	No	Included

N/A: Variable that does not apply to this sub-group.

--- : No reference category because no trends were observed, even after response categories were grouped together.

TABLE 5 **EXCLUDED/INCLUDED VARIABLES FOR THE LOGIT MODEL**
FRANCOPHONES OUTSIDE QUEBEC – GENERAL READER RATE

FRANCOPHONES OUTSIDE QUEBEC READERS	Reference Category	LOGIT Model Included Variables
▪ Official language	N/A	
▪ Age	55 years +	
▪ Type of household	---	
▪ Education level	High school -	
▪ Employment activity	Unemployed	
▪ Visible minority	---	
▪ Access to the Internet	No	
▪ Family income	$20K -	Included
▪ Gender	Man	Included
▪ CMA	No	
▪ Region	---	
▪ Language community	N/A	
▪ Resides near a bookstore (Q4.1.14)	---	
▪ Resides near a public library (Q4.1.15)	---	
▪ Resides near a bookstore with French-language books (Q4.1.17)	---	
▪ Resides near a public library with French-language books (Q4.1.18)	---	
▪ Never skips over unfamiliar words (Q5.1)	Disagree	Included
▪ Parents read (Q6.1)	Never / almost never	
▪ TV shows or videos (Q2.3)	---	
▪ Listens to music	---	
▪ Arts activities or hobbies	No	
▪ Does volunteer work	No	
▪ Chats on-line	---	
▪ Goes to the movies	No	
▪ Museums and art galleries	No	
▪ Arts performances	No	

N/A: Variable that does not apply to this sub-group.

--- : No reference category because no trends were observed, even after response categories were grouped together.

TABLE 6 **EXCLUDED/INCLUDED VARIABLES FOR THE LOGIT MODEL**
FRANCOPHONES OUTSIDE QUEBEC – REGULAR READERS

FRANCOPHONES OUTSIDE QUEBEC REGULAR READERS	Reference Category	LOGIT Model Included Variables
▪ Official language	N/A	
▪ Age	---	
▪ Type of household	No children	
▪ Education level	High school -	Included
▪ Employment activity	---	
▪ Visible minority	---	
▪ Access to the Internet	---	
▪ Family income	$20K$ -	
▪ Gender	Man	Included
▪ CMA	---	
▪ Region	Ontario	
▪ Language community	N/A	
▪ Resides near a bookstore (Q4.1.14)	---	
▪ Resides near a public library (Q4.1.15)	Disagree	
▪ Resides near a bookstore with French-language books (Q4.1.17)	---	
▪ Resides near a public library with French-language books (Q4.1.18)	Disagree	
▪ Never skips over unfamiliar words (Q5.1)	Disagree	Included
▪ Parents read (Q6.1)	Never / almost never	
▪ TV shows or videos (Q2.3)	---	
▪ Listens to music	---	
▪ Arts activities or hobbies	No	
▪ Does volunteer work	No	
▪ Chats on-line	---	
▪ Goes to the movies	No	Included
▪ Museums and art galleries	No	
▪ Arts performances	No	

N/A: Variable that does not apply to this sub-group.

--- : No reference category because no trends were observed, even after response categories were grouped together.